

Dealing With the Concept of Deviation in the Mining Process

Dr.K.Sathish Kumar

Assistant Professor,
Department of Computer Science
(Informatics), Sri Gayathri Degree &
PG College, Mulugu Road,
Hanamkonda, Warangal - 506001.

Kiran Kumar Dharanikota

Associate Professor,
Department of Computer Science &
Engineering,
Gurunanak Institute of Technical
Campus, Hyderabad.

G.Rajesh

Assistant Professor,
Dept of Information Technology,
Gurunanak Institutions Technical
Campus, Hyderabad.

ABSTRACT:

Most business processes change over time, and contemporary practices in the mining process, however, they are in a constant state of analysis processes. A process that can change suddenly or gradually. It is drifting in the league, or one of its kind (for example, the new law) might be (for example, due to seasonal effects) is. Perform the procedure, and it is necessary to understand the intentions of the concept of practical presentations. This paper presents the process of transformation and the process to find out the general framework and specific techniques to localize the changed components. The proposed measures describe the relationship between different features. It is to find the differences in the characteristics of the population.

I.INTRODUCTION:

In this process, it is stable and sufficient, for example, and the implications of the event log entry, ITI can be used to analyze high-quality presentation style, and verification of compliance, and it is possible to find a signal. Unfortunately, a lot of activities are not in a state of instability. In the dynamics of the market today, IT companies need to cut costs and improve the performance of operations to make faster. In this paper, we, mining, and software to the process, which introduced the concept of change in the analysis of the records are in the lower region, the process is based.

We acknowledge the event log for the feature set of the change and the proposed methods to explore the result of changes in the process. In contrast to the situations that arise during the process of the evolution of the concept of motion detection can be handled effectively. Or to support any point of time to get an accurate idea of the performance of operational processes and improve operating procedures.

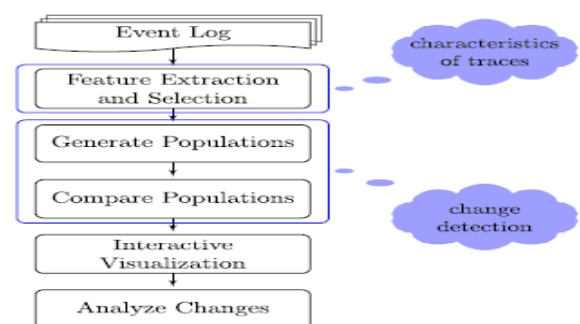


Fig 1: System Architecture

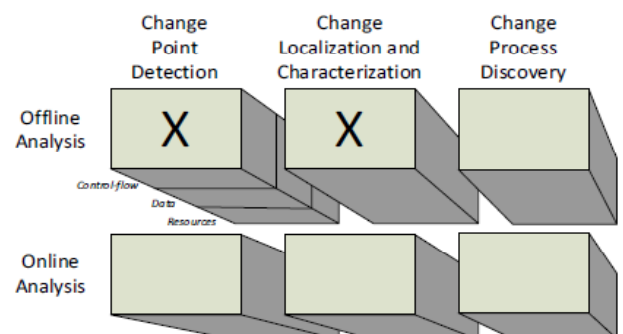


Fig 2: Different dimensions of concept drift analysis in process mining.

II.IMPLEMNTATION:

Concepts, have realized the paper PriM6 movement within the concept of plug-ins. Environmental aspects of the ceremony, all types of mining operations and mining techniques to provide a common basis and be able to import and export, and filter event logs (process) models to analyze and visualize the results. Over the years, the party has emerged as the de facto standard for the mining operations. Dam flow, and ease of implementation of the proposed additional items in the sense of all the steps (for example, easily add new features), and the extension. In support of the hypothesis that the perception of the importance of the test plot down to the bridge as possible.

III.SYSTEM PREMELIRIES:

A.PROCESS MINING:

Data mining and modeling of business processes and mining operations [6] to serve as a bridge between. Commercial operations data sources (for example, audit logs, database, and transaction records) in a variety of holiday songs. It is a set of events to the event log, and a starting point for the process of mining. Some of the name of the activity has been described, the events of treatment cases (often known cases) relate to assume that. For example, the premium in the process. Therefore, the range of activities and is often referred to as the impact of an unprecedented process.

B.CONCEPT DRIFT:

The concept drift [12] The goal of machine learning and data entry data and variable over time in unexpected ways when trying to predict the pattern refers to the mining of the cases, the relationship between changes. Therefore, the prediction accuracy may degrade over time. To prevent this, you will always be able to update itself with new data, predictive models to adapt to the Internet. 1) to receive the new data, and the data stream is usually more than one endless as follows shave preparations. 2) Make predictions. 3) response (actual value) can be obtained and 4) to update the predictive models. Ask the expected: 1 operating conditions Q) design concept and responds to movement (if necessary), as soon as possible, the amendment 2) once and adapt to changes in the difference between the noise of intentions, but it is a strong voice. And 3) the arrival of a limited amount of memory for data storage, and use less time and labor.

C.NATURE OF DRIFTS:

How to activate the change, you can change the immediate and permanent classification. Changes in the short term at the moment of change and survival [31] permanent, except in cases that would affect a very small. In this paper, we can make permanent changes often do not meet the moment of transition from outliers in the data mining, data.2 shift / compatible with the concept of noise can not be detected by sight. It can cause a deviation in the concept of change (practice process) is in. This is shown in the picture. 3, we have identified four categories of intentions.

1)Sudden drift:

This corresponds to a substitution of an existing process M1 with a new process M2, as shown in Fig. 3(a). M1 ceases to exist from the moment of substitution. In other words, all cases (process instances) from the instant of substitution emanate from M2. This class of drifts is typically seen in scenarios such as emergencies, crisis situations, and change of law. As an example, a new regulation by the finance ministry of India mandates all banks to procure and report the customer's personal account number in their transactions.

2)Gradual drift:

This refers to the scenario, as shown in Fig. 3(b) where a current process M1 is replaced with a new process M2. Unlike the sudden drift, here both processes coexist for some time with M1 discontinued gradually. For example, a supply chain organization might introduce a new delivery process. This process is, however, applicable only for orders taken henceforth. All previous orders still have to follow the former delivery process.

3)Recurring drift:

This corresponds to the scenario where a set of processes reappear after some time (substituted back and forth), as shown in Fig. 3(c). It is quite natural to observe such a phenomenon with processes having a seasonal influence. For example, a travel agency might deploy a different process to attract customers during Christmas period. The recurrence of processes may be periodic or nonperiodic. An example of a nonperiodic recurrence is the deployment of a process subjected to market conditions. The point of deployment and the duration of deployment are both dependent on external factors (here, the market conditions). Periodic drifts may be caused by seasonal effects.

IV. IMPLEMENTATION:

1) Change-pattern specific features: In this paper, we presented very generic features (based on follows/precedes relation). These features are neither complete nor sufficient to detect all classes of changes. An important direction of research would be to define features catering to different classes of changes and investigate their effectiveness.

A taxonomy/classification of change patterns and the appropriate features for detecting changes with respect to those patterns are needed. 2) Feature selection: The feature sets presented in this paper result in a large number of features. For example, the activity relation count feature type generates $3 \times |A|$ features whereas the WC and J measure generate $|A|^2$ features (corresponding to all activity pairs). On the one hand, such high dimensionality makes analysis intractable for most real-life logs. On the other hand, changes being typically concentrated in a small region of a process make it unnecessary to consider all features. There is a need for tailored dimensionality reduction techniques [44], [45] that can efficiently select the most appropriate features.

3) Holistic approaches:

In this paper, we discussed ideas on change detection and localization in the context of sudden and gradual changes to the control-flow perspective of a process. As mentioned in Section IV, the data and resource perspectives are also, however, equally important. Features and techniques that can enable the detection of changes in these other perspectives need to be discovered. Furthermore, there could be instances where more than one perspective (e.g., both control and resource) change simultaneously. Hybrid approaches considering all aspects of change holistically need to be developed.

4) Recurring drifts:

When dealing with recurring drifts, in addition to change point detection and change localization, it is important to identify the variant(s) that recur. This requires robust metrics to assess the similarity between process variants and/or event logs.

5) Change process discovery:

As mentioned earlier, after detecting the change points and the regions of change, it is necessary to put them together in perspective. Organizations would be interested in discovering the evolution of change (e.g., as an animation depicting how the process has changed/evolved over time). In addition, there are other applications such as deriving a configurable model for the process variants. A configurable process model describes a family of similar process models.

The process variants discovered using concept drift can be merged to derive a configurable process model.

IV..CONCLUSION:

In this paper, we succumb to distractions in the mining process, and the process of change that in any analysis based on the event log. We acknowledge the event log for the feature set of the change and the proposed methods to explore the result of changes in the process. Our preliminary findings to determine the intentions of the change process that can be handled effectively concept refers to a variety of situations. Change, the event log can be broken down and analyzed. And monitoring and evaluation of efforts to deal with the first step in the process of change. To change the angle of the flow control gradually drifts seen suddenly it appears. Therefore, if you take into account the starting point of a new branch in the mining process, our analysis is taken, and there are still many challenges that need to be addressed to take for. There are some of these challenges.

REFERENCES:

- [1] (2010). All-in-one Permit for Physical Aspects: (Omgevingsvergunning) in a Nutshell [Online]. Available: [http://www. answersforbusiness.nl/regulation/all-in-one-permit-physical-aspects](http://www.answersforbusiness.nl/regulation/all-in-one-permit-physical-aspects).
- [2] United States Code. (2002, Jul.). Sarbanes-Oxley Act of 2002, PL 107-204, 116 Stat 745 [Online]. Available: <http://files.findlaw.com/news.findlaw.com/cnn/docs/gwbush/sarbanesoxley072302.pdf>.
- [3] W. M. P. van der Aalst, M. Rosemann, and M. Dumas, "Deadline-based escalation in process-aware information systems," Decision Support Syst., vol. 43, no. 2, pp. 492–511, 2011. 170 IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS, VOL. 25, NO. 1, JANUARY 2014.
- [4] M. Dumas, W. M. P. van der Aalst, and A. H. M. Ter Hofstede, Process- Aware Information Systems: Bridging People and Software Through Process Technology. New York, NY, USA: Wiley, 2005.
- [5] W. M. P. van der Aalst and K. M. van Hee, Work-flow Management: Models, Methods, and Systems. Cambridge, MA, USA: MIT Press, 2004.



International Journal of Research in Advanced Computer Science Engineering

A Peer Reviewed Open Access International Journal

www.ijracse.com

- [6] W. M. P. van der Aalst, Process Mining: Discovery, Conformance and Enhancement of Business Processes. New York, NY, USA: Springer-Verlag, 2011.
- [7] B. F. van Dongen and W. M. P. van der Aalst, "A meta model for process mining data," in Proc. CAiSE Workshops (EMOI-INTEROP Workshop), vol. 2. 2005, pp. 309–320.
- [8] C. W. Günther, (2009). XES Standard Definition [Online]. Available: <http://www.xes-standard.org>
- [9] F. Daniel, S. Dustdar, and K. Barkaoui, "Process mining manifesto," in BPM 2011 Workshops, vol. 99. New York, NY, USA: Springer-Verlag, 2011, pp. 169–194.
- [10] R. P. J. C. Bose, W. M. P. van der Aalst, I. Žliobaitė, and M. Pechenizkiy, "Handling concept drift in process mining," in Proc. Int. CAiSE, 2011, pp. 391–405.
- [11] J. Carmona and R. Gavalda, "Online techniques for dealing with concept drift in process mining," in Proc. Int. Conf. IDA, 2012, pp. 90–102.
- [12] J. Schlimmer and R. Granger, "Beyond incremental processing: Tracking concept drift," in Proc. 15th Nat. Conf. Artif. Intell., vol. 1. 1986, pp. 502–507.
- [13] A. Bifet and R. Kirkby. (2011). Data Stream Mining: A Practical Approach, University of Waikato, Waikato, New Zealand [Online]. Available: <http://www.cs.waikato.ac.nz/~abifet/MOA/StreamMining.pdf>
- [14] I. Žliobaitė, "Learning under concept drift: An Overview," CoRR, vol. abs/1010.4784, 2010 [Online]. Available: <http://arxiv.org/abs/1010.4784>
- [15] J. Gama, P. Medas, G. Castillo, and P. Rodrigues, "Learning with drift detection," in Proc. SBIA, 2004, pp. 286–295.