Anomaly Detection Service for Financial Data Streams

Paweł Gławiński, Marek Wojciechowski, Maciej Zakrzewicz
Softman S.A., Poznan University of Technology

PRESENTATION OUTLINE

- Introduction: business anomaly detection
- Background: SOA-based BPMS
- Contribution: business anomaly detection service
- Solution topology and architecture
  - Rule-based anomaly detection
  - Aggregation rules: aggregate materialization
  - Learning model rules: learning performance
- Summary

INTRODUCTION: BUSINESS ANOMALY DETECTION

- **Identification** of business items/events which **do not conform to a valid data pattern**
  - credit card frauds, purchase card frauds, telecommunication subscription fraud, phone call fraud, financial reporting fraud, insurance fraud, fraudulent claims for health care, credit applications fraud, credit transactional fraud, etc.
- To **timely detect** anomalous/fraudulent activities attempted by performers of business processes fed with streams of complex data
- Anomalies in **business process execution**
  - unusual process object state, unusual process execution path, unusual performer-to-activity assignment, etc.

BACKGROUND: SOA-BASED BPMS

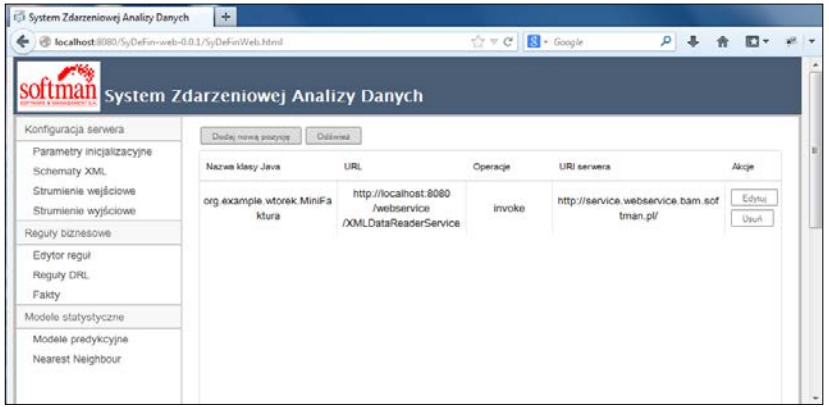
- **Business Process Management System** (BPMS):
  - services (typically Cloud Web Services) orchestrated by BPMN/BPEL into complex business processes performed by humans and applications
- Monitoring of the flow of business processes **still lacks** usability and flexibility in **current BPMSs**
  - eg. detecting anomalous business behavior
- Existing solutions for business process execution anomaly detection
  - **explicit calls** to rule evaluation systems (Complex Event Processing)
  - **external tools** for near-real time business reporting (Business Activity Monitoring)
  - anomaly detection **outside** functional requirements

Our Contribution

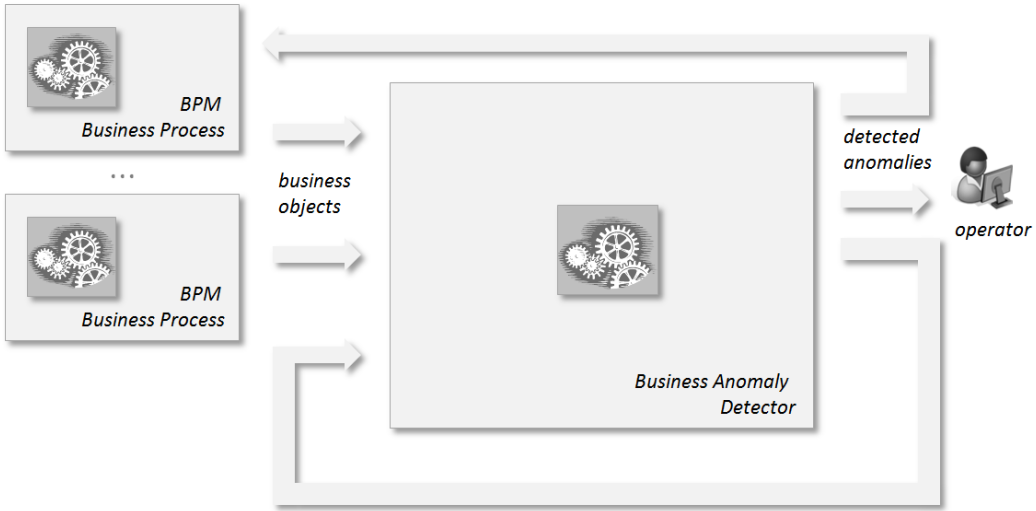
- Monitoring capabilities of BPMSs should be **expanded** with functions to **automatically monitor** every single business process instance
  - **detect** anomalous activities
  - **report** findings to other components/processes
- Business Anomaly Detection **Solution Pattern**
  - successful implementation: synchronous/asynchronous Java EE Web Service
  - easily injected into existing Business Process Management System (BPMS) environments
  - automated detection of anomalous behavior
  - **performance optimization**: aggregate materialization and offline learning

Rule-Based Anomaly Detection

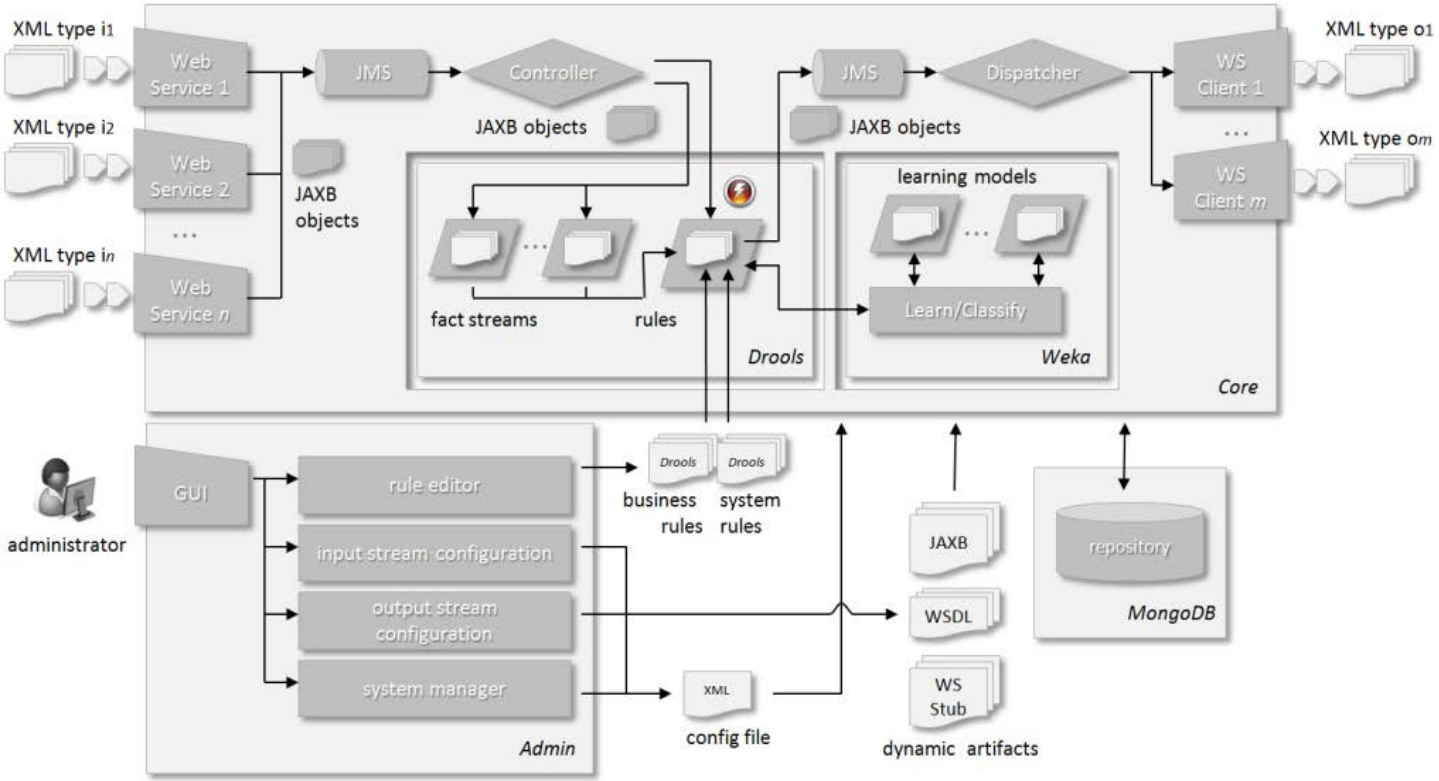
- **Simple rules** based on the current business object only
- **Aggregation rules** based on moving window aggregates calculated from collections of business objects received in the past
- **Calendar rules** based on schedules to validate business objects received recently
- **Learning model rules** based on patterns learnt from business objects received recently



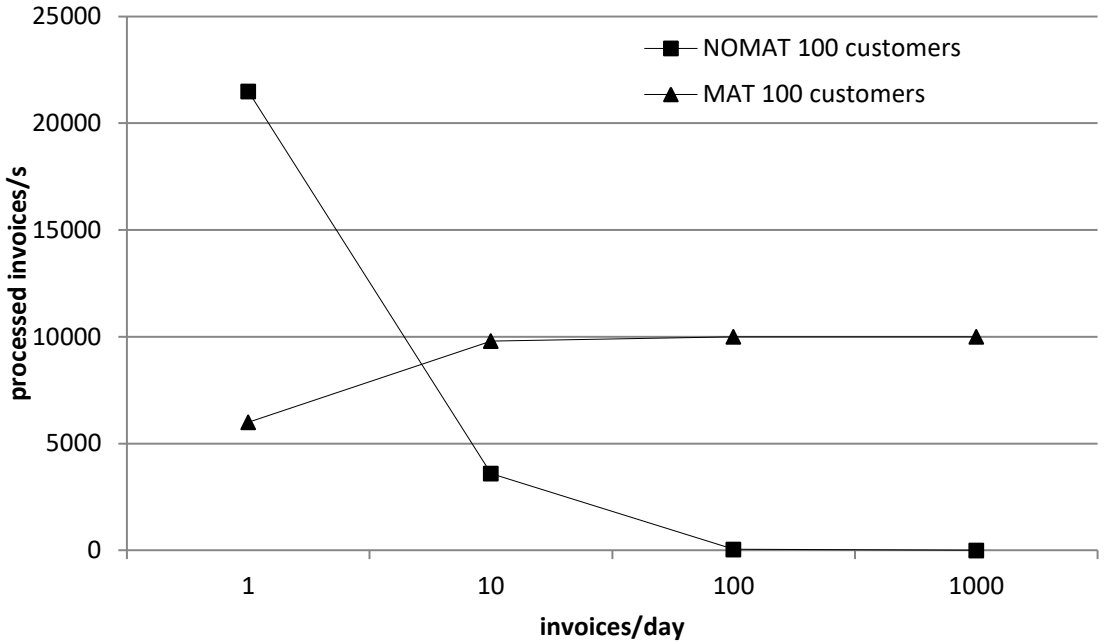
SOLUTION TOPOLOGY AND ARCHITECTURE



SOLUTION TOPOLOGY AND ARCHITECTURE

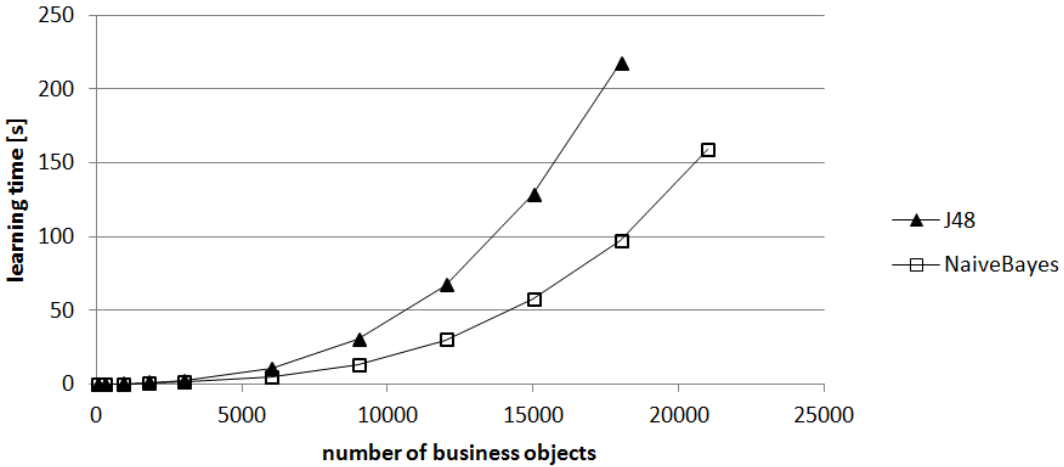


AGGREGATION RULES: AGGREGATE MATERIALIZATION



- Example: check if an invoice is 50% above 30-day moving average
- Drools aggregate evaluation inefficient for heavy data streams
- Our solution:
  - daily **aggregates materialized** in MongoDB store

LEARNING PERFORMANCE: OFFLINE LEARNING REQUIRED



- Example: check if an invoice looks *different* than other invoices seen before
- Anomaly detection by machine learning cannot be performed on-line for heavy data streams
- Our solution:
  - **shadow model learning offline**
  - primary model for classification

SUMMARY

- Business Anomaly Detector solution pattern
  - infrastructural service, part of BPMS
  - intercepting (explicitly or implicitly) business objects from business process flows in order to detect anomalous behavior
  - synchronous/asynchronous architecture
  - four types of anomaly detection rules
  - aggregate materialization
  - off-line learning of discovery-based business rules
- A prototype system implemented and validated in a real-life environment
  - 10K events/s processed using 4-core Intel CPU