International Journal of Engineering and Techniques - Volume 4 Issue 2, Mar – Apr 2018

RESEARCH ARTICLE OPEN ACCESS

DISCOVERING CHANGES IN UNPREDICTIVE BUSINESS PROCESS

SanjanaBaratam., Vaishnavi K., Varsha Roy., Nivethitha.K.,
Department of Computer Science and Engineering,
Panimalar Engineering College,Chennai

Abstract:

A business process is a collection of linked tasks for which find their end in the delivery of a service or product to a client.A business process has also been defined as a set of activities and tasks that, once completed to achieve the defined business outcomes. It can be achieved by analyzing control flow, data flow, and resource perspectives. Nowadays customers expect organisations to be flexible and adapt the changes. Extreme variations in supply and demand, natural calamities, seasonal effects, disasters and so on, are forcing organizations to change their processes. The changes can be classified into two types: Sudden drifts and Gradual drifts. We detect the changes in business process by analysing event logs extracted from the systems that support the execution of the process. Event logs consists of sequence of labels which contains number of traces and it represents the execution of one activity.

Keywords — **Drifts, sudden drifts, gradual drifts, traces.**

I.INTRODUCTION

This project deals with the detection of the drifts between the business processes andclassify them into sudden and gradual drifts.It is an automated method for detecting drifts. The gradual drift detection method relies on the assumption that gradual drift is delimited by two consecutive sudden drifts, such that the distribution of runs in-between these two drifts is a linear mixture of the distributions of runs before the first drift and the second drift. The sudden drift detection method is to identify a time when there is a statistically significant difference between the observed behaviour before and after this point. A gradual drift is not recognized easily. At first it affects some cases and consequently it affects the rest of the cases hence making the drift visible. On the other hand sudden drifts are recognized easily. It occurs abruptly at a time.

II.RELATED WORKS

1.[3]R. P. Jagadeesh Chandra Bose, Wil M. P. van der Aalst, Indr˙e Žliobaite, and Mykola Pechenizkiy presented that most business activities change over time, concurrent process mining techniques incline to analyze these processes as if they are in a steady state. Processes may change suddenly or gradually. The drift may be periodic (e.g., because of seasonal influences) or one-of-a-kind (e.g., the effects of new legislation). For the process management, it is pivotal to discover and understand such concept drifts in processes. This paper presents a generic framework and specific techniques to detect when a process changes and to localize the parts of the process that have changed. Different features are proposed to characterize relationships among activities. These features are used to discover differences between successive populations. approach has been implemented as a plug-in of the ProM process mining framework and has been evaluated using bothsimulated event data exhibiting controlled concept drifts and real-life even data from a Dutch municipality.

2.Marwan Hassani∗, Sergio Siccha†, Florian Richter‡ andThomas Seidl§ Data Management and Data Exploration Group RWTH Aachen University, Germany Process mining is an emerging research area that brings the well-established data mining solutions to the challenging business process modeling problems. Mining streams of business processes in the real time as they are generated is a necessity to obtain an instant knowledge from big process data.In this paper, we introduce an efficient approach for exploring and counting process fragments from a stream of events infer a process model using the Heuristics Miner algorithm. Our novel approach, called StrProM, builds prefix-trees to extract sequential patterns of events from the stream. StrProM uses a batch-based approach to continuously update and prune these prefix-trees. The models are generated from those trees after applying a decaying mechanism over their statistics. The extensive experimental evaluation demonstrates the superiority of our approach over a state-of- the-art technique in terms of execution time using a real dataset, while delivering models of a comparable quality.

3.[18]Georg Krempl , University Magdeburg, Germany proposed that Every day, huge volumes of sensory, transactional, and web data are continuously generated as streams, which need to be analyzed online as they arrive. Streaming data can be considered as one of the main sources of what is called big data. While predictive modeling for data streams and big data have received a lot of attention over the last decade, many research approaches aretypically designed for well-behaved controlled problem settings, overlookingimportant challenges imposed by real-world applications. This article presents adiscussion on eight open challenges for data stream mining. Our goal is to identifygaps between current research and meaningful applications, highlight openproblems, and define new application-relevant research directions for data streammining. The identified challenges cover the full cycle of knowledge discovery andinvolve such problems as: protecting data privacy, dealing with legacy systems, handling incomplete and delayed information, analysis of complex data, andevaluation of stream mining algorithms. The resulting analysis is illustrated bypractical applications and provides general suggestions concerning lines of future research in data stream mining.

4.Asef Pourmasoumi and Ebrahim Bagheri Department of Electrical and Computer Engineering, Ryerson University, Canada Laboratory for Systems, Software and Semantics (LS3) proposed that most organizations spend a lot of resources for implementing, analyzing and managing their business process models. Hence, tools or techniques that can help managers reach these goals are desirable. Process mining is a new research agenda, which helps managers gain more insight about their organization's processes. The main goal of process mining is to extract process-centric knowledge from event logs of existing information system of organizations. Process mining can be considered to be the x-ray machine, which shows the reality that occurs within the organization. In many cases, the process that is executed in an organization can have many differences with the process that is expected to be running. This can be because of several reasons such as management changes, infractions and so on.Process mining extracts valuable knowledge for managers and brings transparency for them by analyzing event logs that are stored in the database of information systems of organizations.

5.Albert Bifet Ricard Gavald`a Universitat Polit`ecnica de Catalunya proposed and illustrated a method for developing algorithms that can adaptively learn from data streams that change over time. As an example, we take Hoeffding Tree, an incremental decision tree inducer for data streams, and use as a basis it to build two new methods that can deal with distribution and concept drift: a sliding window-basedalgorithm, HoeffdingWindow Tree, and an adaptive method, Hoeffding Adaptive Tree. Our methods are based on using change detectors and estimator modules at the right places; we choose implementations with theoretical guarantees in order to extend such guarantees to the resulting adaptive learning algorithm. A main advantage of our methods is that they require no guess about how fast or how often the stream will change; other methods typically have several user-defined parameters to this effect. In our experiments, the new methods never do worse, and in some cases do much better, than CVFDT, a well-known method for tree induction on data streams with drift.

III.EXISTING SYSTEM

In existing systems for business process drift detection extract patterns characterizing each trace. One possible feature is for example that task A occurs before task B in the trace, or that B occurs more than once in the trace. To achieve a suitable level of accuracy, existing techniques either explore large feature spaces automatically or they require the users themselves to identify a specific set of features that are likely to characterize the drift. Another drawback is that it detects sudden drifts or gradual drifts but not both at a time.

IV.PROPOSED SYSTEM

It is automated method for finding sudden and gradual drifts in the business process in a single framework. An analysis on synthetic logs proves that the method accurately detects the process changes and its composition thereby, outperforming the baselines for both sudden and gradual drifts. A separate evaluation on event logs demonstrates the method’s capability to detect drifts that correspond to the user recognisable changes in the drifts. It increases the scalability. It addresses the concept drift and classifies the drifts. The event logs that are generated by the employee is checked with the synthetic logs and the percentage of the drift is found out. The reason for the drift may be the absence or inability of the employee to outperform the work. This is found out by an analysis conducted when a sudden drift happens. The data of the employees are encrypted and can be viewed only by the employees but not by any other people.

V.ARCHITECTURAL DIAGRAM

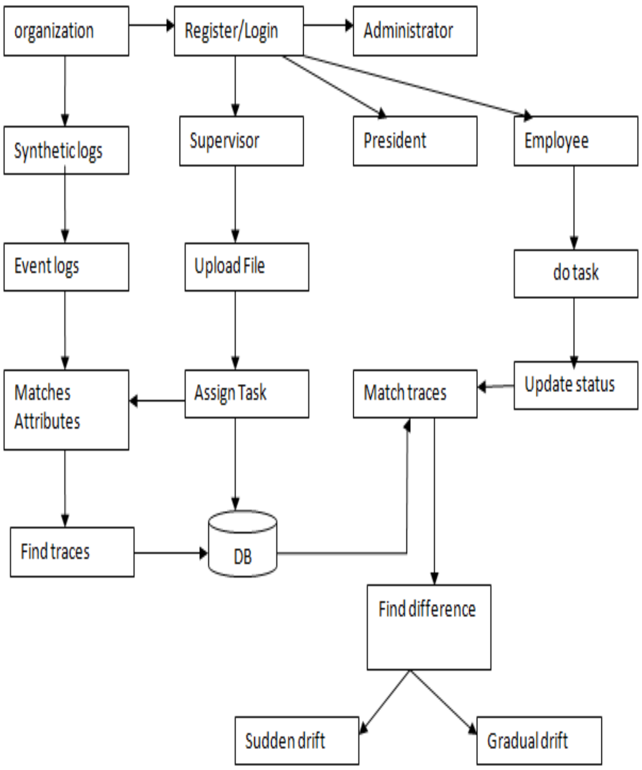

Fig.1-The architecture diagram for finding drifts in business process

VI.GROUP KEY ALLOCATION

Supervisorhas an initial level Registration Process at the web end. The supervisor provides their own personal information for this process at the time the group key created. Likewise employee registration process at the process supervisor is allocated to the employee and the group key is shared. After registration supervisor can upload a files to the server. The supervisor can assign a task to the employee and supervisor also updates the task of the employee.

VII.EMPLOYEE AUTHENTICATION

After assigning a task, the system matches the attributes of current task to existing task attributes in an event logs. If the attributes matches, the system will collect the execution traces and assign them to current process. Employee Registration time employee id is generated. Employee login into the application. At the time of login employee id, password can be validated by the administrator. After login Employee can able to view their task, starting date, ending date, task id, and task completed percentage. Employee do their task and completed the task status are maintained.

VIII. ZERO KNOWLEDGE PROOF

Administrator and President can able to view the employee details. Administrator can able to edit or update the employee details. President can able to view the employee task and completed status. President also view who is the high performer of the day and between two dates. Zero-knowledge proof a method by which one party can prove to another party that a given statement is true, without conveying any information apart from the fact that the statement is indeed true.

IX.DRIFT DETECTION

Each task contains n number of execution traces. Traces means collects information about what is happening in your task. The employees do their tasks and update their status on daily basis. Every time the system matches the current traces to previous execution traces. The system identify the significant difference between traces, the will detect the sudden drift and send the information to the supervisor. Sudden drifts are the one kind of drifts that will change the whole process from the starting event to the ending event. The system will find the low level difference the will detect the gradual drift and send the information to the supervisor.

X.RESULT

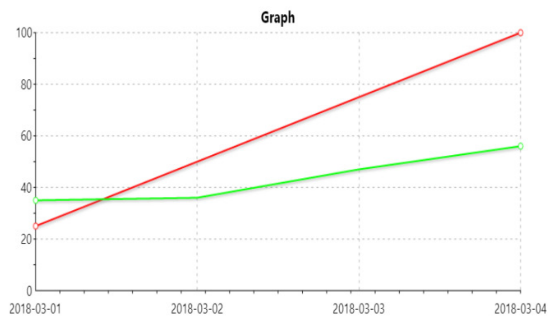

Fig. 2-The final graph when the drift is found out.

XI.CONCLUSION

Thus the drifts in the business process has been detected. The drifts detected are represented in graphs. The graph compares the event logs with the synthetic logs. The data of all the employees are encrypted using diffie helman algorithm. The supervisor is solely responsible for dividing the work among employees and he checks the progress of work. The advantage is the drifts are detected beforehand rather than not checking until the dead-end to find the drifts. This helps in preventing delay in work.

XII.REFERENCES

[1] R. J. C. Bose, W. M. van der Aalst, I. _Zliobait_ e, and M. Pechenizkiy, "Handling concept drift in process mining," in Proc. Int. Conf. Adv. Inf. Syst. Eng., 2011, pp. 391–405.

[2] J. Carmona and R. Gavalda, "Online techniques for dealing with concept drift in process mining," in Proc. Int. Symp. Intell. Data Anal., 2012, pp. 90–102.

[3] R. J. C. Bose, W. M. van der Aalst, I. Zliobaite, and M. Pechenizkiy, "Dealing with concept drifts in process mining," IEEE Trans. Neural Netw. Learn. Syst., vol. 25, no. 1, pp. 154–171, Jan. 2014.

[4] J. Martjushev, R. Bose, and W. P. van der Aalst, "Change point detection and dealing with gradual and multi-order dynamics in process mining," in Proc. Int. Conf. Bus. Informat. Res., 2015, pp. 161–178.

[5] A. Maaradji, M. Dumas, M. La Rosa, and A. Ostovar, "Fast and accurate business process drift detection," in Proc. Int. Conf. Bus. Process Manage., 2015, pp. 406–422.

[6] J. Gama, I. _Zliobait_ e, A. Bifet, M. Pechenizkiy, and A. Bouchachia, "A survey on concept drift adaptation," ACM Comput. Surveys, vol. 46, no. 4, 2014, Art. no. 44.

[7] L. L. Minku, A. P. White, and X. Yao, "The impact of diversity on online ensemble learning in the presence of concept drift," IEEE Trans. Knowl. Data Eng., vol. 22, no. 5, pp. 730–742, May 2010.

[8] R. Klinkenberg and T. Joachims, "Detecting concept drift with support vector machines," in Proc. 17th Int. Conf. Mach. Learn., 2000, pp. 487–494.

[9] M. Baena-Garc_ıa, J. del Campo-_ Avila, R. Fidalgo, A. Bifet, R. Gavald_a, and R. Morales-Bueno, "Early drift detection method," in Proc. 4th ECML PKDD Int. Workshop Knowl. Discovery Data Streams, 2006, pp. 77–86.

[10] D. Ienco, A. Bifet, B. Pfahringer, and P. Poncelet, "Change detection in categorical evolving data streams," in Proc. 29th Annu. ACM Symp. Appl. Comput., 2014, pp. 792–797.

[11] D. Kifer, S. Ben-David, and J. Gehrke, "Detecting change in data streams,"
in Proc. 30th Int. Conf. Very Large Data Bases, 2004, pp. 180–191.
[12] A. Bifet and R. Gavald_a, "Learning from time-changing data with adaptive
windowing," in Proc. SIAM Int. Conf. Data Mining, 2007, pp. 443–448.
[13] J. Gama, R. Sebasti~ao, and P. P. Rodrigues, "Issues in evaluation of
stream learning algorithms," in Proc. 15th ACM SIGKDD Int. Conf. Knowl.
Discovery Data Mining, 2009, pp. 329–338.
[14] R. Accorsi and T. Stocker, "Discovering workflow changes with time-based
trace clustering," in Proc. Int. Symp. Data-Driven Process Discovery Anal., 2012,
pp. 154–168.
[15] A. Burattin, A. Sperduti, and W. M. van der Aalst, "Control-flow discovery
from event streams," in Proc. IEEE Congr. Evol. Comput., 2014, pp. 2420–2427.
[16] A. Burattin, M. Cimitile, F. M. Maggi, and A. Sperduti, "Online discovery of declarative process models from event streams," IEEE Trans. Services
Comput., vol. 8, no. 6, pp. 833–846, Nov./Dec. 2015.
[17] A. Ostovar, A. Maaradji, M. La Rosa, A. H. ter Hofstede, and B. F. van
Dongen, "Detecting drift from event streams of unpredictable business
processes," in Proc. 35th Int. Conf. Conceptual Model., 2016, pp. 330–346.
[18]Georg Krempl , University Magdeburg, Germany georg.krempl@iti.cs.uni-
magdeburg.de,"Open Challenges for Data Stream Mining Research"