University of Sindh Journal of Information and Communication Technology
(USJICT)
Volume 3, Issue 1, January 2019

ISSN-E: 2523-1235, ISSN-P: 2521-5582
Website: http://sujo2.usindh.edu.pk/index.php/USJICT/

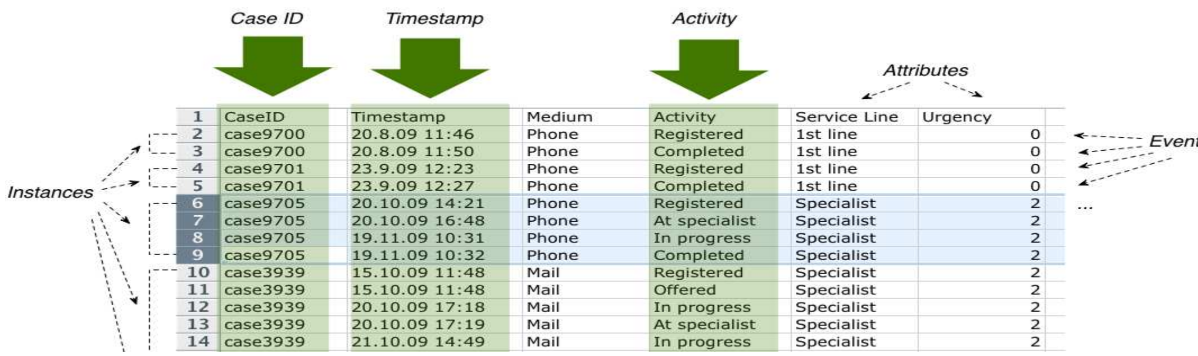
PROCESS MINING APPROACH TOWARDS OPTIMIZATION OF ERP BUSINESS PROCESSES: A CASE STUDY OF HEALTHCARE

Ayesha Rashid, Naveed Anwer Butt, Nauman Riaz Choudhary, Reema Choudhary, Huma Jabeen
Department of Computer Science, University of Gujrat[1,2,3,4], Institute of Management Science and Technology, Lahore[5]
ayesha.rashid@uog.edu.pk, naveed@uog.edu.pk, nauman.riaz@uog.edu.pk reema@uog.edu.pk, hjabeen2@gmail.com

Abstract: Today enterprises are growing in size supporting globally distributed and vigorous businesses involving a huge number of events. Enterprise Resource Planning (ERP) systems have a particular way of recording activities in the form of event logs. Organizations having ERP systems apply process mining techniques to extract information from event logs and analyze the actual business processes underlying the organizations. In this paper, we have focused over one of the process mining application features i.e. discovery of process model with control flow perspective. Acase study of a Pakistani Enterprise has been used to analyze the results of three applied discovery algorithms; Alpha Miner, Fuzzy Model, and Heuristic Miner. After being applied these three discovery algorithms has determined the most frequent behaviors underlying the complaint monitoring process. It is observed that for ideal process model discovery, preprocessing is the most essential and preliminary phase. The results of process mining techniques have been analyzed and visualized by using ProM framework and DISCO tool.
Keywords: Process Mining, Conformance, ERP systems, Data Mining, Alpha Miner, Petri-net

I. INTRODUCTION

Process mining is a channel between techniques of model-based process and data-oriented analysis [1,2]. It provides a powerful mean to analyze existing business processes on the basis of the actual execution logs. Different techniques are used to choose an optimal process model for process mining. the main objective of process mining is to extract the process related useful information from user event logs. [3,4,5]. Health care market uses the process mining techniques to sleeking their processes in order to deliver the high-quality products at minimal cost. Hospitals used many communication Technology tools such as electronic clinical charts, computerized guidelines and decision support systems to craft huge volume of data [8,9]. Process mining tools are applied on collected data afterward Process analysis is used to inspect the event data log. Data should be consisting of at least following three attributes ***Case Id***, ***Activity***and ***Time Stamp*** to get a precise knowledge [5, 10]. **Case id** exemplifies process instance, **Timestamp** shows the time of event occurrence and **Activity** represents the status of the event as shown in Figure 1. To cope with any change, there must be some approach to discover the real business process structure. For that process, mining is one of the most suitable and efficient approaches to optimize business processes. ERP solution implementation is crucial for the improvement and efficiency of an organization, but it is important that the ERP solution represents the real business processes that depict the real workflow. Information about processes extract from information system's event logs, technique, tools, and methods for modeling are used to analyse the bottlenecks in the current business structure that would be useful for further enhancement of said system [11,12]. The purpose of this research is to improve and check the conformance of activity-oriented process models from event logs. Data set used for this research is collected from the Centers for Medicare & Medicaid Services [14,15]. The data collected by this way deal with the comparison of different hospitals data. Replay log on Petri-net for conformance analysis applied to healthcare systems data [16].



	CaseID	Timestamp	Medium	Activity	Service Line	Urgency
2	case9700	20.8.09 11:46	Phone	Registered	1st line	0
3	case9700	20.8.09 11:50	Phone	Completed	1st line	0
4	case9701	23.9.09 12:23	Phone	Registered	1st line	0
5	case9701	23.9.09 12:27	Phone	Completed	1st line	0
6	case9705	20.10.09 14:21	Phone	Registered	Specialist	2
7	case9705	20.10.09 16:48	Phone	At specialist	Specialist	2
8	case9705	19.11.09 10:31	Phone	In progress	Specialist	2
9	case9705	19.11.09 10:32	Phone	Completed	Specialist	2
10	case3939	15.10.09 11:48	Mail	Registered	Specialist	2
11	case3939	15.10.09 11:48	Mail	Offered	Specialist	2
12	case3939	20.10.09 17:18	Mail	In progress	Specialist	2
13	case3939	20.10.09 17:19	Mail	At specialist	Specialist	2
14	case3939	21.10.09 14:49	Mail	In progress	Specialist	2

Figure 1: Event Log Data of Three Attributes

II. PROCESS MINING PROCEDURES

Process mining can be categories in the following of three forms

A. *Discovery*: Discovery used to define the process model based on an event [12].

B. *Conformance*: Conformance is the process to detects the inconsistencies of the existing model by comparing it to newly constructed process model [12, 13].

C. *Enhancement*: The process model is enlarged with additional information or aspects, such as performance data [13].

This paper emphasizes on conformance parameter to takes an event log and its existing process model to checks lack of consistency among them. Numerous deterministic, heuristic and genetic algorithms along with tools are used to construct a process model. ProM 6.5.1 toolkit and Disco Fluxicon have widely used conformance tools of process mining [14]. ProM generates dotted charts and calculates originator to task matrix by forming the interaction between different plug-ins. In this paper, ProM 6.5.1 used to dig out the discrepancy from event log data [15].

III. RELATED WORK

The α-algorithm on audit trails to process the security efforts for high-level fraud prevention used in to mined the workflow net for process execution and showed that process conformance can be checked by comparing process fragments with the discovered Petri net[1]. It 's believed that it would be helpfulfor organizations to monitor and store event logs for furthered use[2].A ProM conformance checker tool is used to measure the conformance of event logs. They worked on “fitness” and “correctness” of BPEL choreography language to trace every event in sequence asan ideal or actual event occurred. Heuristic miner algorithm expresses the main behavior registered in an event log [3]. Heuristic miner can deal with noise and low frequent behavior and can mine a long distance dependencyby Mining a dependency graph and put on AND/XOR-split/join and non-observable tasks.

Thevenetet.al [4] proposed a representation of protocol, based on message exchange and finite state automata. They focus on essential properties for the verification of interoperability of a set of services. They proposed a conformance test that can guarantee the interoperability of a set of services by verifying properties of the single service against the protocol. This is particularly relevant in open environments, where services are identified and composed dynamically on demand. Protocols have been formalized in the simplest possible way to capture the essence of interoperability and to define a fine-grain conformance test. Most of the information systems log events to audit and monitor their supported processes. Their Conformance can be measured with two parameters “fitness” and “appropriateness”.

Suraj et al. [5] used a ProM framework conformance checker to audit the structural and behavioral appropriateness. Although process mining is an ongoing procedure to mine the behavioral activities all the activities in one process cannot be measured.

Hosian and Martin [6] identified that most of the process mining techniques cause “overfitting” and “unfitting”. They proposed a two-step approach to control the flow of process mining. The firststep, detected that dependencies on the desired Properties of the model and the characteristics of the log, the proposed algorithm can be used to provide a more suitable model. In the second step, a process model is synthesized from the transition system resulting from the first step. They used the theory of region implemented in ProM framework.

Burattin et al. [7] reveal that in the context of the health care market, the focal point of hospitals is to deliver the high-quality products in condensed costs. To achieve this target, different communication tools are used that are not only examining thedata but also do the process analysis. Families of a-posteriori analysis techniques are used on clinical and pre-hospital datasets to improve the effectiveness of the data. Results of clinical dataset were generated using heuristic miner algorithm Whereas, Petri net was used for pre-hospital datasets. Structure and control parameters of software development processes are very important to improve efficiency and reduce the cost of software.

Complexities of processes force the organizations to adopt business process analysis techniques. Yanget al. [9] presented an analysis based ProM framework technique to measure the conformance of an event log for a specific given process model. Conformance checking defines, how good a model of a process is with respect to an event log to records the executions of the process. The benefit of always finding the best possible sequence of transitions to fire the cost parameters comes at the cost of computational complexity.

Sami et al. [10] developed aProM BMP model alignment technique to achieve insights into the control flow and performance of the process execution. The model showed the deviation and alignments between event logs and process models from process discovery. On the basis of these deviations, logs are repaired in the form of a model process. Process mining techniques used to discover a process model, check the conformance of a process model to enhance its performance. The addressed case study shows that projection of alignments onto process model delivers reliable Performance information.

Process mining can be considered as the **"missing link"** between data mining and business process management. Conformance checking techniques verify whether the observed behavior recorded in an event log matches a modeled behavior. Sugiyantoet al.[11] exposed that although there exist solid conformance checking techniques for procedural models it is needed to improve the conformance checking for declarative models. The alignment of log and declarative model provides sophisticated diagnostics that pinpoint where deviations occur and how severe they are. A ProM based approach implemented on both synthetic and real-life logs from Dutch municipalities. Based on such alignments a novel diagnostic is provided at the trace level to show why events need to be inserted/ removed in a trace, and at the model level, coloring constraints and activities in the model based on their degree of conformance.

The aim of Sowjanya et al. [12] is to reveal the applicability of conformance check techniques to quality management in Production Company. They used a ProM process conformance check and project management technique to check the deviations of business inventory.

Cristina et al. [13] have presented two case studies to provide the authenticity of data-flow mining algorithms. naive algorithms are proposed that focused on data-flow perspective and presented a general solution for the extraction of the data-flow model from different data sources.

Kingsley et al. [14] have proposed a novel approach to automate the learning processes by extracting useful patterns from different data sources. With the help of semantic reasoning, they applied semantic annotation of activity logs within the learning process to discover learning patterns automatically. OWL and SWRL web languages are used with process mining techniques to discover various learning patterns.

Luís et al. [15] extracted the business transactional data from the system data and applied a process discovery algorithm to determine a business process model using the ProM framework.

IV. RESEARCH METHODOLOGY

This section illustrates the complete process to mine the conformance activities from event logs of health care information system. Steps involve completing the activities are as follow:

A. *Data Source and Collection*

Health care flat file data collected from a reliable source of Data.Medicare.gov available on following *link.https://data.medicare.gov/HospitalCompare/Complications-Hospital/632h-zaca#*

This whole dataset consists of 19 attributes, 32721 rows, and 621,699 entries. Provider ID, Hospital Name, Address, City, State, Zip Code, Country Name, Phone Number, Measure Name, Measure ID, Compared to National, Denominator, Score, Lower Estimate, Higher Estimate, Footnote, Measure Start Date, Measure End Date are the attributes used to mine the process of health care corpus.

B. *Data Transformation*

As the extracted data is in CSV file format, this research needed an XES type of file. ProM6.5.1 toolkit transforms the file in the required format as shown in Figure 2.

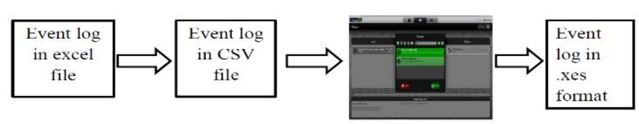


Figure 2: Event Log File Transformation

The converted file represents event log span start from period April 1,2011, to June 30, 2014, consisted on 1 originator and only one process, 1463 cases with 20000 events, 14 events classes and 2 event types.

C. *Data Preprocessing*

Data preprocessing is the process to examine and filter the XES log data to remove the unnecessary attributes from a file and fill the missing values by mining techniques. For example, to calculate the throughput time of cases if the log does not contain information about the dates and time of event occurrence. The next step is to inspect the preprocessed XES format file to view the log summary. Log summary can express the:

i. Number of cases (or process instances) in the log
ii. Number of tasks (or audit trail entries) in the log
iii. Number of resources in the log
iv. Running cases in the log
v. Which resources work on which tasks?

Several activities have been performed on data to preprocess all the outliers from corpus to apply process mining techniques.

The first step is to sort the log files according to occurred events. As events is the main entity so all the preprocessing activities applied to log events. The next step is to remove the unnecessary data from the extracted corpus. Event logs are observed manually to remove the empty and repeated events. Then the whole corpus re-examined to detect the remaining deficiencies. Now data is ready to apply multiple Generalization rules of activities to filter the required fields. event logs are filtered as discussed in data filtration section, outliers are removed. Now the cleaned log is available for process mining. Figure 3 shows the complete process of data preprocessing.

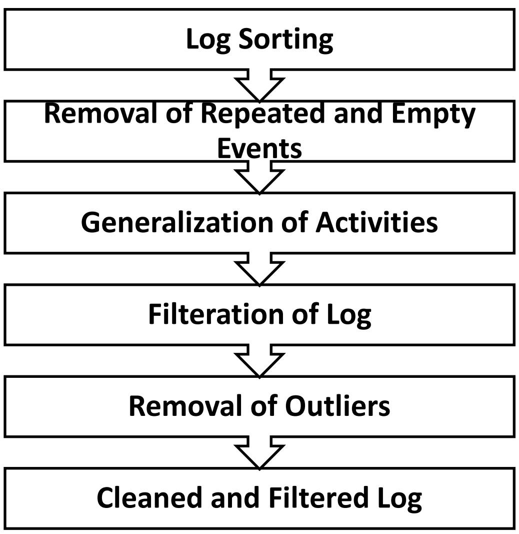


Figure 3: Steps of Data Preprocessing

D. Data Filtration:

This phase filtered data using ProM 6.5.1 toolkit. Filters can add, remove or modify an event, discard the non-required process instances and can filter logs by attributes. This research implements four types of filters on data, as they are mentioned below:

i. The first log filter is the event type filter to keep the instances of required events in the corpus.
ii. The second log filter is the start event to keep only such cases in data which start with the indicated tasks.
iii. The third log filter is the end event filter to keep only such traces (or cases) which end with the indicated tasks.
iv. The fourth filter is the event filter, which filters all unselected events from the log.

E. Data Modeling Techniques

Process mining plug-ins are used for the Process Discovery and Conformance Analysis. For log analysis, different algorithms can be used depending upon the nature of data. It is a real-time data set because of what there is a chance of having noise in it. In order to deal with noise Alpha Miner and fuzzy set, algorithms are used for process discovery and conformance analysis.

F. Process Discovery Analysis by using Disco

Disco is a complete process mining toolkit from Fluxicon that makes process mining fast, easy, and simply fun [37]. It is based only on fuzzy miner and provides reliable and trustworthy results in a short time span. Figure 4 depicts the detailed information of each event for data discovery analysis using the fuzzy set technique. It also shows the summarized view of data. Which includes 10,000 Events, 1463 cases,7 Activities, mean (average time of processing of events) and median duration with start and end time.

Figure 4: Statistical Analysis in Disco

G. Variants

Maps show the overview of the process flow between multiple activities by all cases together. A variant is a rich feature of Disco used for specific classification of activities. It can be perceived as one path from the start to the very end of the process. Usually, a huge portion of cases in the corpus are following just a few variants, and it is useful to know which are the most frequent ones.

Event logs are distributed in the form of variants. Variants are classified on the basis of cases. Figure 5 shows the complete detail of events

➔Variant 1 has 1422 cases, 7 events, mean duration 4 years and 156 days and median duration 4 years and 156 days.

➔Variant 2 has 39 cases, 1 event, mean duration 1 years and 185 days and median duration 1 years and 185 days.

➔Variant 3 has 1 case, 6 events, mean duration 4 years and 156 days and median duration 4 years and 156 days.

➔Variant 4 has 1 case, 1 event, mean duration 1 years and 154 days and median duration 1 years and 154 days.

H. Measure fitness

In this study to measure the fitness between event logs and process models, Replay a log on Petri net for conformance analysis Plugin in ProM 6.5.1. The replay of each logical log trace begins with the marking of the primary place in the model. Then, the logged event belonged transitions are fired one after another. During Replay execution, artificially created tokens are counted. (i.e., the transition belonging to the logged event was not enabled and therefore could not be successfully executed) and the remaining number of tokens in the model indicate that the process was not properly completed [9].

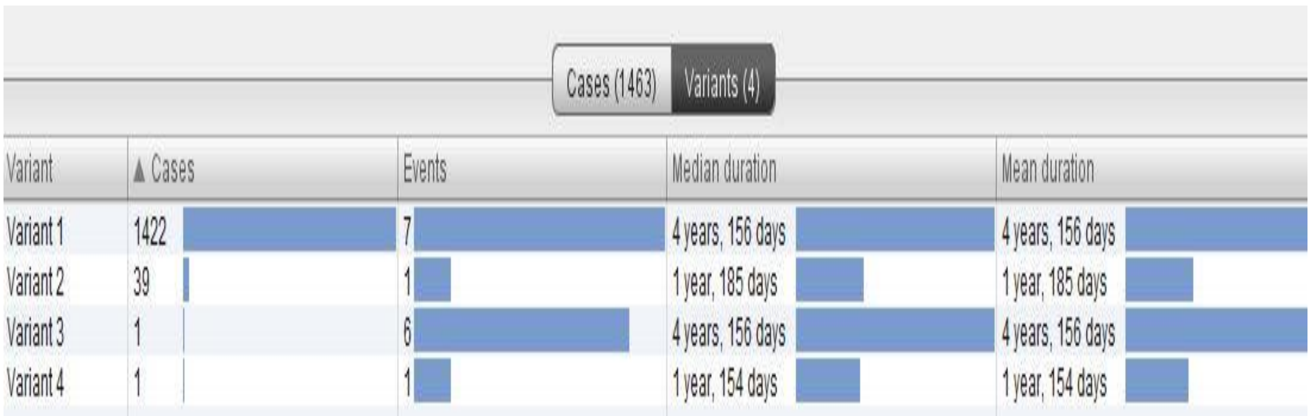


Figure 5: Variants W.R.T Cases

V. Results

Result section represents multiple techniques' results applied to log event prepared data. Data filtration, ProM, Petri-net log data, global statistics and element statistics results are discussed. Filters applied on log data in different categories like by attributes, by event attribute values and by using simple heuristics. The said categories showed the results in processes, cases, events, event cases, event types and originators as shown in table 1.

TABLE 1: FILTER RESULTS

Filter	**Processes**	**Cases**	**Events**	**Event classes**	**Event Types**	**Originators**
Filter Log by Attributes	1	1463	20,000	14	2	1
Filter Log on Event Attribute Values	1	1463	20,000	14	2	1
Filter Log using Simple Heuristics	1	1463	20,000	14	2	1

The above results conclude that all the filters represent the same results of event log which shows that data is noise free. Then a process discovery alpha algorithm applied to preprocessed data. The alpha algorithm is chosen to mine the process data because the filtration results have shown that sample data is not too much noise and most of the event traces were complete. The said miner mines the results in the form of Petri-nets as shown in Figure 6.

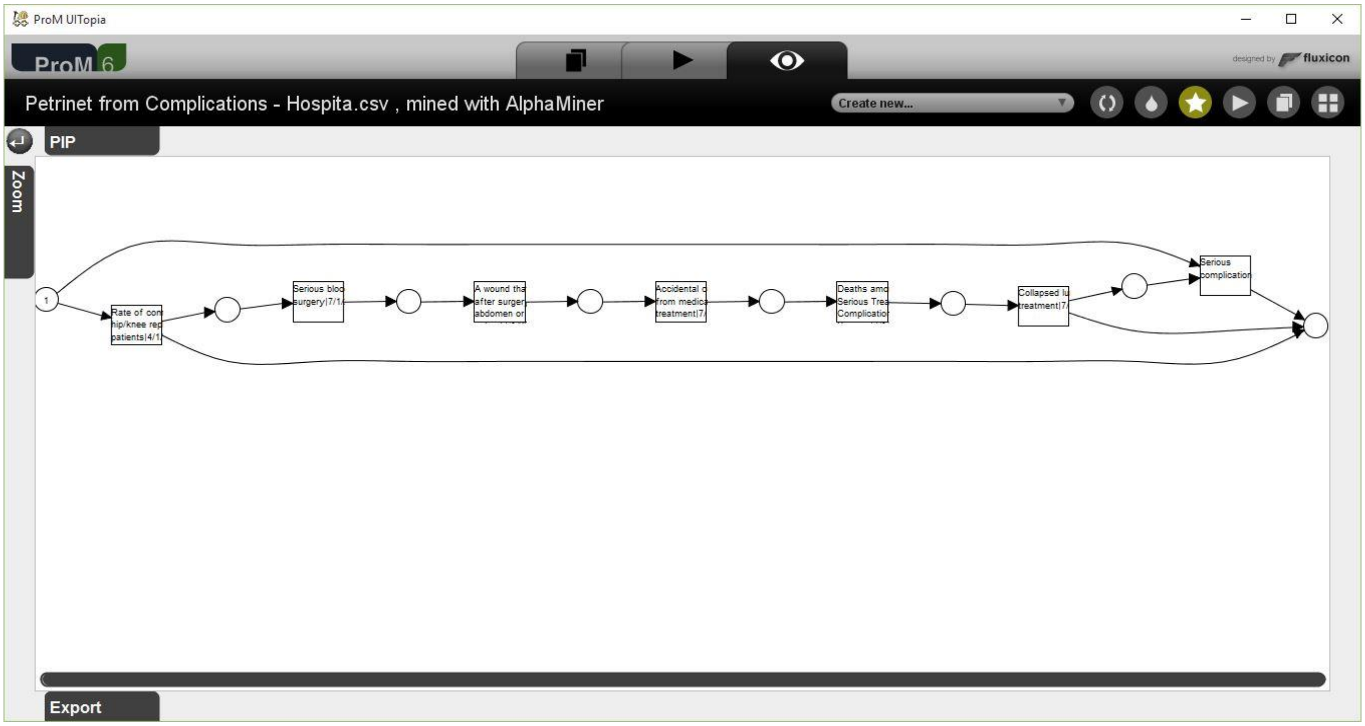


Figure 6: ProM Result

Petri net and log data are taken as input to map the logs. After mapping data, a conformance algorithm i.e. Replay log on Petri-net for conformance analysis is used to get the fitness and model of log data. The model shows the deviation of between events as shown in Figure 7.

The yellow place shows where the move log occurred. Red border line indicates the non-alignment of events also occurred in these event classes with respect to log data. White place indicates the state where No move log occurred. In figure 7 yellow circles indicate the total alignment of logs and events whereas red border lines indicate the non-aligned events. Alignment represents moves in logs related to moves in the model. These moves are called legal moves. There are seven event classes; all of them aligned and three of them have some nonaligned moves as mentioned below

Aligned Events:

i. Accidental cuts and tears from medical treatments

ii. Among patients with serious treatable complications after surgery.

iii. Collapsed lungs due to medical treatments.

iv. Serious complications.

v. The rate of complications for hip and knee replacement patients (1462 move are aligned).

vi. Serious blood clot after surgery (1423 moves are aligned)

vii. A wound that splits open after surgery on abdomen or pelvis (1423 moves are aligned).

Non-Aligned Events:

i. The rate of complications for hip and knee replacement patients (1 move is not aligned).

ii. Serious blood clot after surgery (40 moves are not aligned).

iii. A wound that splits open after surgery on the abdomen or pelvis (40 moves are not aligned).

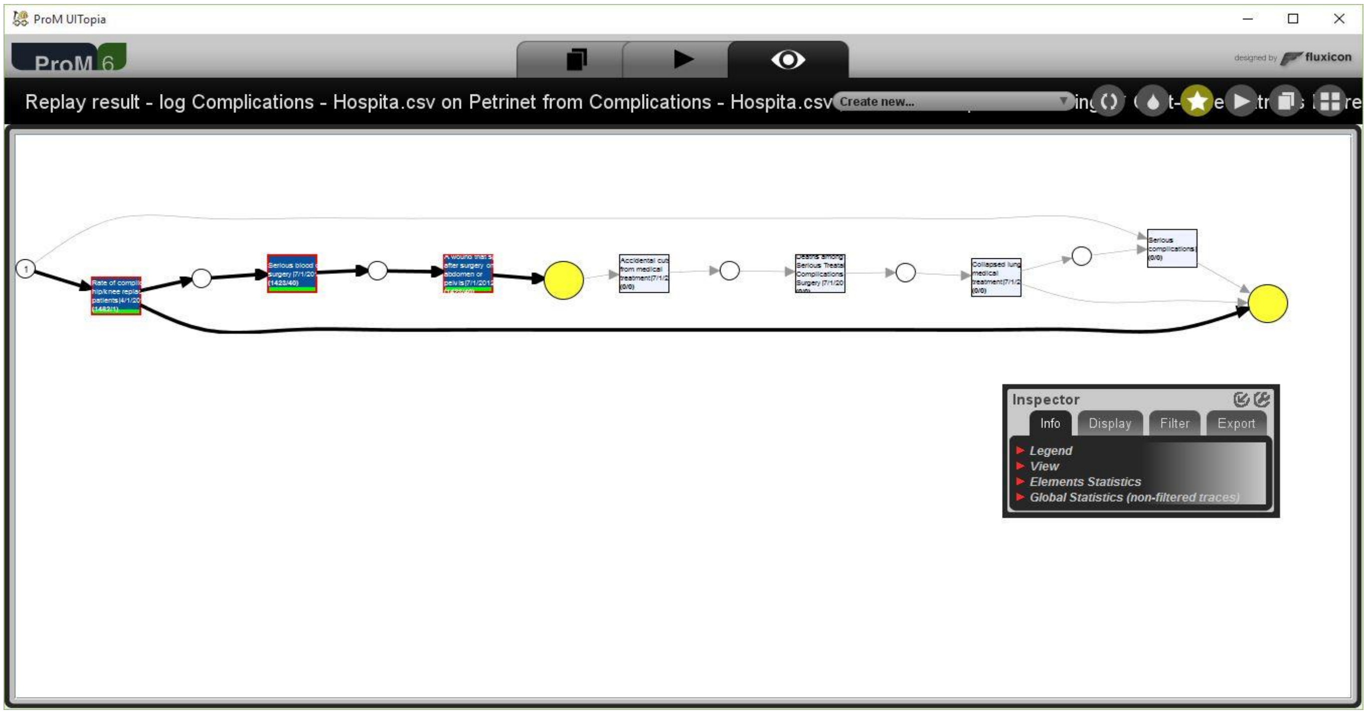


Figure 7: Petri-Net and Log Data Results

Alignment to Log

After applying the algorithm, the visualizer changed to project alignment (a standard visualizer) to check further results. In Figure 8, it shows the results which represented the fitness of log and model as well as the log model alignment that how much align they are. The above results in figure categorized the caseids into 4 levels and further define them. Secondly, it defines the legend in which different types of moves are defined including log and model moves. Further shows the average log fitness, average model fitness, average trace fitness, average trace length.

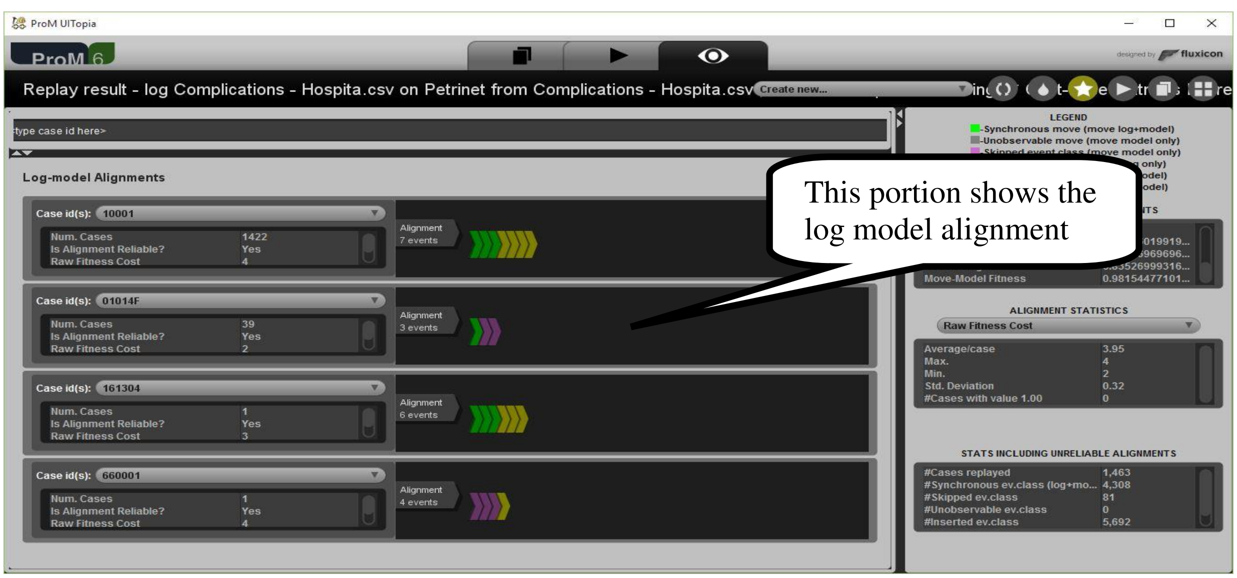


Figure 8: Element Statistics W.R.T Project

A. Category 1 (10001-161303)

From the case id 10001 to 161303, 7 events are aligned. It has log fitness 0.43 and Model fitness of these ids is 1. It includes 1422 number of cases. This shows that these cases have same log and model fitness i.e. 0.43 and 1. The next question is what are they all cases are reliably aligned or not? The result shows yes, these all cases are reliably aligned. In this case ids, 7 events are aligned. Green portion shows the synchronous moves of log and model means to log and model is both move here to complete the flow of path. It consists of 3 synchronous moves and 4 inserted event classes. Inserted activities refer to activities that occur in the log but should not happen according to the model.

B. Category 2 (01014F-16004F)

In case id 01014F to 16004F, 3 events are aligned, and it has log fitness 1 and model fitness 0.33. It consists of 39 cases which have log fitness 1 and model fitness is 0.33. The alignment for these cases is reliable. It consists of one synchronous move and 2 skipped event classes. Skipped activities refer to activities that should be performed according to the model, but do not occur in the log.

C. Category 3 (161304)

In case ID 161304, 6 events are aligned, and it has log fitness 0.5 and model fitness 1. It contains only 1 case id with the alignment of 6 events. It consists of 3 synchronous moves and 3 inserted moves.

D. Category 4 (660001)

In case ID 660001, 4 events are aligned, and it has the model and logs fitness 0. It also contains only 1 case id with the alignment of 4 events. It consists of 3 skipped moves and 1 inserted move. So, after considering the results from the above figure 19, we have come to know that in case id 161304 and 10001 to 161303, log and model are more aligned than the other cases.

TABLE 2: SUMMARIZED VIEW OF LOG MODEL ALIGNMENT

Case Id	Aligned events	Log fitness	Model fitness
10001 to 161303	7	0.43	1
01014F to 16004F	3	1	0.3
161304	6	0.5	1
660001	4	0	0

Log and model fitness summarized the data into different forms of parameters to define the accuracy of the model. Table 2 represents the log fitness and model fitness along with measurements. Table 3 represents Alignment statistics of Average Log Fitness signified that log fitness per average case is 0.44 and the cases are consisting of value 1 are 39. Maximum fitness is 1 and minimum fitness of log is 0. Model fitness per average case for Average Model fitness is 0.98 and the cases consist of value 1 are 1423. Maximum fitness of model is 1 and minimum is 0. Trace fitness per average case is 0.6, Max trace fitness is 0.67, Min trace fitness 0 standard deviations 0.02. Trace length per average is 6.84, Max trace fitness is 7, Min trace fitness is 1, and the standard deviation is 0.98.

TABLE 3: SUMMARIZED VIEW OF LOG MODEL ALIGNMENT

Alignment Statistics				
	Move log fitness	Move model fitness	Trace fitness	Trace length
Average/case	0.44	0.98	0.60	6.84
Max	1	1	0.67	7
Min	0	0	0	1
Std. Deviation	0.09	0.11	0.02	0.98
#cases with value	39	1423	0	40

VI. CONCLUSION

This paper emphasized on preprocessing of event log in order to apply process mining algorithms. It is crucial to completely clean and filters the event log before applying discovery algorithms. Tuning of the log for process mining techniques application can be performed to have more precise, simple, generic, refined and so-called ideal Petri-nets at discovery phase. Three discovery algorithms are applied to model the control flow of a complaint monitoring department of the ERP system and have determined the most frequent behaviors underlying complaint monitoring process. After applying fitness through a plug-in (Replay a log on Petri-net for conformance analysis) it is proved that log data and model are somewhere deviating from each other which gave the real insight of data and model.

REFERENCES

[1] Van Dongen B., de Medeiros A., Verbeek H., Weijters A., van der Aalst W.: The prom framework: A new era in process mining tool support, In Applications and Theory of Petri Nets 2005, vol.3536, Springer-Verlag, pp. 444–454 (2005)

[2] Van der Aalst, W. M.: Formalization and verification of event-driven process chains, InInformation and Software Technology, vol.41, N°10, pp. 639-650 (1999)

[3] Van der Aalst W., Van Dongen B.: Discovering workflow performance models from timed logs, Engineering and Deployment of Cooperative IS, pp.107-110 (2002) Thevenet L. H., Salinesi C.: Aligning IS to Organization's Strategy: The I nS t A l Method, In:Advanced Information Systems Engineering, Springer Berlin, Heidelberg, pp. 203-217 (2007)

[4] Suraj Sarda, ToshishAkhand, Sumit Rathod, Khushbu Prajapati, Prof.BharatKapse. "ERP Application for Limestone Mining Industry." International Journal on Recent and Innovation Trends in Computing and Communication, vol. 3, pp. 356-358, Feb. 2015.

[5] HosianAkeel, Martin Wynn. "ERP Implementation in a Developing World Context: a Case Study of the Waha Oil Company, Libya." eKnow 2015 7th International Conference on Information, Process and Knowledge Management. ThinkMind, 2015.

[6] A.Burattin, F. M. Maggib, A. Sperduti. "Conformance Checking Based on Multi-Perspective Declarative Process Models." Mar. 2015.

[7] Anna A. Kalenkova _ W. M. P. van der Aalst _ Irina A. Lomazova _Vladimir A. Rubin. "Process Mining Using BPMN: Relating Event Logs and Process Models." Mar. 2015.

[8] Yang Liu1, Junichi Iijima. "A Case Study of Business Process Simulation in the Context of Enterprise Engineering." 2015.

[9] Sami Laine, JuhaSoikkeli, Toni Ruohonen. "Timestamp accuracy in healthcare business Process improvement." 2015.

[10] Sugiyanto, RiyanartoSarno. "Avoiding Over-generalization ERP Business Processes Using Process Mining Trace Clustering." International Journal of Advancements in Computing Technology 6.2. 2014.

[11] S.Sowjanya Chintalapati, Dr.Ch.G.V.N.PrasadR.Vineela, J.Sowjanya. "Process mining: Business Intelligence (BI) and Business Process Management (BPM)." International Journal of Advanced Research in Computer Science and Software Engineering, vol. 4, May 2014.

[12] MiekeJans, Michael Alles, Miklos Vasarhelyi. "A field study on the use of process mining of event logs as an analytical procedure in auditing." The Accounting Review, vol. 89, pp. 1751-1773, Sep. 2014.

[13] Anna A. Kalenkova, Irina A. Lomazova, and Wil M.P. van der Aalst. "Process Model Discovery: A Method Based on Transition System Decomposition." Application and Theory of Petri Nets and Concurrency. Springer International Publishing, pp.71-90, 2014.

[14] Elham Ramezani, Vladimir Gromov, Dirk Fahland, andWil M. P. van der Aalst. "Compliance Checking of Data-Aware and Resource-Aware Compliance Requirements." in On the Move to Meaningful Internet Systems: OTM 2014 Conferences, vol. 8841. Robert Meersman, HervéPanetto, Tharam Dillon, Michele Missikoff ,Lin Liu, Oscar Pastor, Alfredo Cuzzocrea,TimosSellis, Ed. berlin heidelberg:Springer, 2014, pp. 237-257.

[15] Anna A. Kalenkova, Irina A. Lomazova. "Discovery of Cancellation Regions within Process Mining Techniques." CS&P. 2013.

[16] Asst. Prof. Esmita .P. Gupta. "Process Mining A Comparative Study." International Journal of Advanced Research in Computer and Communication Engineering, vol. 3, Nov. 2014.

[17] Joerg Evermann Ghazal Assadipour. "Big Data meets Process Mining: Implementing the Alpha Algorithm with Map-Reduce." Proceedings of the 29th Annual ACM Symposium on Applied Computing. ACM, 2014.

[18] Dolean, and Cristina-Claudia. "Mining Product Data Models: A Case Study," Informatica Economica, vol.18.1, pp. 69-82, 2014.

[19] Okoye, Kingsley, et al. "A Semantic Rule-based Approach Supported by Process Mining for Personalised Adaptive Learning." Procedia Computer Science, vol. 37, pp. 203-210, 2014.

[20] Luís Filipe Nascimento da Silva, "Process Mining: Application to a case study,"

[21] M.S. thesis, Dept. Faculdade de Economia, Universidade do Porto, Portugal, 2014.

[22] Ewa Ziemba, Iwona oblak. "Critical Success Factors for ERP Systems Implementation in Public Administration." Interdisciplinary Journal of Information, Knowledge & Management, vol. 8, p1, 2013.

[23] S.Hanumanth Sastry, Prof.M.S.PrasadaBabu. "Implementation of CRISP Methodology for ERP Systems." International Journal of Computer Science Engineering (IJCSEm), Vol. 2, No.05 Sep. 2013.

[24] Niels Mueller-Wickop and Martin Schultz. "ERP Event Log Preprocessing: Timestamps vs. Accounting Logic." in Design Science at the Intersection of Physical and Virtual DesignLecture Notes in Computer Science, vol. 7939, Jan vomBrocke, RiittaHekkala, Sudha Ram, Matti Rossi, Ed. berlin heidelberg:Springer, 2013, pp. 105-119.

[25] Rafael Accorsi, Thomas Stocker and Günter Müller. "On the Exploitation of Process Mining for Security Audits: The Process Discovery Case." Proceedings of the 28th Annual ACM Symposium on Applied Computing. ACM, 2013.

[26] Martin Schultz. "Enriching Process Models for Business Process Compliance Checking in ERP Environments." in Design Science at the Intersection of Physical and Virtual Design, vol. 7939. Jan vomBrocke, RiittaHekkala, Sudha Ram, Matti Rossi, Ed. berlin heidelberg:Springer, 2013, pp. 120-135.

[27] Gehrke, StB Prof Dr Nick, and Michael Werner. "PROCESS MININg,"2013.

[28] Santiago Aguirre, Carlos Parra, and Jorge Alvarado. "ERP

Combination of Process Mining and Simulation Techniques for Business Process Redesign:A Methodological Approach," in Data-Driven Process Discovery and Analysis, vol.162. Philippe Cudre-Mauroux, Paolo Ceravolo, Dragan Gašević, Ed. berlin heidelberg: springer, 2013, pp.24-43.

[29] Khaled F. Hussain, Mohamed H. "Improving Genetic Process Mining Using Honey Bee Algorithm." In proc. Informatics and Applications (ICIA), 2013, pp. 59-65.

[30] Fernando Szimanski , Célia G. Ralha, Gerd Wagner, Diogo R. Ferreira." Improving Business Process Models with Agent-Based Simulation and Process Mining," in Enterprise, Business-Process and Information Systems Modeling ,vol.147. SelminNurcan, Henderik A. Proper, PninaSoffer, John Krogstie, Rainer Schmidt, Terry Halpin, Ilia Bider, Ed.berlinheidelberg:springer, 2013, pp.124-138.

[31] Jorge Munoz-Gama, JosepCarmonaz, Wil M. P. van der Aalst." Conformance Checking in the Large: Partitioning and Topology," in Business Process Management, vol.8094. Florian Daniel, Jianmin Wang, Barbara Weber, Ed.berlinheidelberg:springer, 2013, pp 130-145..

[32] Diogo R. Ferreira, Fernando Szimanski and CéliaGhediniRalha.“ Mining the low-level behavior of agents in high-level business processes". International Journal of Business Process Integration and Management, vol.6, pp.146-166, Jan 2013.

[33] Jorge Munoz-Gama, Josep Carmona and Wil M.P. van der Aalst. "Single-Entry Single-Exit decomposed conformance checking". Information Systems, vol.46, pp.102–122, Dec. 2014.

[34] Jianmin Wang, Raymond K. Wong, Jianwei Ding, Qinlong Guo, and Lijie Wen. "Efficient Selection of Process Mining Algorithms". Services Computing, IEEE Transactions on, vol.6, pp. 484 - 496, Oct-Dec 2013.

[35] MiekeJans, Michael Alles and Miklos Vasarhelyi. "The case for process mining in auditing: Sources of value added and areas of application". International Journal of Accounting Information Systems, vol.14, pp.1-20, March 2013.

[36] MiekeJans, Michael Alles and Miklos Vasarhelyi. "Process Mining of Event Logs in Auditing: A Field Study of Procurement at a Global Bank." in proc 9th ICESAL, 2012.

[37] MiekeJans, Michael Alles and Miklos Vasarhelyi. "Process Mining Of Event Logs In Internal Auditing: A Case Study." in proc. European Accounting Association, 2012.

[38] J.C.A.M. Buijs, B.F. van Dongen and W.M.P. van der Aalst. "Towards Cross-Organizational Process Mining in Collections of Process Models and Their Executions" in Business Process Management Workshops, edition, vol.100, Florian Daniel, Kamel Barkaoui and SchahramDustdar, Ed.berlinheidelberg:springer, 2012, pp. 2-1.

[39] Carlos Fernandez-Llatas, Teresa Meneu, Salvatore F. Pileggi, Jose Miguel Benedi, Bernardo Valdivieso and Vicente Traver. "Using Process Mining for Automatic Support of Clinical Pathways Design" in Data Mining in Clinical Medicine,vol. 1246. Carlos Fernández-Llatas and Juan Miguel García-Gómez, Ed.NewYork:springer, 2015, pp.79-88.

[40] Adriansyah and J.C.A.M Buijs.“ Mining Process Performance from Event Logs," in Business Process Management Workshops, vol.132. Marcello La Rosa,PninaSoffer ,Ed. Berlin Heidelberg: Springer, 2013, pp. 217-218.

[41] Mehdi Ghazanfari, Saeed Rouhani, Mostafa Jafari and Peyman Akhavan. "Association Rule Approach for Evaluation of Business Intelligence for Enterprise Systems". The IUP Journal of Computer Sciences, vol.5, pp.21-38, April 2011.

[42] MiekeJans, Michael Alles and Miklos Vasarhelyi. "Process Mining Of Event Logs In Auditing: Opportunities And Challenges". SSRN, 15 Jan. 2010.

[43] Nick Gehrke."Basic Principles of Financial Process Mining a Journey through Financial Data in Accounting Information Systems."AMCIS, 2010, pp.289.

[44] Wil M.P. van der Aalst. "Configurable Services in the Cloud: Supporting Variability While Enabling Cross-Organizational Process Mining" in On the Move to Meaningful Internet Systems: OTM 2010, vol. 6426, Robert Meersman, Tharam Dillon and Pilar Herrero, Ed.berlinheidelberg:springer, 2010 , pp.8-25.

[45] W.M.P. van der Aalst." Process-aware information systems : lessons to be learned from process mining" in Transactions on Petri Nets and Other Models of Concurrency II, vol.5460, K. Jensen and W.M.P. van der Aalst, Ed.berlinheidelberg:springer, 2009, pp.1-26.

[46] Abdullah S. Al-Mudimigh, B. Zahid Ullah, and C. Farrukh Saleem. "A Framework of an Automated Data Mining Systems Using ERP Model". International Journal of Computer and Electrical Engineering, vol.1, pp.1793-8163, 5 Dec 2009.

[47] Laura M˘aru¸ster and Nick R. T. P. van Beest. "Redesigning business processes: a methodology based on simulation and process mining techniques". Knowledge and Information Systems , vol. 21, pp. 267-297, Dec.2009.

[48] Mans, Ronny, et al. "Process mining techniques: an application to stroke care," Studies in health technology and informatics, vol.136, pp. 573, 2008.

[49] Conforti, R., Dumas, M., Garcia-Baneulos, L., La Rosa,M. " Beyond tasks and gateways: Discovering BPMN models with sub processes, boundary events and activity markers," in business process management, vol. 8659, Shazia Sadiq , PninaSoffer , Hagen Völzer , Ed. Switzerland: Springer International Publishing, 2014, pp. 101–117.

[50] Van der Aalst W., Stahl C. Modeling Business Processes: A Petri Net Oriented Approach, MIT Press, Cambridge, MA, 2011.