Process Drift: User guide

The program is the implementation of the process drift detection method in:
Lu, Y., Chen, Q., & Poon, S. (2021). A Robust and Accurate Approach to Detect Process Drifts from Event Streams arXiv:2103.10749.

Note: Only events with "complete" lifecycle information will be processed by the program. If your log does not contain such information, please add the "lifecycle: transition" to the events first.

The guide is for Windows only, the command to run the program on Linux is slightly different.

Firstly, please download the ProcessDrift.zip file and decompress it.

How to replicate our results

In this part, we show how to run all 4148 artificial logs in the 24 different parameter settings presented in the paper.

1. Make sure you have a folder named "Evaluation" in the "bin" folder
2. In the cmd, enter the "bin" folder
3. Simply run the program using the following command:
   java -Xmx16G -cp ".;.\lib*" Test.Main evaluateAll add

   -Xmx16G: heap space (In our experiment, we set it to 16Gb)
   add: choose between "add", "remove", "noise-free" (type "add" for adding events, "remove" for removing events, "noise-free" for noise-free event logs. For example, if you type "add", the program will run all the artificial event logs with added events as noises.)

By default, the program will detect the number of available CPU threads on your computer. If you want to customize the number of threads used, use the following command:

java -Xmx16G -cp ".;.\lib*" Test.Main evaluateAll add numOfThreads

4. You will see the results in the "Evaluation" folder

How to run experiment on a user-specified parameter setting

In this part, we show how to run all 4148 artificial logs in your specified parameter setting.

1. Make sure you have a folder named "Evaluation" in the "bin" folder
2. In the cmd, enter the "bin" folder
3. Simply run the program using the following command:
   java -Xmx16G -cp ".;.\lib*" Test.Main evaluate add numOfConsecutiveTests windowSize

   -Xmx16G: heap space (In our experiment, we set it to 16Gb)
   add: choose between "add", "remove", "noise-free" (type "add" for adding events, "remove" for removing events, "noise-free" for noise-free event logs. For example, if you type "add", the program will run all the artificial event logs with added events as noises.)
   numOfConsecutiveTests: For example, if you type 2, it means windowSize / 2 of consecutive statistical tests will be performed in one direction

By default, the program will detect the number of available CPU threads on your computer. If you want to customize the number of threads used, use the following command:

java -Xmx16G -cp ".;.\lib*" Test.Main evaluate add numOfConsecutiveTests windowSize numOfThreads

4. You will see the results in the "Evaluation" folder

How to apply the method to other datasets

In this section, we show how to run our method on a log provided by the user.

Note: Only events with "complete" lifecycle information will be processed by the program. If your log does not contain such information, please add the "lifecycle: transition" to the events first.

1. In the cmd, enter the "bin" folder
2. Simply run the program using the following command:
   java -Xmx16G -cp ".;.\lib*" Test.Main apply path_to_log numOfConsecutiveTests windowSize

   -Xmx16G: heap space (In our experiment, we set it to 16Gb)
   numOfConsecutiveTests: For example, if you type 2, it means windowSize / 2 of consecutive statistical tests will be performed in one direction

As described in our paper, we recommend setting numOfConsecutiveTests to 2

By default, the program will detect the number of available CPU threads on your computer. If you want to customize the number of threads used, use the following command:

   java -Xmx16G -cp ".;.\lib*" Test.Main apply path_to_log numOfConsecutiveTests windowSize numOfThreads

3. You will see the results in the console

If you want to replicate the results based on the BPIC-2020 data, simply use the following settings:

EventLog	Path_to_log	numOfConsecutiveTests	windowSize
Requests for Payment	request.mxml	2	2000
Domestic Declarations	domestic.mxml	2	2000
Prepaid Travel Cost	prepaid.xes	2	1500
International Declarations	international.mxml	2	2000
Travel Permits	permit.xes	2	2000

*We preprocessed the logs by adding a "complete" lifecycle transition to the events.