

**Filtering out Infrequent Behaviour from Business Process Event Logs**

**Abstract:**

In the era of “big data”, one of the key challenges is to analyze large amounts of data collected in meaningful and scalable ways. The field of process mining is concerned with the analysis of data that is of a particular nature, namely data that results from the execution of business processes. The analysis of such data can be negatively influenced by the presence of outliers, which reflect infrequent behavior or “noise”. In process discovery, where the objective is to automatically extract a process model from the data, this may result in rarely travelled pathways that clutter the process model. This paper presents an automated technique to the removal of infrequent behavior from event logs. The proposed technique is evaluated in detail and it is shown that its application in conjunction with certain existing process discovery algorithms significantly improves the quality of the discovered process models and that it scales well to large datasets.