IoE: A Novel Approach for Software Process Improvement

Shashank Sharma
Manipal University
Jaipur

Sumit Srivastava, PhD
Manipal University
Jaipur

ABSTRACT
Day by day the number of organizations is increasing which using process mining to understand the way of execution of operational processes. For ssystematically drive the innovation in the digitalized world is used by Process Mining. The aim of Process mining is to get the knowledge of an organization's business processes by transforming the event data recorded in information systems. This leads to improve process performance or compliance to organization standards where process mining analysis is implemented. Process Mining techniques are relying on the existence of event data. We require putting too many efforts on making our systems to record historic data at all. The need to understand and improve their processes of businesses requires the process analysis techniques. This paper presents an approach how software event log is analyzed to understand & improve the software process by using classification which later used for the software code clone optimization**.**

General Terms
Data Mining, event log, event classification, processes.

Keywords
Process Mining, PRoM, IoE (Internet of Events), Clone, Clone Optimization.

1. INTRODUCTION
Business Process Management is rapidly changing by the spectacular growth of event data. Without using this valuable information hidden in software systems, it makes no sense to focus on modeling based approach, method & procedure. Organizations are competing mostly on analytics and only organizations that intelligently & efficiently use the vast amounts of available data will survive [5]. Intelligently exploitation of the sudden availability of event data is the today's main innovations. Out of the blue, Big Data" has become a topic in board-level discussions. This data abundance will change many jobs across all industries. Just like computer science emerged as a new discipline from mathematics when computers became abundantly available, we now see the birth of data science in action as a new discipline driven by the torrents of data available in our increasingly digitalized world. There is a huge demand for data scientists, which is rapidly increasing day by day. However, process-orientation should not be blur by the focus on data analysis. In the end, smooth procedure is much more important than data analysis. The old phrase "It's the process stupid" is still valid. Hence, we preach the urgent need for process scientists that will drive process innovations while exploiting the Internet of Events(IoE). The IoE is composed of:

- The Internet of Things(IoT): Every entity(physical) connected to the network. It includes Radio-Frequency Identification(RFID), Near Field Communication(NFC) etc.
- The Internet of Locations(IoL): Every data related to the location i.e. GPS(Global Positioning System) for mobile phone which comprises data in the form of spatial coordinates.
- The Internet of People(IoP): Every processed data for social interaction. It includes electronic mail, instagram, watsapp etc.
- The Internet of Content(IoC): Every processed data created by humans for knowledge advancement on particular domain/ topic. It includes web pages, blogs etc.

Note that the IoT, the IoL, the IoP and the IoC, the partially overlap. For example, a facebook post from the user or the location name tagged from the facebook post will cross verify by using the user IP (Internet Protocol). It is not enough to gather all the event data, but most important is that how we use those data for process improvement. The new approach targeting to validate this issue is Process Mining. Various methods or approach for Process mining together assembles a toolkit for future's process scientist. Process mining connects modeling of processes and data analytics. It can be used:

- Automated process discovery without modeling,
- Searching chokepoint & finding factors responsible for chokepoint,
- Measure the asperity after detection and make understanding of deviations,
- Access the whole degree of abidance,
- Anticipate risks, costs, and delays,
- Suggest measures to avert inefficiencies, and
- Sustain redesign.

Today, there are many mature process-mining techniques that can be directly used in everyday practice [1]. Despite the excess of powerful mining techniques for process and success stories in a variety of application areas, a limiting factor is still some area is not focused properly. In this paper, we focus on the problem of extracting or classifying information from the various event logs. After classification we get the classified data in the form of classes & also we get petri net from the event log. First, we introduce process mining in a somewhat more detailed form (Section 2). Section 3 presents procedure for classification. Section 4 introduces result & outcomes. In Section 5 we conclude this paper.

.

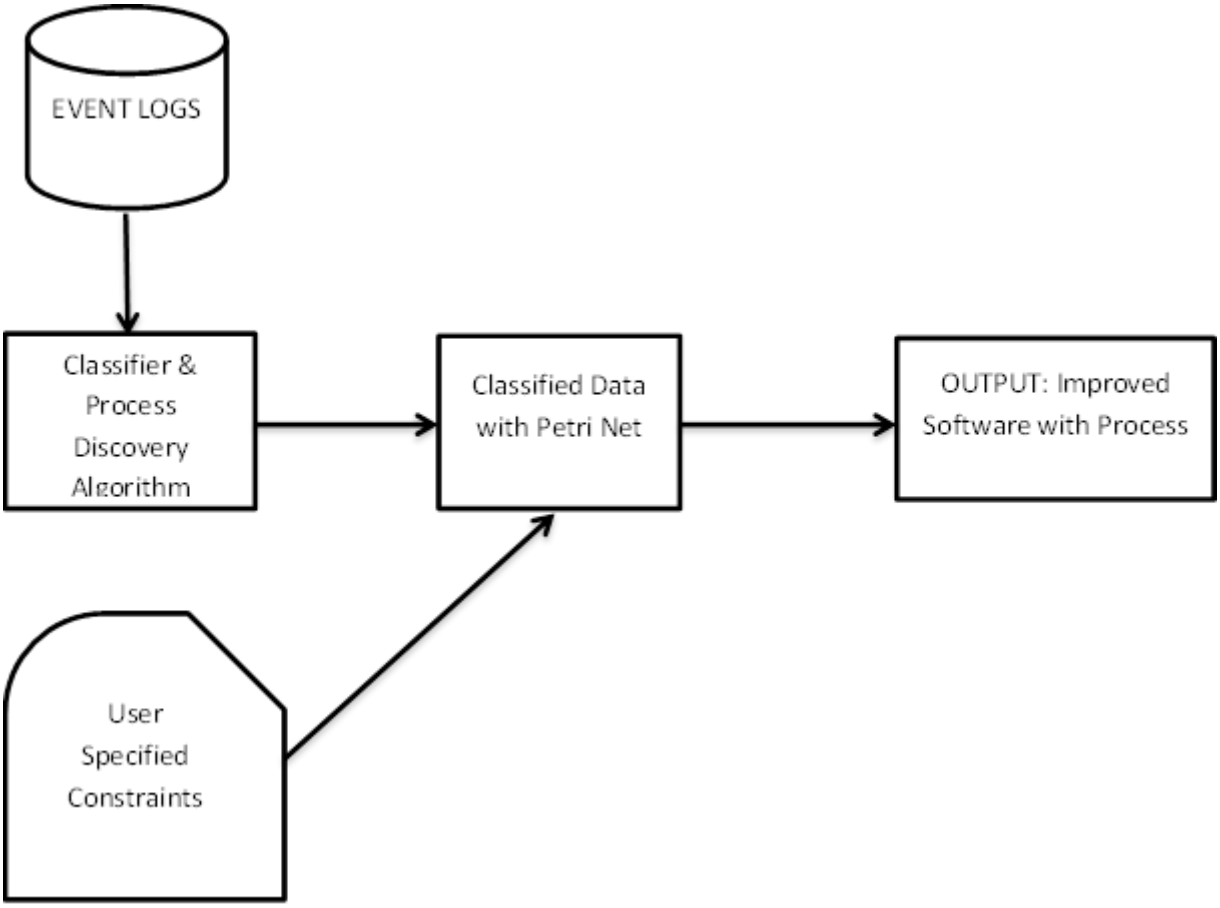
2. PROCESS MINING

Process mining aims to bridge the gap between big data analytics and traditional business process management. This field can primarily be categorized into (1) process discovery,

(2) Conformance checking and (3) enhancement [1]. This allows the extraction of insights about the overall and inner behavior contained in any given process. Process discovery techniques focus on using the event data in order to discover process models. Conformance checking techniques focus on aligning the event data on a process model to verify how well the model fits the data and vice versa [3]. Whereas enhancement techniques use event data and process models to repair or enrich the process model. Hence, process mining provides the bridge between data mining or machine learning techniques and the business process management discipline.

3. PROCEDURE WITH CASE STUDY

In this first we take the event log of Audit Trail Entry case of information system. In this particular event log we have 42 events or events flow. Our aim is to classify the events on the basis of type of events. We start with event log then we select MXML legacy classifier & process discovery algorithm for the extraction of petri net. Along with this we apply some user specified constraints to get expected result.



Flow Chart 1: Procedure for Process Mining

```
<AuditTrailEntry>
    <Data>
        <Attribute name="Activity">examine casually</Attribute>
        <Attribute name="Resource">Mike</Attribute>
        <Attribute name="Costs">400</Attribute>
    </Data>
    <WorkflowModelElement>examine casually</WorkflowModelElement>
    <EventType>complete</EventType>
    <Timestamp>2010-12-30T15:06:00.000+01:00</Timestamp>
    <Originator>Mike</Originator>
</AuditTrailEntry>
```

Figure 1: Event Log Sample from Case Study Event Data Set

4. RESULT & OUTCOMES

In this we received classified result for that particular Audit Trail Entry case. We got 8 classified event classes.

After interpreting the petri net we find that which work flow model we need to work upon to improve the process of information system. Please find below the outcomes.

Table 1: Classified event log data

Model Element	Event Type	Occurrences (Absolute)	Occurrences (Relative)
Check ticket	Complete	9	21.429%
Decide	Complete	9	21.429%
Register Request	Complete	6	14.286%
Examine Casually	Complete	6	14.286%
Reinitiate Request	Complete	3	7.143%
Examine Thoroughly	Complete	3	7.143%
Pay Compensation Request	Complete	3	7.143%
Reject Request	Complete	3	7.143%

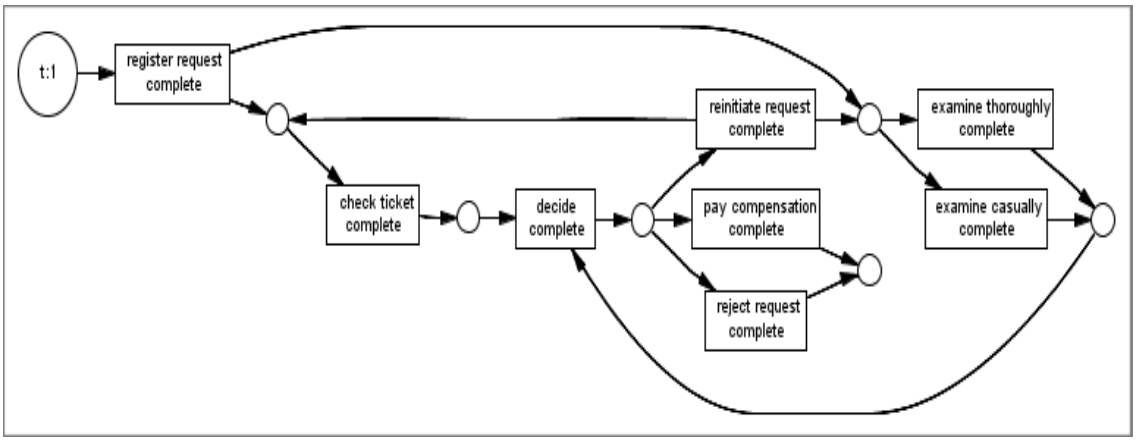


Figure 2: Petri net from Case Study Event Data Set

5. CONCLUSION & FUTURE WORK

To drive innovation in an increasingly digitalized world, the "process scientist" needs to have specific/ precise tools. In this paper we only conceptualized our idea. This allows us to analyze the operational processes of information systems under real-life scenarios, and use mining techniques for processes to obtain precise and formal software improvement models. This paper focused on supporting the systematic extraction of event classes from event data for the software process improvement by using the Petri nets flow model approach. In the future, we aim at conducting additional experiments using different variety of event log data sets. A logical next step is to develop tool support for specific information management systems. One case study demonstrated how we implement our idea. In the future, we would like to incorporate domain knowledge at different stages.

6. REFERENCES

[1]W.M.P. van der Aalst. Process Mining: Discovery, Conformance and Enhancement of Business Processes. Springer-Verlag, Berlin, 2011

[2] W.M.P. van der Aalst, A.J.M.M. Weijters, and L. Maruster. Workflow Mining: Discovering Process Models from Event Logs. IEEE Transactions on Knowledge and Data Engineering, 16(9):1128-1142, 2004

[3] Adriansyah, A., van Dongen, B.F., van der Aalst, W.M.P.: Towards robust conformance checking. In: Business Process Management Workshops, Lecture Notes in Business Information Processing, vol. 66, pp. 122-133. Springer Berlin Heidelberg 2011

[4] W.M.P. van der Aalst and C. Stahl. Modeling Business Processes: A Petri Net Oriented Approach. MIT press, Cambridge, MA, 2011

[5] W.M.P. van der Aalst. Data Scientist: The Engineer of the Future. In K. Mertins, F. Benaben, R. Poler, and J. Bourrieres, editors,Proceedings of the I-ESA Conference, volume 7 of Enterprise Interoperability, pages 13-28. Springer-Verlag, Berlin, 2014

[6] W.M.P. van der Aalst, A.J. Mooij, C. Stahl, and K. Wolf. Service Interaction: Patterns, Formalization, and Analysis. In M. Bernardo, L. Padovani, and G. Zavattaro, editors, Formal Methods for Web Services, volume 5569 of Lecture Notes in Computer Science, pages 42-88. Springer-Verlag, Berlin, 2009

[7] W.M.P. van der Aalst, V. Rubin, H.M.W. Verbeek, B.F. van Dongen, E. Kindler, and C.W. Gunther. Process

Mining: A Two-Step Approach to Balance Between Under_tting and Over_tting. Software and Systems Modeling, 9(1):87-111, 2010

[8] R. Agrawal, D. Gunopulos, and F. Leymann. Mining Process Models from Workow Logs. In Sixth International Conference on Extending Database Technology, volume 1377

of Lecture Notes in Computer Science, pages 469{483. Springer-Verlag, Berlin, 1998

[9] J. vom Brocke and M. Rosemann, editors. Handbook on Business Process Management, International Handbooks on Information Systems. Springer-Verlag, Berlin, 2010

[10] T. Calders, C. Guenther, M. Pechenizkiy, and A. Rozinat. Using Minimum Description Length for Process Mining. In ACM Symposium on Applied Computing (SAC 2009), pages 1451{1455. ACM Press, 2009