Process Mining with Applications to Automotive Industry

Michael Siek*, Ryan Malik Gunadharma Mukti
Bina Nusantara University Jakarta, Indonesia 11480

*michael.s@binus.edu

Abstract. Process mining as a modeling and analysis tool can be used to improve the business performance by looking at the actual business processes. This paper presents the applications of process mining in automotive industry. Using event log data with timestamps, process mining algorithms, like inductive miner and fuzzy miner were able to automatically generate car manufacturing processes, automatically checking the conformance between the actual processes and the predefined standard ones, and identify and solve any bottlenecks and issues in the car manufacturing processes. A number of car manufacturing issues were considered in this research, such as process delays, stagnant workflow, mismanagement, faulty production and labor insufficiency. The modeling and statistical results show promising leverage of process mining in automotive industry that can lead to the autonomous car manufacturing with abilities for real-time process auditing and reengineering.

Keywords: automated process model generation, real-time process auditing and reengineering, inductive miner, fuzzy miner, identification of process bottleneck

1. Introduction

Mining a business process or process mining often leads to new discoveries and insights into the real processes occurred in many businesses[1]. These findings could explain some detailed reasons why the businesses run stagnant or without any progress. Analyzes of these findings and insights are crucial in order to give some benefits to most of the enterprises, such as detecting the process bottleneck and how to solve this issue [2]. Some bottleneck or interferences in the processes may disrupt the whole of business performance.

As an analysis tool used in process management, a process mining utilizes simulation methods of playing out or playing in of events, which usually occur in many companies’ processes and determine what to do in tackling any process problems [3]. The process mining requires event log data as inputs for generating business process model and its further objects for analysis. The event logs are sets of event or process data with timestamps to indicate when a process starts and ends [4]. Some example of events or processes in automotive industry, such as: casting, stamping, engine assembling and quality controlling [5].

The objective of this research is to apply process mining algorithms for automotive industry in the purposes of automated generation of car manufacturing process and automated conformance checking of the actual manufacturing process leading to process auditing and reengineering. The main principles of car manufacturing in Toyota were chosen and utilized for creating production scenarios and event logs, testing process mining algorithms and their resulting process models. There are several scenarios included in the case of Toyota Manufacturing, which visage a number of certain hindrances causing delays and stagnant workflow in the car manufacturing processes. Some hindrances include...
faulty engine manufacture, prolonged wait time of quality control inspection, mismanagement of certain parts that require attention, and faulty production overall [6].

2. Experimental method

2.1. Process mining methodologies

Despite having a widely-used Business Process Model (BPM), we employ Petri net as process model representation as it is more universal [7]. A Petri net is able to visualize how events are transition to one another based on the event logs. Some reliable tools like ProM and RapidProM were utilized for mining manufacturing processes from event log data and checking the conformance of the actual processes with the standard manufacturing processes in Toyota [8]. Two algorithms for process mining like inductive miner and fuzzy miner [9-11] were exploited to obtain a consistent process model results from event logs. In fuzzy model, a rough diagram of process can be generated from event logs with the frequency information of each process transition for producing every individual car. In other words, it allows us to verify how often a particular process travels to the other process[12]. In fuzzy model diagram, the frequency information is depicted by the arrows between events or processes; the thicker the arrow the more frequent the process leading to the next process.

2.2. Event log data collection

Event log data is utilized as the inputs for process mining algorithms to produce car manufacturing processes represented by Petri net. These event logs are specific database employed to identify and check on how the businesses operates and find in case of the company has some unused or unnecessary process [13]. Though most companies may not have event logs, however they have time stamped records of events that are part of the company’s business processes. For instance, the events can be taken from the meetings held, reports made, items manufactured, and so forth [14].

Table 1. Fragment of synthetic event log data in automotive industry.

| Scenario ID | Case ID | Activity               | StartTimestamp | EndTimestamp | Cost | Resources |
|-------------|--------|------------------------|----------------|-------------|------|-----------|
| 1           | 311656 | Casting                | 2019-05-07 07:45 | 2019-05-07 09:25 | 100  | Mike      |
| 1           | 311657 | Engine production      | 2019-05-07 09:30 | 2019-05-07 11:30 | 200  | Pietro    |
| 1           | 311658 | Stamping               | 2019-05-07 07:45 | 2019-05-07 08:55 | 50   | John      |
| 1           | 311659 | Welding                | 2019-05-07 09:00 | 2019-05-07 10:00 | 100  | John      |
| 1           | 311660 | Painting               | 2019-05-07 10:05 | 2019-05-07 11:00 | 50   | James     |
| 1           | 311661 | Assembly               | 2019-05-07 13:15 | 2019-05-07 15:45 | 200  | Charles   |
| 1           | 311662 | Quality Control        | 2019-05-07 16:00 | 2019-05-07 18:00 | 100  | Charles   |
| 1           | 311663 | Shipping and delivery  | 2019-05-07 18:05 | 2019-05-07 19:00 | 50   | Susan     |
| 2           | 311665 | Engine production      | 2019-05-08 08:00 | 2019-05-08 09:45 | 200  | Pietro    |
| 2           | 311666 | Stamping               | 2019-05-08 08:00 | 2019-05-08 09:00 | 50   | John      |
| 2           | 311667 | Welding                | 2019-05-08 09:05 | 2019-05-08 09:50 | 100  | John      |
| 2           | 311668 | Painting               | 2019-05-08 09:50 | 2019-05-08 10:30 | 50   | James     |
| 2           | 311669 | Assembly               | 2019-05-08 10:35 | 2019-05-08 11:58 | 200  | Charles   |
| 2           | 311670 | Quality Control        | 2019-05-08 13:00 | 2019-05-08 14:00 | 100  | Mike      |
| 2           | 311671 | Shipping and delivery  | 2019-05-08 15:00 | 2019-05-08 15:55 | 50   | Susan     |

For data collection in this research, we investigated the car manufacturing processes in Toyota, specifically on how the company manufactures its cars, how many cars is possibly produced, how often something can go wrong in each process, and how the issues can be resolved. Subsequently, we created various types of scenarios in the event logs covering all possible standard and non-standard
procedures of car manufacturing that could occur within a certain time frame for one production line, with production scenarios with production issues. Using this event log data, we conducted some experiments and testing for automated generation of car manufacturing processes and checking the conformance or compliant of the actual car manufacturing processes with the standard car manufacturing processes [15].

A fragment of synthetic event log data utilized for process mining in automotive industry is listed in Table 1. This data consists of four main attributes: case ID, activity, starting timestamp and ending timestamp. The other attributes are used to create the scenarios for various types of car manufacturing processes. For instance, scenario ID 1 illustrates the “perfect” process scenario from the process of casting until shipping and delivery of the car. In other case, scenario ID 2 indicates an event which the manufacturing plant chooses to re-use an asset rather than casting a new engine. In scenario ID 3, we consider the manufacturing processes and packing of spare parts. Lastly, scenario ID 4 is a scenario in which a certain car manufacturing requires to build a new engine after the first production is considered faulty.

3. Results and discussion

3.1. Automated generation of manufacturing processes

Two process mining algorithms: inductive miner and fuzzy miner, were performed to automatically generate car manufacturing processes from event log data as partly listed in Table 1.

3.1.1. Inductive miner. The automated generation of car manufacturing processes using inductive miner from event log data is depicted as Petri net in Figure 1. The Petri net diagram produced by inductive miner provides us with how the manufacturing processes flow with some information of the time duration and frequency for each process.

Figure 1. Automated generation of automotive manufacturing processes using process mining induction represented as Petri net.

The green icon on the left indicates the start of the car manufacturing process, while the red icon indicates the end of the car manufacturing process. What comes after the green icon is a ‘+’ icon, this icon indicates that the process is split and will work in parallel with each other. For this case, it can be seen from Figure 1 that the processes that are working in parallel is casting, stamping, and engine production. It is also shown that after casting is completed; the process continues to engine production; while stamping ends with its own route. Once they finish, the process ends back to a new ‘+’ icon indicating that the parallel sequences come together and follow by further process to be executed. Some subsequent processes include welding, painting, assembly, quality control, and end process. These manufacturing processes look simple since we hide some process paths for better understanding of the generated car manufacturing processes from the used event log data.

In addition to Petri net diagram automatically generated by inductive miner from event log data, we can check some anomalies in the car manufacturing processes as illustrated in Figure 2. It is a dotted chart for the event logs indicated by index traces versus timestamps of the events or processes with the information of each process’s time duration and frequency. Using this chart, we can find the manufacturing process issues and analyze the abnormal processes in comparison to the standard processes in car manufacturing. In more detailed example, we can obviously see that on May 17th the
production line appears to take longer than the rest of the production scenarios. With closer inspection, we can look that after conducting a quality control, there is a need to recast a part for the car, therefore insinuating that the previous casting caused an error in the car production.

![Event: index in trace versus time: timestamp depicting some variations on time duration and frequency of the car production processes.](image)

**Figure 2.** Event: index in trace versus time: timestamp depicting some variations on time duration and frequency of the car production processes.

Another important dotted chart is the one indicated by event name versus time stamps as illustrated in Figure 3. This chart allows us to check how often a particular activity in car manufacturing need to be performed in regards to all scenarios. From this dotted chart, we can see the list of events that have occurred in both the process and event log, with the timestamp placed as the handle to separate the events based on its frequency and occurrence. It becomes apparent in this dotted chart that the least frequent event that executed in the process is ‘Packing and Vanning’, occurring only 3 times within the entire processes.

3.1.2. **Fuzzy miner.** The other process mining algorithm utilized in this research is fuzzy miner. In similar case to inductive miner algorithm, we used event log data to automatically generate fuzzy process model as depicted in Figure 4. In the fuzzy process model, the frequency information of a particular process is indicated by an arrow connecting between events or processes; the thicker the arrow the more frequent the process leading to the next process.
It can be inferred that the car manufacturing processes can continue progressively. How often a certain process follows the standard predefined process indicates that there is high conformance between the actual and standard car manufacturing procedures. The current process does not need to impact any drastic changes since the processes indicated here still follow the standard ones. In other aspect, there are some arrow connections in the fuzzy model diagram that loop back onto itself. The most likely reason is due to the timestamps and the sequence of activities overlapped in the event log data. In case of some processes do not continue to any other process, the fuzzy model creates a loop in which it ends at its own process once it completes.

Table 2. Statistical evaluation measures to indicate the alignments.

| Measure                      | Average | Max. | Min.  | Std. Deviation | #Case Value of 1.00 |
|------------------------------|---------|------|-------|----------------|---------------------|
| Raw Fitness Cost             | 2.75    | 5    | 0     | 2.22           | 0                   |
| Calculation Time (ms)        | 2,378.50| 2,677| 2,031 | 274.59         | 0                   |
| Num. States                  | 2,771.00| 4,813| 1,244 | 1,531.30       | 0                   |
| Move-Model Fitness           | 0.97    | 1    | 0.92  | 0.04           | 2                   |
| Trace Fitness                | 0.94    | 1    | 0.9   | 0.04           | 1                   |
| Move-Log Fitness             | 0.97    | 1    | 0.93  | 0.03           | 1                   |
| Trace Length                 | 42      | 48   | 30    | 8.49           | 0                   |
| Queued States                | 7,806.50| 9,574| 6,516 | 1,416.13       | 0                   |

3.2. Automated conformance checking of manufacturing processes

For conformance checking, we utilize some statistical evaluation measures to indicate the process alignments, such as: raw fitness cost, move-model fitness, trace fitness and move-log fitness. Table 2 lists all of these statistical evaluation measure values that indicate how the actual processes to the standard car manufacturing processes conform. The move-model fitness of 0.97, trace fitness of 0.94 and move-log fitness 0.97 indicate reliable conformance or alignment, whereas the raw fitness cost of
2.75 did not provide good indicator of conformance. The latter is due to the fact that the number of cases for every scenario in event log data is insufficient for the process mining.

![Automotive manufacturing processes generated by fuzzy-based process mining from event log data.](image)

**Figure 4.** Automotive manufacturing processes generated by fuzzy-based process mining from event log data.

There are also some other statistics to measure the speed of process model generation, number of states, trace length and queued states. It can be inferred from these alignment statistics that the generated Petri net took time on average of 2.3 milliseconds to create with an average trace length of 42. As the alignment statistic shows, there are 7,806 queued states on average, which explains why the the Petri net diagram sometimes looks scrambled and messy. Figure 5 shows the conformance checking of automotive manufacturing processes generated from event log data, represented as Petri net.
Figure 5. Conformance checking of automotive manufacturing processes from event log data.
4. Conclusion

The applications of process mining in automotive industry are quite straightforward. There are at least three applications: automated generation of car manufacturing processes, automated conformance checking of the actual car manufacturing processes to the standard ones, and to identify and solve any bottleneck in the manufacturing processes. The leverages of process mining in automotive industry can lead to the autonomous manufacturing with abilities for real-time process auditing and reengineering. A number of car manufacturing issues such as process delays and stagnant workflow due to mismanagement, faulty production and labor scarcity can be immediately detected by the process mining system. This indicates process mining can be a reliable tool towards the automated car manufacturing technology in the near future.

Acknowledgment

This research has been conducted and funded through a research grant program supported by Bina Nusantara University with a number of 033/VR.RTT/IV/2019.

References

[1] W. M. van der Aalst, H. A. Reijers, A. J. Weijters, B. F. van Dongen, A. A. De Medeiros, M. Song, et al., "Business process mining: An industrial application," Information Systems, vol. 32, pp. 713-732, 2007.
[2] W. M. van der Aalst, "Process mining: Discovery, conformance and enhancement of business processes," ed: Springer-Verlag Berlin, 2011.
[3] S. Park and Y. S. Kang, "A study of process mining-based business process innovation," Procedia Computer Science, vol. 91, pp. 734-743, 2016.
[4] W. M. van der Aalst, Process Mining. Data science in action: Springer-Verlag Berlin, 2016.
[5] Toyota. (2019). How we manufacture. Available: http://www.toyotaindonesiamanufacturing.co.id/manufacturing/how-we-manufacture?lang=id
[6] A. Valencia-Parra, B. Ramos-Gutiérrez, A. J. Varela-Vaca, M. T. Gómez-López, and A. G. Bernal, "Enabling Process Mining in Aircraft Manufactures: Extracting Event Logs and Discovering Processes from Complex Data," in Proceedings of the Industry Forum at BPM 2019, 2019, pp. 166-177.
[7] S. J. Leemans, E. Poppe, and M. T. Wynn, "Directly follows-based process mining: Exploration & a case study," in 2019 International Conference on Process Mining (ICPM), 2019, pp. 25-32.
[8] W. M. van der Aalst, A. Bolt, and S. J. van Zelst, "RapidProM: mine your processes and not just your data," arXiv preprint arXiv:1703.03740, 2017.
[9] D. P. Solomatine and M. B. Siek, "Modular learning models in forecasting natural phenomena," Neural networks, vol. 19, pp. 215-224, 2006.
[10] R. Mans, W. M. van der Aalst, and H. Verbeek, "Supporting Process Mining Workflows with RapidProM," in BPM (Demos), 2014, p. 56.
[11] S. M. Tedjojuwono, "Virtual Lines Sensors for Moving Object and Vehicle Counters," IAENG International Journal of Computer Science, vol. 44, 2017.
[12] D. Chapela-Campa, M. Mucientes, and M. Lama, "Mining frequent patterns in process models," Information Sciences, vol. 472, pp. 235-257, 2019.
[13] E. González López De Murillas, H. A. Reijers, and W. M. Aalst, "Connecting databases with process mining: a meta model and toolset," Software and Systems Modeling (SoSyM), vol. 18, pp. 1209-1247, 2019.
[14] S. van Zelst, "Process mining with streaming data," Technische Universiteit Eindhoven, 2019.
[15] Á. J. Varela-Vaca, J. A. Galindo, B. Ramos-Gutiérrez, M. T. Gómez-López, and D. Benavides, "Process mining to unleash variability management: discovering configuration workflows using logs," in Proceedings of the 23rd International Systems and Software Product Line Conference-Volume A, 2019, pp. 265-276.