Abstract

Process mining techniques have usual notable attention within the literature for their ability to help within the redesign of complex processes by mechanically discovering models that specify the events registered in some log traces provided as input. Process mining refers to the extraction of process models from event logs. Now real-life processes tend to be less structured and a lot of flexible. Traditional process mining algorithms have issues dealing with such unstructured processes and generate "spaghetti-like" process models that are exhausting to understand. An approach to beat this is often to cluster process instances specified every of the ensuing clusters correspond to coherent sets of process instances which will every be adequately represented by a process model. To overcome these issues projected system aims to produce associate automatic means for code engineers to get mined models from systematic event logs specification embrace drawback finding, operating to learn others and technical challenge. This technique at first converts the Systematic Event Logs into some intermediate type like translated tokenized log file and keyword filtered log file. Then this filtered log file format is analyzed to extract the knowledge like Similarity matrix, Frequency count, Most read/write information, database queries and these event logs data measure accustomed build the clusters. Any system would generates the clusters using ActiTraC algorithm to produce refined description of generated models therefore incorrectness and additional overhead in
analysis part of model development is removed to extended extent. This is supported on an repetitious, graded, refinement of the process model, where, at every step, traces sharing similar behavior patterns are clustered along and equipped with a specialized schema. The formula guarantees that every refinement results in an progressively sound model, so attaining a monotonic search.

References

- Ferreira et al, Sequence Clustering: "Techniques for Process Mining Sequence clustering" (2007).
- Can Wang, Xiangjun Dong, Fei Zhou, Longbing Cao: "Coupled Attribute Similarity Learning on Categorical Data" (2014)
- Yongkweon Jeon and Sungroh Yoon, "Multi-Threaded Hierarchical Clustering by Parallel Nearest-Neighbor Chaining" (2013)
- Wil van der Aalst, Senior Member, "Service Mining: Using Process Mining to Discover, Check, and Improve Service Behavior" (2013)
- G. Greco, A. Guzzo, L. Pontieri, and D. Sacca: "Discovering Expressive

Index Terms

Computer Science
Information Sciences

Keywords
Process mining  event log  process discovery  trace clustering  process model data mining.