An Introduction to MPM - MEHRWERK ProcessMining

Janna Meyer
Mehrwerk AG
Karlsruhe
Germany
janna.meyer@mehrwerk-ag.de

Josua Reimold
Mehrwerk AG
Karlsruhe
Germany
josua.reimold@mehrwerk-ag.de

Constantin Wehmschulte
Mehrwerk AG
Karlsruhe
Germany
constantin.wehmschulte@mehrwerk-ag.de

***Abstract*—Qlik Sense® platform based MEHRWERK ProcessMining is the first tool to offer self-service process mining including the data governance necessary for enterprise-wide use. Fast implementation through various data extraction options and integrability with machine learning functions make MPM a customizable tool for every use case. Basic process discovery and conformance checking algorithms in combination with a powerful business intelligence platform allow in-depth analysis as flexible as the investigation requirements themselves.**

keywords: self-service process mining, process discovery, conformance checking, variants analysis, log extraction

1. Introduction

As process mining is a highly explorative analysis technique, process mining software should provide interactive and visual analysis to support insights in complex processes. MEHRWERK ProcessMining (MPM), deployed on the Qlik Sense® platform, is designed to offer comprehensive analytics to the process analyst. MPM combines market leading self-service business intelligence (self-service BI) with the insights achieved by process mining algorithms and visualizations for process discovery and conformance checking and, thus, becomes a powerful tool for professional process analysis. Through Qlik Sense®'s APIs it is possible to enhance the solution according to the given use case's requirements, for example, by integrating data mining scripts from R or Python. To introduce MPM properly to the interested reader we will firstly discuss the innovation our software offers to the process mining universe by shortly explaining the platform and the implemented functions. Then a description of the tool's maturity is presented to outline its usability and its adaptability to future developments. Future developments are shortly mentioned to elaborate our vision and in the concluding part of the paper is a link to a 25 minute introduction video.

mds
April 05, 2019

2. Discussion of Innovation

MPM is characterized by the combination of interactive self-service BI with lean process mining algorithms. As flexibility and usability regarding analysis capabilities are key to successful generation of process insights, employing Qlik Sense® for process mining purposes is convincing. Even non-experts find answers rapidly in the appealing visualizations of MPM due to easily understandable associative analysis.

2.1. Qlik Sense® Platform

The Qlik Sense® platform provides user-friendly analysis based on governed data discovery. With the Qlik Associative Engine it is possible to discover relations in even vast amounts of data by technically performing a many-to-many full outer join in the back-end [1]. This approach enables powerful and interactive search, selection and filter functions due to a patented in-memory technology including compressed binary indexing, logical deduction and dynamic KPI calculation and aggregation [1]. Recognizing the importance of such capabilities for process analysis, the enhancement of Qlik Sense® with process mining capabilities to create an innovative tool is the logical conclusion. As a result, analysts can select any dimension and every visualization, diagram, table or metric is calculated at run-time with respect to the choosen analysis perspective and data set. A colouring supports the comprehension of relations within the data as the selected data is coloured green, the excluded dark and the associated data light grey. The benefit for process mining is obvious: for example, to analyze the context of a given case or the influence of a resource's involvement on process delay. Another advantage provided by Qlik is creating ad-hoc analysis in a governed environment with validated KPIs and dimensions without the need of coding. With modern drag and drop functionality process analysts are able to generate new visualizations within minutes. Therefore, MPM is more than a simple reporting tool on processes. By defining and sharing bookmarks and ad-hoc reports analysts communicate their insights to colleagues. Furthermore, they can export the visualizations and reports to present them for example to the management. To support

mobile analytics, Qlik Sense®'s apps offer a responsive design that can be used seamlessly on iOS or Android. Offline analysis is provided likewise.

As business analytics is dependent on up-to-date and high-quality data, Qlik Sense® offers the import of various data formats as well as connectors to a large number of enterprise information systems. Official connectors can be found at [2]. Profiting from this connectivity, MPM allows fast, tolerant and scheduled event log extraction from source systems and data processing in a scalable manner.

2.2. Log Extraction and Enrichment

MPM provides two options of feeding event logs into the process mining algorithm. The first is to run the code on an existing event log that is read into Qlik Sense® and transformed into a minimal log with information about case ID, activity type and activity timestamps such as start or end of the event. The second option is to extract data directly from sources, as databases, XML, xlsx or csv with the MPM RuleEngine and create the event log. Doing this, there is no need to provide a process mining-ready event log beforehand and the import effort is significantly reduced. Advantageous in this *data-directly from source scenario* is the opportunity to easily alter the scope. If time periods or activities of interest change, a few clicks lead to a new event log that is automatically analyzed by our process discovery algorithm. Also, due to the MPM RuleEngine the event log generation is implicitly and comprehensively documented. Through these distinct initiation possibilities, MPM has a short implementation time since it adapts to the use case and the existing data architecture.

To offer the best insights in process data, MPM connects the process log to context data by Qlik-typical data modeling. Omitted KPIs are, if needed, fastly calculated and appended to the data model, often without the need to go back to the data extraction step. Thus, the enrichment of event logs, process logs or data models for e.g. root cause analysis purposes is rapidly accomplished given some knowledge in Qlik Sense® formular syntax.

2.3. Functionality

Through MPM, the complex task of integrating process mining with modern business intelligence is achieved. By defining reusable template applications the analyst is enabled to perform exactly the same analysis on top of different event logs – some refer to this as *history mode*. For example, the comparison of different time periods before and after a certain process intervention is simply achieved. Since performance and quality analysis, compliance checking as well as process monitoring are important areas of process mining applications, the opportunities provided by the integration with BI that are offered by MPM are outstanding.

2.3.1. Process Discovery. Process Discovery in MPM is performed by fuzzy mining the event log and extracting the resulting directly-follows graph. Hence, the deterministic MPM Process Discovery Algorithm is rather simple but efficient. The MPM ProcessAnalyzer shows the real processes as-is and enhances them by standard metrics like process step duration, lead or idle times and further use case-dependent performance indicators such as automation rates. Filter and selection options help reducing the spaghetti-diagram to relevant process variants. Through Qlik's Associative Engine the whole application content is then recalculated for the selected data. The MPM QueryBuilder provides the useful capability to search for activity patterns of concern thus, allowing to fastly check critical process variants and their context information.

2.3.2. Conformance Checking. Another helpful functionality is conformance checking. With the MPM ProcessModeler the user is able to define a happy path via drag and drop which is displayed in the process visualization to investigate process deviations. If a should-be process model is defined, its list of distinct process variants can be used to calculate the process conformance between reality and model. The MPM ConformanceChecking Algorithm encompasses alignment based logic and, therefore, evaluates activities and moves that are either synchronous or in model or log. With these information a fitness metric for the most fitting happy path and each real-life process variant is calculated. Hence, the analyst can rapidly identify highly deviating process variants due to deviating activities or changed activity ordering and compares process variants that should, but, do not correspond to one common happy path.

2.3.3. Process Analysis Perspective. Given the analytic flexibility of Qlik Sense®, analysts can use MPM to evaluate processes from different perspectives. At first, the control flow perspective can be taken to get an overview on the general process by generating the process visualization and calculating the process variants. The ordering of activities and the deviation from a happy path get clear in the visualization and can be further assessed by a variety of diagrams, KPIs and the alignment based conformance checking. With the process instance as central investigation object, the case perspective is the easiest perspective to take with MPM. Along with resource data, the use case dependent context, e.g. supplier, material, machine or weather, can be included in the analysis. To cover the organizational or rather resource perspective the analyst can employ a social network graph, an extension provided by an open source extension on Qlik Branch® Garden [3]. To investigate bottlenecks, resource utilization or service degree, the time perspective is also supported by MPM. Qlik's standard analysis provides the required functionalities when used on the event log timestamps and the process indicators, which were calculated by the process discovery algorithm.

Figure 1 demonstrates the seamless interaction of the Qlik selection (green, white, light and dark grey), the MPM ProcessAnalyzer, and the MPM ConformanceChecking. The left process visualization exhibits the most common process

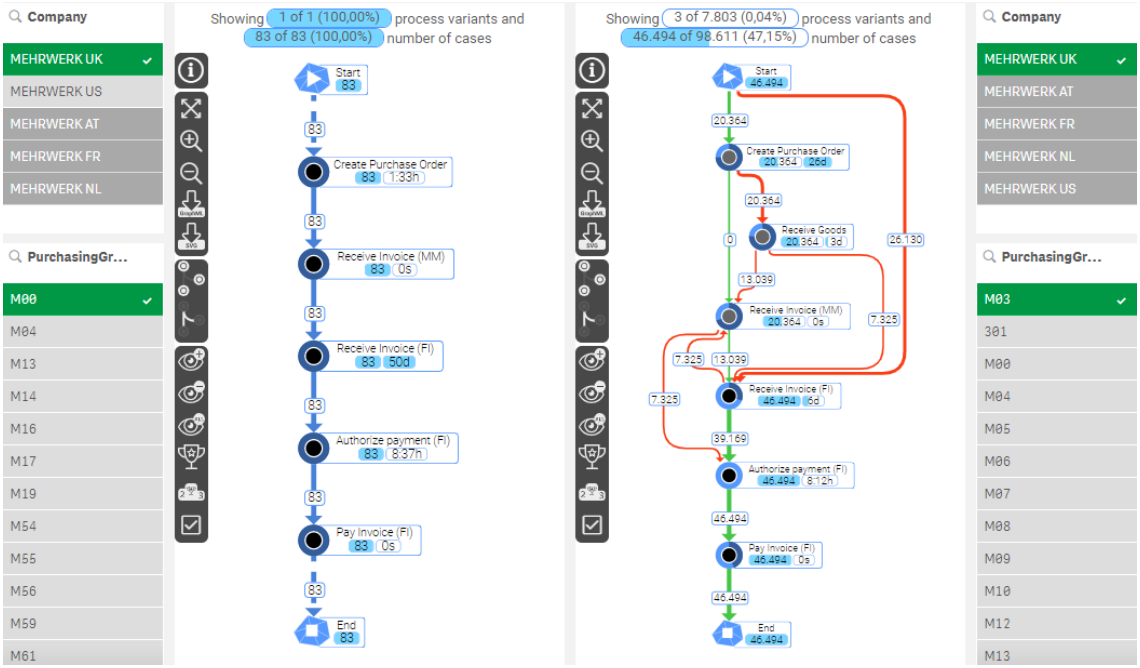

Figure 1. Comparing process deviations between purchasing organizations

variant in company MEHRWERK UK's purchasing group M00. This variant is set as happy path for the right visualization to compare it with the process variants for purchasing group M03. Red moves in the right graph show where M03 deviates from the process of M00 which is marked greenly.

3. Description of Maturity

As Qlik Sense® is the platform for MPM, the maturity of the analytic capabilities is eminently high with more than 50,000 customers of different sizes and industries trusting in Qlik Sense®'s functionality. Advantageous for any development actions is the large community of Qlik users which also provides open source code for extension development. The MPM functionality itself has been tested in various real-life projects receiving strong positive feedback by customers with environments varying from small up to large and e.g. multiple SAP® systems and heterogenous data sources. The deterministic MPM ProcessDiscovery Algorithm creates a dependency graph as process visualization showing the *spaghetti-like* process variants. The filter and selection functions of Qlik Sense® help to reduce the process variants to the relevant scope. The MPM ConformanceChecking Algorithm handles should-be process models by splitting them into the distinct process variants which are then compared to the real-life process variants. Planned but not yet implemented is conformance checking that also includes the should-be process step durations or process step idle times in the fitness calculation. In none of the projects where MPM was implemented problems with respect to scalability occurred. Nevertheless, to meet the future demands by Big Data, MPM is prepared to confront the upcoming challenges with Qlik Sense®'s strategies.

3.1. Handling Big Data

To overcome performance problems due to big data, Qlik follows different approaches like *on demand app generation* and *large scale architectures for horizontal* or *vertical scalability*. On-demand apps allow the user to load and analyze big data sources, where aggregated, representative visualizations are created for the whole data set to identify interesting data subsets on which detailed analysis is carried out afterwards by interactively generating an on-demand app with the full Qlik in-memory capabilities [4]. With respect to large scale architecture, Qlik Sense® recently adopted Kubernetes cluster using containers [5]. Thereby, high-elastic deployments into Kubernetes clusters become possible, running in either public or private clouds on customer managed infrastructures [6]. Having access to this architecture MPM can handle big data respectively. The interested reader finds more information about scalability at Qlik's Whitepaper [7].

3.2. Case Studies

One of the first projects realised with MPM was to rapidly identify compliance issues as well as to ensure and prevail corporate governance in a *purchase-to-pay* process of a large German energy supplier. The audit, governance and industrial control system area of this enterprise initially comprised 1,828,864 cases and 3,036,204 events composed by 27 activities. The data were obtained from SAP® ERP

and Excel sources. With MPM the company achieved full as-is purchasing process transparency across multiple company codes and purchasing organisations combined with automated calculation and evaluation of risk indicators.

Another application was an *order-to-cash* project at a German raw material supplier where MPM was used to identify optimization potential for days of sales outstanding (DSO), order lead time, order processing cost and the general process. 477,345 process instances with more than 4.4 million events leading to over 45,000 process variants have been analyzed.

Another project by a German automotive logistics supplier was the optimization of their *logistic process*. The small logistics data set came from a SAP® ERP including 6,159 cases, 49,711 events which were built up of 15 activities and formed 85 process variants. MPM generated insights in optimization potential, for example, by identifying the refueling process as bottleneck. Major achievements were the discovery of "forgotten" cars on parking sites and a new benchmark system due to understanding which transport system - train or truck - was the better option in certain situations.

One project that was not based on SAP® input data was the analysis of a *production process* which' data were recorded in two different information systems – a computer-aided quality system and a manufacturing execution system. During production, the process instance changed from batch process to single production process. Therefore, the requirement was to achieve an overview on the overall process by investigating the subprocesses in one single app. Hence, two event logs were extracted by the MPM RuleEngine from the source systems, the first with 6,483 batches as process instance, 11 activities and 51,334 events and the second with 49,396 single production numbers as process instances, 17 activties and 322,201 events. The two event logs were processed and led to two process logs that were associated to one another via the batch information. The processes were visualized seperatly but in the same app so the analysts could clearly see the overall process and track the single production number down to batch events. The aim of employing MPM was to generate a starting point for continuous process improvement. Slow process variants as well as production bottlenecks were discovered and selected for improvement activities.

3.3. The Vision

A first version of decision tree forecasting and root cause analysis capabilities is in testing at the moment. Planned for the future is an intuitive root cause analysis modeling functionality and stronger prediction capabilities with deep learning algorithms for decision support and automated activity triggering. The enhancement of MPM's process visualization with hierarchical functionalities that allow for the aggregation of activities to a super-node is in development. Aggregating and deaggregating the view on as-is processes will thereby be supported.

Having recent developments of Qlik Sense® regarding chat bots and alerting functions [8] in mind, we are looking forward to real-life projects to embed these capabilities within process mining scenarios. Hand in hand with the planned export of process models, MPM will not only be a tool to generate insights and appealing diagrams but to automatically communicate important process informations directly to the responsible stakeholders.

4. Conclusion

MEHRWERK ProcessMining based on the Qlik Sense® platform is first class self-service process mining that provides the data governance necessary for enterprise-wide use. MPM offers fast implementation through flexible data extraction options for a wide range of use cases with or without a previously prepared event log and is easily adaptable to specific analysis requirements due to integratability with machine learning functions (e.g. DataRobot/Python/R).

We see application possibilities of MPM in technically every context where IT systems track events. Experiences have been collected in P2P and O2C processes, classic logistic and production processes, banking portfolio management and the monitoring of user-interaction with machines. MPM can be generally helpful for improving processes due to algorithmic process discovery and analysis. Furthermore, it works as enabler for root cause analysis. Improving auditing and compliance by algorithmic process comparison is an application field as well as supporting process automation or digital transformation by discovering opportunities and linking strategies to operations. An interesting use case would be the monitoring and controlling of S/4 Hana migrations. A 25 minutes introduction video can be found at:
https://mpm-processmining.com/demo-video-mpm-2019-en/,
password: MPMDemo2019

References

[1] QlikTech International AB, *The associative difference: Freedom from the limitations of query-based tools*. https://www.qlik.com/us/resource-library/?search=associative+difference\&Resource+Types=-1, 01.11.2018.

[2] QlikTech International AB, *Qlik Connectors®*. https://www.qlik.com/us/products/qlik-connectors, 01.04.2019.

[3] Qlik Branch®, *Network Vis Chart*. https://developer.qlik.com/garden/56cea95aaaacff050825f6f8, 01.04.2019.

[4] QlikTech International AB, *Managing big data with on-demand apps*. https://help.qlik.com/en-US/sense/November2018/Subsystems/Hub/Content/Sense_Hub/DataSource/Manage-big-data.htm, 01.04.2019.

[5] QlikTech International AB, *Qlik Sense® Enterprise for elastic deployments*. https://help.qlik.com/en-US/sense/February2019/Subsystems/PlanningQlikSenseDeployments/Content/Sense_Deployment/Deploying-Qlik-Sense-multi-cloud-Efe.htm, 01.04.2019.

[6] QlikTech International AB, *Qlik Sense® Enterprise architecture and scalability*. https://www.qlik.com/us/resource-library/qlik-sense-enterprise-architecture-and-scalability, 2018.

[7] QlikTech International AB, *Qlik Sense® Performance Benchmark*. https://www.qlik.com/us/resource-library/qlik-sense-performance-benchmark, 2017.

[8] QlikTech International AB, *Qlik Sense® Bot - Power of the Qlik Platform*. https://www.youtube.com/watch?v=aoZ-LaicXHs, 18.04.2017.