

УДК 004.891

Koage Samuel

Ternopil Ivan Puluj National Technical University, Ukraine

MAPPING DATA SOURCES TO XES IN A GENERIC WAY

Коаге Самуель

МЕТОД ВІДОБРАЖЕННЯ ДЖЕРЕЛ ДАНИХ У ФОРМАТ XES

Information systems are taking a prominent place in today's business process execution. Since most systems are complex, enterprise-wide systems, very few users, if any, have a clear and complete view of the overall process. In the area of process mining several techniques have been developed to reverse engineer information about a process from a recording of its execution. To apply process mining analysis on process-aware information systems, an event log is required. An event log contains information about cases and the events that are executed on them.

Although many systems produce event logs, most systems use their own event log format. Furthermore, the information contained in these event logs is not always suitable for process mining. However, since much data is stored in the data storage of the information system, it is often possible to reconstruct an event log that can be used for process mining. Extracting this information from the business data is a time consuming task and requires domain knowledge. The domain knowledge required to define the conversion is most likely held by people from business, e.g. business analysts, since they know or investigate the business processes and their integration with technology. In most cases business analysts have no or limited programming knowledge. Currently there is no tool available that supports the extraction of an event log from a data source that doesn't require programming.

This paper discusses important aspects to consider when defining a conversion to an event log. The decisions made in the conversion definition influence the process mining results to a large extent. Defining a correct conversion for the specific process mining project at hand is therefore crucial for the success of the project. A framework to store aspects of such a conversion is also developed in this thesis. In this framework the extraction of traces and events as well as their attributes can be defined. An application prototype, called 'XES Mapper' or 'XESMa', that uses this conversion framework is build.

The XES Mapper application guides the definition of a conversion. The conversion can be defined without the need to program. The application can also execute the conversion on the data source, producing an event log in the MXML or XES event log format. This enables a business analyst to define and execute the conversion on their own. The application has been tested with two case studies. This has shown that many different data source structures can be accessed and converted.

References

1. Wikipedia. List of ERP software packages — Wikipedia, The Free Encyclopedia, 2010. [Online; accessed 6-January-2010].
2. C. Gunther. Process Mining in Flexible Environments. PhD thesis, Eindhoven University of Technology, 2009