Interactive Data Mining and Analysis System

Tool enables parameter-centric visual rule association exploration in real-time

TITLE

PARAS: Parameter Space-based Association Rule Mining Methodology

FIRE: Interactive Visual Support for Parameter Space-Driven Rule Mining

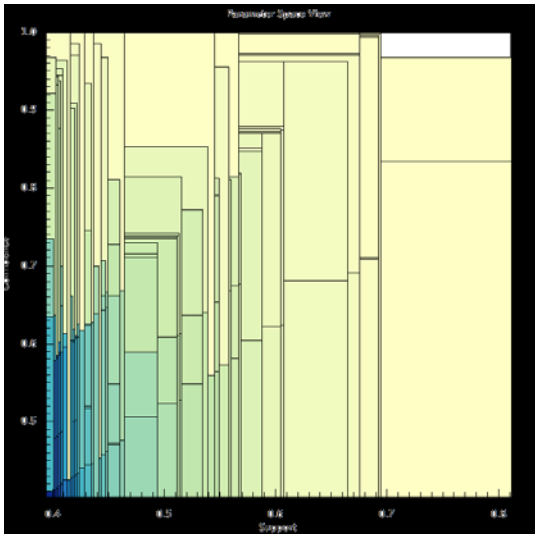
INVENTORS

Elke A. Rundensteiner
Matthew O. Ward
Abhishek Mukherji
Xika Lin
Jason Whitehouse
Christopher R. Botaish

COPYRIGHTED SOFTWARE

CONTACT

Todd Keiller, Director
Director, Intellectual Property & Innovation
Worcester Polytechnic Institute
100 Institute Rd
Worcester, MA 01609
tkeiller@wpi.edu
+1 508 831 4907



Learn more: http://paras.cs.wpi.edu

SUMMARY

- FIRE front-end visual interface is powered by PARAS back-end computing technology to create an effective and robust data analysis tool
- System ingests data sets and performs rule extraction and compression ahead of time, storing information in a compact and quickly accessible index
- Interactive parameter space display organizes rules by support and confidence parameters, enabling analysts to gain insights into rule distributions and relationships
- Separate glyph view allows analysts to visually drill-down and compare rules using shapes and colors, as well as to produce clusters of similar rules
- Supporting tool suite enables analysts to explicitly explore, filter, and compare and contrast rules and rule sets without trial-and-error parameter exploration

BACKGROUND

In today's highly networked and predominantly digital world, big data is big business. Yet, current data mining technologies offer limited performance due to high response times and the need for guesswork to determine optimal parameter settings for rule exploration. These impediments can cause delays that put analysts at a competitive disadvantage when decisions need to be derived quickly.

The PARAS and FIRE tools represent a leap forward that makes interactive rule exploration using successive parameter settings both feasible and practical. By pre-processing data and rulesets and presenting interactive displays containing recommendations for parameter tuning, the tools enable analysts to make informed decisions in real time, exploring large data sets at the speed of thought.

ADVANTAGES

- Offers significant performance improvements in CPU and memory consumption by several orders of magnitude over state-of-art approaches
- A comprehensive user study found that subjects performed tasks in significantly less time with increased accuracy using FIRE compared to state-of art rule-miner
- Interactive displays provide broad overviews that can be further drilled down, analyzed, and compared, offering effective sense making support
- Has the potential to reduce time and costs for data mining and analysis in any market where correlations in data relationships exist, including finance, business, science, engineering, and medicine
- PARAS and FIRE can be licensed and used separately for distinct needs