Modeling Digital Shadows in Manufacturing Using Process Mining

Tobias Brockhoff[1], Merih Seran Uysal[2], Wil M.P. van der Aalst[3]

Abstract: Friction in shopfloor-level manufacturing processes often occurs at the intersection of different subprocesses (e. g., joining sub-parts). Therefore, considering the Digital Shadows (DSs) of individual materials/sub-parts is not sufficient when analyzing the processes. To this end, holistic views on shopfloor-level processes that integrate multiple DSs are needed. In this work, we discuss how material-centric DSs supported by discrete assembly events can be integrated using techniques from process mining. In particular, we propose to utilize DSs that contain additional structural information to overcome the main challenges of concurrency and the presence of many different objects.

Keywords: Digital Shadow; Process Mining; Bill of Materials; Manufacturing Process Discovery

1 Introduction

With the advent of Industry 4.0, Digital Shadows (DSs) become increasingly important for decision-making in production. At the same time, companies collect increasing amounts of data on their operational processes. Despite the importance and feasibility in terms of data availability, realizing compatible DSs, which can be integrated and linked to create new insights, often remains difficult. To facilitate the use and create a common foundation, a DS meta model has been proposed in [Be21] (see Fig. 1). Still, implementations can become very and specific, and, therefore, realizing DSs remains difficult. For example, predicting waste for a machine can require to consider and fuse many different sensors and models.

In our research, we focus on general-purpose DSs that exist on the shopfloor-level of manufacturing processes—namely, *DSs of assembly executions* and *DSs of sub-components/materials composition as well as production line plans*. In particular, we consider assembly execution DSs built on discrete event data (e. g., assembly activity events). To gain insights into the production process, the DS on the assembly execution and the structural material composition for multiple products need to be combined. To this end, we use techniques from Process Mining (PM) which is an emerging discipline that leverages event data to improve processes.

In PM, there are three major concepts: *events*, *cases*, and *process models*. *Events* are recordings of discrete business operations and their time of occurrence. Multiple events

[1] RWTH Aachen University brockhoff@pads.rwth-aachen.de
[2] RWTH Aachen University uysal@pads.rwth-aachen.de
[3] RWTH Aachen University wvdaalst@pads.rwth-aachen.de

(e. g., events related to a product) constitute a *case*. Finally, *process models* describe the behavior of a process. As depicted in Fig. 1, these concepts perfectly align with the meta model proposed in [Be21]. However, from the PM perspective, there are two main challenges

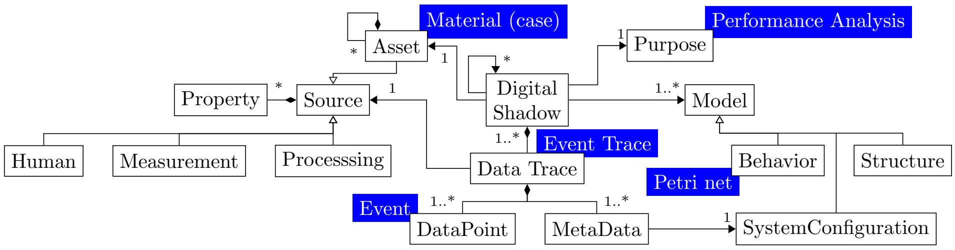

Fig. 1: Alignment of process mining concepts (blue) with the entities of the Digital Shadow (DS) concept model proposed in [Be21] (black).

for integrating the DSs: *concurrency* and *object-centricity* [Aa21]. Sub-materials can be assembled concurrently, and, when analyzing the performance of shopfloor-level processes, we can consider either the sub-materials or the product as case notions. When selecting the product as the case, the events related to its sub-materials may arbitrarily interleave; while, when selecting the materials as the case notion, friction at the intersection of materials remains unnoticed. In our research, we investigate bridging the gap towards a comprehensive production model by leveraging additional sources of information—namely, structural DSs on material composition (e. g., bill of materials) or the assembly line (e. g., assembly line plans). We thereby address the following research questions:

RQ1 How can dynamic DSs of process executions be combined with structural DSs of material composition and production lines?

RQ2 How can we create a performance-aware DS of a production line that reveals friction particularly at the intersection of subassembly boundaries?

While the first question focuses on modeling production lines, the second question targets the enhancement of this model such that it realizes the *purpose* of performance monitoring.

2 Related Work

Digital Shadows and Twins in manufacturing are increasingly gaining attention. Key applications are production planning (e. g., by simulation), control, and optimization [Kr18]. In this regard, one approach to describe shopfloor-level activities is by means of discrete events (e. g., events for starting or completing assembly activities). Discrete event simulation can then be applied to plan and optimize the process [KA16; YAL16]. While the required accurate process models would be valuable for a data-driven performance analysis, such models often do not exist. Moreover, tuning the parameters (e. g., service times) to obtain

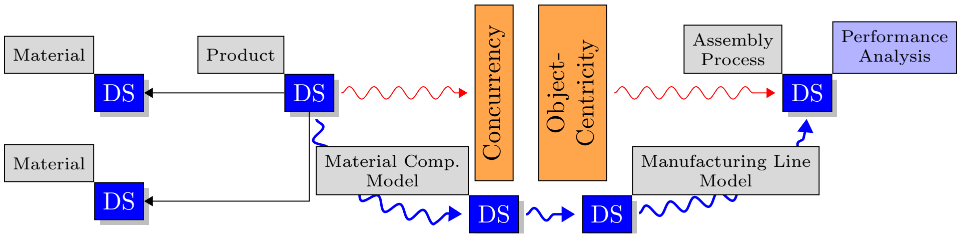

Fig. 2: Exploiting additional structural information to overcome the challenges of concurrency and object-centricity. We propose (blue snake line) to structure and integrate Digital Shadows (blue) for different assets (gray) by exploiting DSs on structural production information.

precise models is often tedious. In our research, we focus on the reverse direction, that is, given event and additional structural data, a model of the manufacturing process is built. Using this model, we generate unbiased performance assessments based on *real data*.

Within the last decade, there have been numerous works on the application of PM for manufacturing processes. Overviews over the potential of PM to improve manufacturing processes and how PM has already been applied can be found in [Aa20; DRG21]. Its focus on the actual dynamics distinguishes PM from classical process mapping that shows statically aggregated data [Lo21]. For example, it can show differences between the designed work flow and the as-is production [Lo21]. However, purely event data-driven model discovery becomes infeasible for large processes. Consequently, these works resort to process-global statistics such as the number of activities in progress or cumulative delay [Pa15]. Process models are only used within limited process scopes. To analyze processes based on comprehensive models, we exploit additional manufacturing-specific information.

3 Methods

In our research, we strive to create comprehensive views on shopfloor-level processes. To this end, we leverage event data generated on the shopfloor as well as additional structural information. Conceptually, we create a new DS that combines event data-based DSs of the process executions with DSs that contain structural material models (e. g., Multi-level Manufacturing Bills of Materials (M^{2}BOMs)). Techniques from PM thereby help to *integrate highly dynamic event data* and rather *rigid structural models*. In particular, using concepts from process mining and information from the structural DSs, we first discover a behavioral model—i. e., a process model—of the manufacturing process. Afterwards, we use PM to enhance the discovered model by performance information creating a performance-aware DS of the production. An illustration of our approach is depicted in Fig. 2. We start from DSs of material and subpart assembly execution that are built on and instantiated by discrete event data (i. e., events of assembly activities). Gaining insights into the process then requires to integrate the obtained DSs. However, the integration faces two major challenges:

concurrency and *object-centricity* [Aa21]. In our research, we exploit DSs that contain additional structural information to disentangle shopfloor-level manufacturing processes. Such models can either be manufacturing line models that disentangle the concurrency of assembly lines or material composition models. Due to physical constraints, it can usually be assumed that the data conforms to the model (e. g., products cannot skip stations at the conveyor belt). For example, in [Uy20], we modeled a car manufacturing process that consists of a general assembly line where some stations depend on concurrent sub-assembly lines. We used the model to replay the event data to compute KPIs (e. g., waiting or idle times) and visualized the evolution over time. While direct modeling is feasible for highly structured processes, it quickly becomes infeasible if the product flexibility increases. In this case, information on the material composition can help to disentangle and, eventually, model the process.

The resulting DS can be visualized, thereby, enabling a backward-looking analysis of the process that can reveal systematic production problems. Moreover, it can be used to query specific production KPIs.

Exploiting Structural Domain Information A common type of material composition information are M^2BOMs. M^2BOMs organize the materials built into a product in trees. Vertices correspond to materials whose assembly depend on the assembly of their child materials. The root vertex is the final product. Depending on the product, M^2BOMs contain hundreds of materials. For such large processes, in particular when subparts are assembled concurrently, automatic model discovery usually fails to find understandable models. Compared to the underlying highly-structured M^2BOM, the models are either unstructured and 'spaghetti' or overly general. This problem is worsened by products having similar but not necessarily equal M^2BOMs (e. g., certain materials are optional, or there might be a choice between different configurations). In [Br21], we investigated how M^2BOMs can be exploited to comprehensively model manufacturing processes for the purpose of performance analysis. We start with a collection of M^2BOMs and, targeting RQ1, output a M^2BOM-like process model. The latter is a tree that contains all materials from the input collection as well as optional materials, material choices, and additional material groupings (e. g., a choice between two material groups). Moreover, each material is endowed with an assembly task vertex that subsumes all activities related to its direct assembly. An example output for the offset printer manufacturing process introduced in [Br21] is depicted in Fig. 3 which shows all occurring (anonymized) materials as well as optional materials and choices between materials. The performance-aware coloring shows process-global bottlenecks as well as differences between similar materials.

Conceptually, we obtain the model by merging M^2BOMs. Based on counting arguments, we automatically identify shared materials and potential choices. The latter are then resolved manually as the resolution can be ambiguous. For example, consider two infrequent features that never occur together. It does not per se clear that these features are mutually exclusive. Since the so-obtained model has a direct correspondence to a process model (i. e., a process

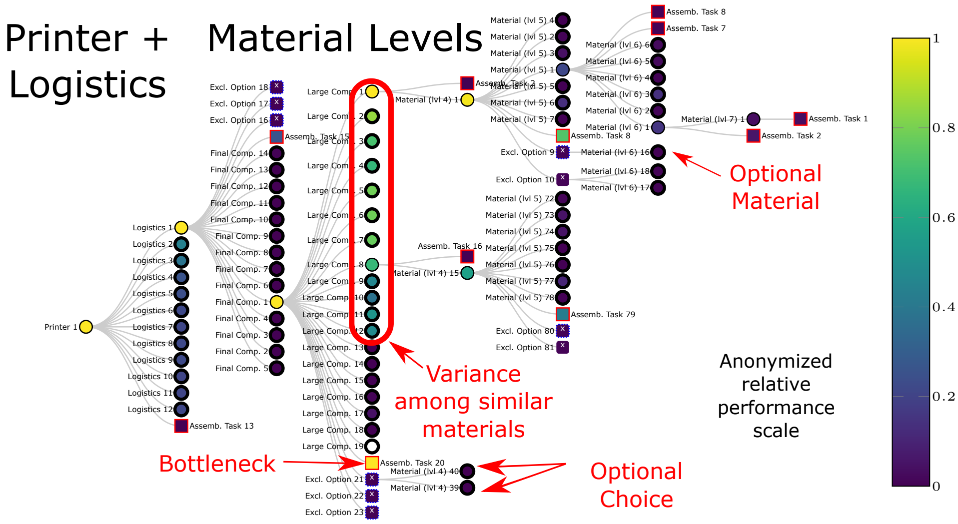

Fig. 3: Overall offset printer manufacturing process model with collapsible vertices. Each vertex corresponds to a material. Special vertices are dedicated to optional materials (blue dashed), choices (×), and assembly tasks (red vertices). The color depicts the cumulative material assembly time.

tree), it can be endowed with performance metrics derived from replaying the event data. Considering RQ2, this enables to use the model to detect performance problems. Since the model comprehensively integrates all materials, it can also be used to compare similar materials across the model as well as to analyze relations between parent and child materials.

4 Challenges and Conclusion

In this work, we presented our research on realizing performance-aware Digital Shadows (DSs) of shopfloor-level manufacturing processes. To this end, we propose to complement techniques from Process Mining (PM) by additional structural data to alleviate the challenges of concurrency and object-centricity. In doing so, we can visualize processes even if standard automatic model discovery fails. In future work, we aim to generalize our work to other sources of structural information. Considering the performance analysis, a major challenge lies in integrating additional process context into the model. While, in process model notations commonly used in PM, different orders are independent, this does not hold in real life. We therefore require models that capture the process context. Moreover, our current work only enables a backward-looking analysis. Even though this is sufficient to yield insights into systematic problems, it does not allow to react to and recover from real-time problems. The latter requires to continuously update the model turning it into a Digital Twin of the assembly line. While techniques from PM facilitate integrating dynamic performance updates with respect to the event data, structural updates can become more challenging.

Acknowledgment We thank the Alexander von Humboldt (AvH) Stiftung for supporting our research. Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC- 2023 Internet of Production – 390621612.

References

[Aa20] van der Aalst, W. M. P. et al.: Removing operational friction using process mining: challenges provided by the internet of production (IoP). In: DATA. Springer, pp. 1–31, 2020.

[Aa21] van der Aalst, W. M. P.: Concurrency and Objects Matter! Disentangling the Fabric of Real Operational Processes to Create Digital Twins. In: ICTAC 2021. Springer International Publishing, pp. 3–17, 2021.

[Be21] Becker, F. et al.: A Conceptual Model for Digital Shadows in Industry and Its Application. In: ER. Springer International Publishing, pp. 271–281, 2021.

[Br21] Brockhoff, T. et al.: Analyzing Multi-level BOM-structured Event Data. In: Process Mining Workshops. Vol. 433, 2021.

[DRG21] Dreher, S.; Reimann, P.; Gröger, C.: Application Fields and Research Gaps of Process Mining in Manufacturing Companies. In: LNI. Gesellschaft für Informatik, Bonn, pp. 621–634, 2021.

[KA16] Kaylani, H.; Atieh, A. M.: Simulation Approach to Enhance Production Scheduling Procedures at a Pharmaceutical Company with Large Product Mix. Procedia CIRP 41/, pp. 411–416, 2016, ISSN: 2212-8271.

[Kr18] Kritzinger, W. et al.: Digital Twin in manufacturing: A categorical literature review and classification. IFAC-PapersOnLine 51/, pp. 1016–1022, 2018.

[Lo21] Lorenz, R. et al.: Using Process Mining to Improve Productivity in Make-to-Stock Manufacturing. IJPR/, pp. 1–12, 2021, ISSN: 0020-7543.

[Pa15] Park, M. et al.: Workload and Delay Analysis in Manufacturing Process Using Process Mining. In: Asia Pacific BPM. Springer International Publishing, pp. 138–151, 2015.

[Uy20] Uysal, M. S. et al.: Process Mining for Production Processes in the Automotive Industry. In: Industry Forum at BPM, Sevilla, Spain. 2020.

[YAL16] Yang, S.; Arndt, T.; Lanza, G.: A Flexible Simulation Support for Production Planning and Control in Small and Medium Enterprises. Procedia CIRP 56/, pp. 389–394, 2016, ISSN: 2212-8271.