

Dealing With Concept Drifts In Process Mining Using Event Logs

Maddu. Uma Lavanya¹, Mr. Sunil Kumar Talluri²

¹Department of Computer Science Engineering,
Specialization In Software Engineering
Vnr Vignana Jyothi Engineering College, Hyderabad
Email id: umalavanya.maddu@gmail.com

² Associate Professor
Department of Computer Science Engineering,
Vnr Vignana Jyothi Engineering College, Hyderabad
Email id: [Email: sunilkumar_t@vnrvjiet.in](mailto:sunilkumar_t@vnrvjiet.in)

Abstract: *Now a days in this e-world most of the business totally related to the process mining trends, it is been generated and checks the process flow in terms of changes in the whole system. Here a few process mining techniques are used to analyze the changes in the drifts, this drifts may be changes according to the process and it can be sudden change, gradually change or the random change. This changes can be visualized by using the Rapid Miner tool, in that a ProM framework is used to get the better Results. Not only it tracks the change it also finds the localization in the business Process. Here Feature Extraction and Generate Population method is used to find the relationship among the activities. Event logs are used to establish the process flow*

Keywords:- *Process Mining ,Drifts ,Event Logs , Feature Extraction, RapidMiner*

1. Introduction

Process mining is a one kind of a process developing the management flow technique that permits for the visualizing and the analysis of business processes and the management based on a particular event logs. The idea is to extract the knowledge of processes from event logs that are recorded by an information process system process depending upon the substance mining. Process mining [1] this process mining aims at checking the changes and also improving this by providing some alpha mining techniques implement the process and tools for discovering data, control, process, and social structures organizational, from event logs. Process mining techniques mostly used when no formal description of the process can be replace by the other approaches, when the quality of process are in readily existing. Example, the databases of a workflow management system of a bank, that records the transaction event logs of the business enterprise. Frame work method is almost based on the characteristics depending upon there is a petri net model and if there is no petri net model then it only checks the demand of the particular system events, how it is used the event log [2] available at any instance of time management of process Then data mining techniques are offence we used to see which system environment data elements are replaceable with one another. As a result, a petri net diagonally tree is developed and generated for each log choice in the process. There are 3 classification Process mining techniques (1) Discovery (2) Conformance Analysis (3) Extension.

Process mining is closely related to, BPI (Business Process Intelligence) [3], BOM (Business Operations Management) [4],

BAM (Business Activity Monitoring) [5] and data workflow mining.

1.1 Business Processing

A business transaction is that it requests the information of a particular from or changes the data in a database. A particular event in a subsequent ways forms the chain of structured in an regular business activities. The event generally changes the state of product or a data and generates some information kind of an output. For Examples of tracking the whole system business processes include delivering, supply of same kind materials, shipping products, data error problems, updating employee information, tracking receiving orders, or setting financial terms. Activity processes occurs at every levels of the profitable organization's activities, data, resources activities include process events that the customer needs and the events that are invisible to the customer.

Business Process Management (BPM) [6] plays a major role in the business environment. The term business process management and the business operational resources are covers how we identify, study, monitor and change, business processes to ensure the process of the event logs and classifies the structure improved over time.

2. Related Work

Operational processes need to change the instance changing circumstances, extreme variations in supply and demand, seasonal effects, etc. When discovering a business process

activity model from a particular event logs actually the process it happens often that the model depends due to the nature known as concept drift. It is assumed that the process at the beginning of the recorded period is the same as the process at the end of the recorded period. The approach has been implemented in ProM6 [7] and evaluated by analyzing an evolving process.

2.1 Concept Drifts

Concept Drifts [8] refers to situation when the relationship between the input data and the target variable, which the model is trying to changes overtime. Drifts are classified changes into both momentary and permanent. Momentary drifts are nothing but the change will be appear at maximum time and after that it resolves the process. Where as Permanent Drifts are it gradually changes the whole process which disturbs the event logs while running the features. Mainly drifts are classified into 4 types

- (1) **Sudden Drifts:** Sudden drifts are the one kind of drifts that will change the whole process from the starting event to the ending event. For Example recently government of India as declared that all the people of Indian country has to enrolled there aadhar card number it's a mandatory. Because of this whole business process should be change according to the requirement.
- (2) **Gradual Drifts:** Gradual drift sometimes coexist the events at a time for a while time and it resolves the process and again it starts executes. For example a new regulation by the finance ministry of India mandates all banks to collect the aadhar number of the customers to seed with their own account numbers.
- (3) **Recurring Drifts:** Recurring drifts are a set of processes reappears after sometime. Depending upon the market condition it changes.
- (4) **Incremental Drifts:** Incremental drifts are nothing but the process will be slowly changes will occur in the events.

There are some Perspectives changes in the perspectives in the context of business processes

- (1) **Control flow perspective:**-Changes deals with the behavioural and structural changes in a process model. For example if we consider an insurance agency that classifies claims as high (or) low. Depending on the claim if the customer pays last year Rs. 1000 this year he will pay Rs.1500 because of the Organization Decisions.
- (2) **Data Perspective:**-Change in the production and Consumption it may no longer be required to have any such as document approving the client.
- (3) **Resource perspective:**-Changes in resources, their roles, and organizational structure, on the execution of a process. For example, there could have been a change pertaining to who executes an activity. Roles may change and people may change roles. Assumes that change logs are available, modifications of the workflow model are recorded.

2.2 Event Logs

Event logs are characterized by the relation between the activities. The starting point for process mining is an event logs which is a collection of events ,event logs can be of any type like databases , transactions , audit trails etc ,here we will explain the event logs with the example if we take any kind of transaction there we have 3 kinds of attributes Associate resource (Person executes the activity) Transaction resource (start complete, suspend etc..) Data attributes (amount, type of customer). Event logs are special files that record significant events on computer when a program encounters in an computer an error occurs. Whenever these types of events presents, Windows automatically records the events in an event log that you can read by using Event Viewer. Although most of the users might find the details in event logs helpful when troubleshooting problems with Windows and other programs.

The event log I have taken in my project to solve the issue is if want to collect the event we can easily get the logs from our own system otherwise we can download the logs from the online services. They are essential to understand the activities of difficulty systems in an organization, particularly in the case of applications with little user interaction (such as server applications).

2.3 Rapid Miner Tool

Rapid miner is a software platform developed by the company of the same name that provides an integrated environment for machine learning, data mining, business analytics text mining, and predictive analytic . It is used for research business and industrial applications as well as for, rapid prototyping, training, education, and application development. And it supports all steps of the data mining process including results optimization, validation visualization, and Rapid Miner [9] is developed on a business source model which means only the previous version of the software is available under an OSI-certified open source license. And also we have extension called ProM 6 by using this we can visualize the process in a better way of analyzing the process to get the result. This tool is a open source software and it is very useful in data mining, text mining as well as in process mining. It reads the log data as input and passes the event in the processor it moderates the whole system in an optimized way. And also gives the best result form in an visualization mode.

3. Architecture

We propose the framework shown in Figure1. for analyzing concept drifts in process mining. The framework identifies the whole process how we are getting the result. First of all we have to import the particular event logs in an data base server then we will follow the frame work to detect the changes in the environment to recognize the event logs in an appropriate visualization and drift changing in an Prom 6 network to regain the very it is being change in the process.

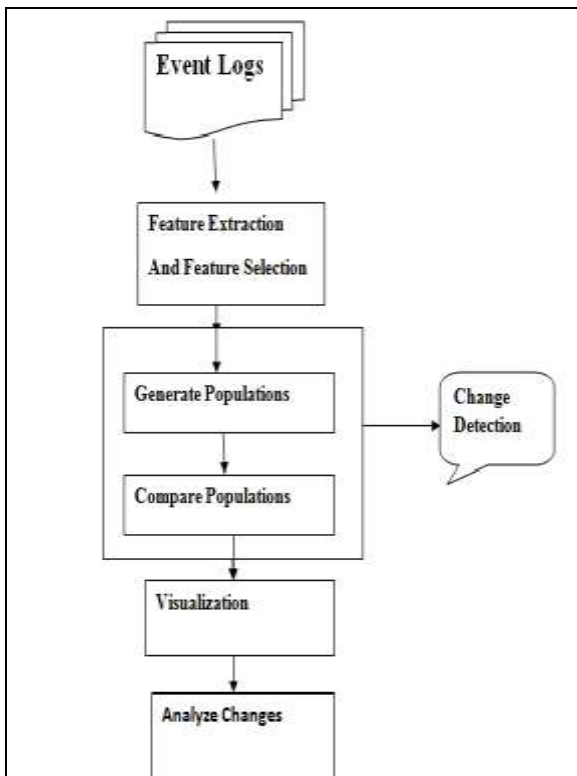


Figure: 1 Framework for executing concept drifts

This framework identifies the 4 modules to detect the drift changes in process mining.

- 1) **Feature extraction and selection:** In This feature extraction and selection step we have to identify what type of process perspectives is suitable like organizational or resource then we have to select any one of them to extract the changes in the environment system.
- 2) **Generating Population and Compare Population:** In this generating populations step we will collect the sample event logs from particular log data and we will generate the populations after that we can see the changes in the process then compare the previous event process.
- 3) **Interactive Visualization:** The results of comparative studies on the populations of the events. And we can visualization Prom [10] the changes what type of drift occurred, by using the ProM6 software.
- 4) **Analyze Changes:** Here we can analyze and identify the exact where the change has been occurred in the form of bar chart we can notice the changes.

4. Implementation

ProM framework works very well in finding out the visualization of the process event logs in an appropriate manner. Below figure 2 shows the process of transforms from one process to another. Process ProM has emerged to be the facto standard for process mining. First we have to download the prom6 extension from the rapid miner tool .Then automatically all the packages will be uploaded from that prom framework we have to read the log data and it should be connected to the multiplier, in that all the operations are involved and it performs to generate the results. And now we have to import two more packages they are Execute the read

file and also the alpha miner operation step to get the relevant features.

This step involves many actions to get the source related process mining drifts from the event logs, and also it follows where the change has been occurring and it tracks the recourses whether it is data perspective or the organizational perspective.

We will notice all the changes in the form of a bar chart and visualizes very clearly with the help of a x-axis and the y-axis in the prom framework according to the detect in the changes in process mining.

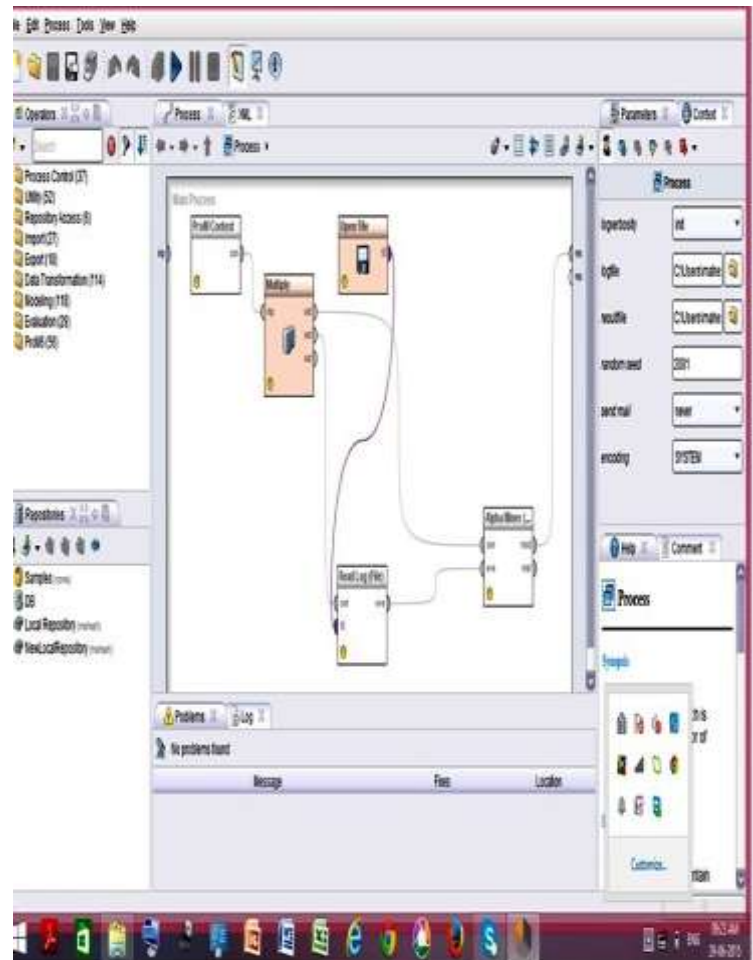


Figure: 2 Executing the process in Prom6 Framework

5. Results

Finally the result of dealing with concept drift in process mining using event logs appears has in the form of x-axis and the y- axis . How the events from one process to another process is been changing and it will exactly localize the change and it visualizes in the form of bar chart the below figure 3 shows the visualization part of the concept drifts. On the x-axis chart we have the change detection over and on the y-axis Change percentage. This can be gradually increases or decreases depending upon the various event logs in the process mining.

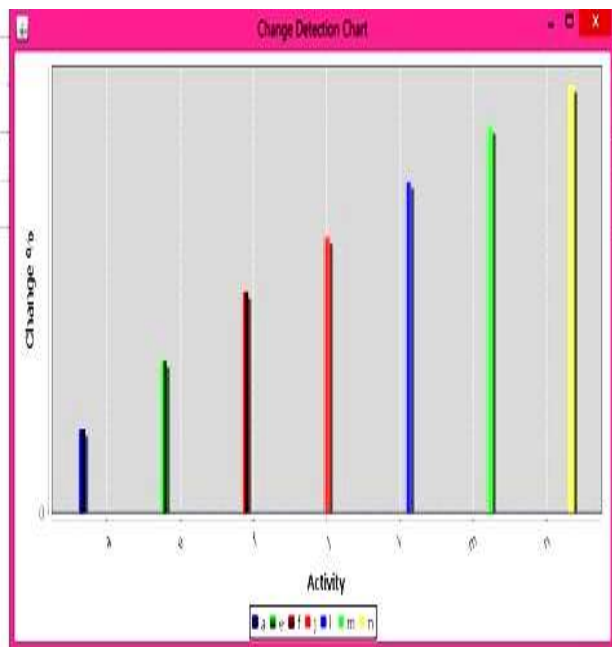


Figure: 3 Visualization of the drifts in Prom6

6. Conclusion

In this paper, the main purpose of the project is to describe the concept drifts in process mining using event logs and this process depends on the event logs. A process may change suddenly or gradually for that purpose it is important to track the process where there is a change. A platform is developed in order to find out the drifts in the event logs using the feature technique and the rapid miner tool. By this we can visualize the changes in the events. Mainly it is used to effectively identify the changes in the events.

References

- [1] Chris J. Turner, Ashutosh Tiwari, Richard Olaiya, Yuchun Xu, (2012) "Process mining: from theory to practice", Business Process Management Journal, Vol. 18 Iss: 3, pp.493 – 512
- [2] Aalst,W. van der, Weijters, A., & Maruster, L. (2004). Workflow Mining: Discovering Process Models from Event Logs. IEEE Transactions on Knowledge and Data Engineering, 16 (9), 1128–1142
- [3] Jin.Gu kang;Dept.of Ind & Syst. Eng., Gyeongsang Nat. Univ., Jinju ; Kwan Hee Han; Convergence and hybrid information technology 008. ICCIT '08. Third International Conference on (Volume:1)
- [4] Osephine E. Olson Associate Dean and Professor of Business Administration The Joseph M. Katz Graduate School of Business University of Pittsburgh Pittsburgh PA 15260 USA volume-3
- [5] Tobias Bucher and Anke Gericke, process centric business intelligence, University of St Gallen, St Gallen, Switzerland, and Stefan Sig SAP AG, Walldorf, Germany.
- [6] W.M.P. van der Aalst, J. Desel, and A. Oberweis, editor business process management , and Empirical Studies , volume 1806 of Lecture Notes in Computer Science. Springer Verlag, Berlin, 2000.

- [7] W.M.P. van der Aalst and K.M. van Hee. Work on Management: Models, Methods, and System MIT press, Cambridge, MA, 2002.
- [8] Bose, R.P.J.C. ; Dept. of Math. & Comput. Sci., Eindhoven Univ. of Technol., Eindhoven, Netherlands ; van der Aalst, W.M.P. ; Zliobaite, I. ; Pechenizkiy, M. Handling with concept drifts Neural Networks and Learning Systems, IEEE Transactions on (Volume:25 , Issue: 1)
- [9] M.Santhanakumar, C.Christopher Columbu, Web Usage Based Analysis Of Web Pages Using Rapidminer , wseas transactions on computers (volume:7 , issue: 6
- [10] B.F. van Dongen, A.K.A. de Medeiros, H.M.W. Verbeek, A.J.M.M. Weijters, and W.M.P. van der Aals, The ProM framework: A new era in process mining tool support 16 (9), 1128–1142

Acknowledgments

I feel privileged to offer my sincere thanks and deep sense of gratitude to **Dr. C. Kiranmai** (Principal & Professor), **Dr. B. Raveendra Babu** (Professor and Head of the Department, CSE Department), **Mrs.R.Vasavi** (Project Coordinator, CSE department, Assistant Professor), of VNRVJIET for extending their co-operation in doing this project. I express my sincere gratitude to my Project Guide, **Mr. T. Sunil Kumar**, (Associate Professor, CSE department), of VNRVJIET for valuable time and guidance, throughout the M.Tech Project. My special thanks to all the Faculty members of the Department for their valuable advices at every stage.

Author Profile



Maddu . Uma Lavanya¹ received the B.TECH degree from the MALLAREDDY ENGINEERING COLLEGE FOR WOMEN under the JNTUH in Computer Science Engineering in the year 2012 , and pursuing M.TECH degree from VNR VJIET COLLEGE (Autonomous), Specialization in Software Engineering. Participate in various fests , and also Participated in INDAS International Conferences.



Mr. Sunil Kumar Talluri² pursuing Ph.D. in Data Mining area from JNTU-A, Ananthapur. Did the Post graduation, M.Tech in Computer Science and Engineering from Osmania University, Hyderabad. has been working in the Department of Computer Science and Engineering, in VNR Vignana Jyothi Institute of Engineering and Technology, affiliated to Jawaharlal Nehru Technological University. Currently as an Associate Professor in the Computer Science and Engineering department.