VTechnik TVue + Mojaloop

VTechnik TVue

Fraud detection/prevention for Mojaloop

Alex Shmelev, CTO
VTechnik Corp.
alex@vtechnik.com

VTechnik TVue – What is it?

- Event-Sequence pattern detection
  - High throughput, low latency event pattern detection
  - “Time” and “Event Sequence” define the pattern “signal”
    - The sequence of events and the time between events is the highest-order indicator of “intent”
- In-memory Event database
  - Real-time pattern evaluation
  - Ultra-low latency (i.e. < 1 ms)
- Distributed, scalable architecture
  - Perfect for high-volume, real-time applications
  - Cloud, infrastructure or hybrid deployment

VTechnik TVue – What does it do?

- TVue consumes “Events”, stores them by “partition key” and timestamp
  - Events can model any user/system interaction
    - Account creation, credit/debit, adjustments, inquiries, overdraft, etc.
- TVue uses continuous queries to inspect Event Sequences for various patterns.
  - Pattern matches can trigger TVue “reactions”
    - TVue can send “Alert” messages, route messages to different queues, add/change/remove attributes
  - Pattern matches can trigger “chained” queries
    - TVue can automatically gather additional information about a pattern match

VTechnik TVue – What does it do?

- TVue supports ad-hoq queries
  - Real-time, analyst initiated queries on live data stream
  - Prefix patterns
    - “What are the most prevalent events after a pattern?”
  - Post-fix patterns
    - “What are the most prevalent actions before a pattern?”
- TVue supports pattern discovery
  - “Which usage patterns are the costliest to serve?”
  - “Which usage patters are the least-frequently observed?”
  - “Find patterns that do not match ‘normal’ sequences.”

VTechnik TVue – How does it work?

- TVue consumes Events
  - Events can be in various formats:
    - CSV, JSON, XML
  - Events can be ingested via various integrations:
    - AMQP, Kinesis, SQS, Kafka, REST, SOAP, CSV files
  - Event data is semi-structured
    - TVue requires: Event Type, Partition Key, Timestamp
    - TVue stores any other values as “event attributes”
    - Any attribute can be part of a pattern
    - Patterns can reference any attribute in any prior event

VTechnik TVue – How does it work?

- Events are stored in an in-Memory database
  - Events are grouped under the “Partition Key”
    - Partition Key can be:
      - Customer ID, Account ID, Agent ID, Device ID, etc
    - Data can be stored by more than one Partition Key
  - Events are in time-order (adjusted by timezone and precedence
  - Event data is distributed evenly across the TVue cluster
    - All queries are run “in parallel” across the cluster for maximum performance

VTechnik TVue – What is a “pattern”?

- Patterns = events + time
- Normal patterns:
  - Most (99.97%+) patterns are “normal usage”
    - For “fraud detection”, normal usage = “noise”
  - Deposit + 1 day + withdrawal + 1 day + withdrawal ... 12 days + deposit...
- Fraud patterns:
  - Previous analysis has identified ~200 fraud patterns
  - Create account + 1 day + deposit 1.00 + 1 day + deposit 1.00 + 1 day + deposit 100,000.00 + withdrawal 100,000.00…
  - Known fraud patters can be “pre-defined” in TVue from first day

VTechnik TVue – Event Sequence

- What does an Event Sequence look like?

Event [type=C, timestamp=1555270345948, pk=100000332, attributes=[100000332, 2019-04-01 12:32:25.948, AC, 33769e45-8007-4e86-b34a-1f702853bdb4]]
Event [type=T, timestamp=1555270383337, pk=100000332, attributes=[100000332, 2019-04-05 12:33:03.337, CR, d3940633-9384-4683-a240-43f362e014d2, USD, 1061, 100000680]]
Event [type=T, timestamp=1555270383362, pk=100000332, attributes=[100000332, 2019-04-07 12:33:03.362, CR, b6f194ac-e7ae-4a5c-8f6c-b7b55f2d27b0, USD, 1025, 100000011]]
Event [type=T, timestamp=1555270393616, pk=100000332, attributes=[100000332, 2019-04-08 12:33:13.616, DB, b4bc782a-3f81-4650-bb08-6f7dd3edd7a0, USD, 424, 100000042]]
Event [type=T, timestamp=1555270394282, pk=100000332, attributes=[100000332, 2019-04-09 12:33:14.282, DB, 2cd4cbf1-388f-4a2b-af18-d3572dafff99, USD, 282, 100000165]]
Event [type=T, timestamp=1555270402993, pk=100000332, attributes=[100000332, 2019-04-10 12:33:22.993, DB, 1dbef77b-2f09-4856-8bdd-c6b0b99b06bb, USD, 466, 100000460]]
Event [type=T, timestamp=1555270403942, pk=100000332, attributes=[100000332, 2019-04-11 12:33:23.942, DB, 2b08a6fa-2c84-4783-8f23-832793351a0e, USD, 316, 100000363]]
Event [type=T, timestamp=1555270412196, pk=100000332, attributes=[100000332, 2019-04-12 12:33:32.196, DB, 9746d443-d75c-4831-a69d-9dd7d06b7e30, USD, 105, 100000752]]
Event [type=T, timestamp=1555270536784, pk=100000332, attributes=[100000332, 2019-04-13 12:35:36.784, CR, 1905371f-d163-401b-8336-ac01cfda5998, USD, 1788, 100000344]]
Event [type=T, timestamp=1555270571301, pk=100000332, attributes=[100000332, 2019-04-14 12:36:11.301, DB, 2aee8373-57cb-4ebd-ad30-7b001cbf0a00, USD, 997, 100000220]]

Events require: Type, Timestamp, Partition Key (PK)

Many other attributes are allowed. Every attribute can be used to define a pattern

VTechnik TVue – Pattern

- A pattern is any sequence of events across time
  - Create account → deposit → withdraw → deposit → etc.
  - A normal pattern like this can be ignored...

VTechnik TVue – Fraud pattern

- Example: Recent account creation…
- Followed by:
  - A significant deposit, and...
  - A large number of payments in quick succession...

VTechnik TVue – Fraud pattern

Fraud patterns can be defined by an operator or discovered automatically. They are described in JSON:

```
{
  "pattern": "Fanout",
  "define": {
    "has": {
      "all": {
        "events": [
          {
            "eventType": "Acct",
            "minInstance": 1,
            "maxInstance": 1
          },
          {
            "eventType": "Tx",
            "eval": {
              "txType": {
                "op": "eq",
                "values": [
                  "CR"
                ]
              },
              "txValue": {
                "op": "gt",
                "values": [
                  "100000"
                ]
              }
            },
            "minInstance": 1,
            "maxInstance": 1
          },
```

… ->

```
{
            "eventType": "Tx",
            "eval": {
              "txType": {
                "op": "eq",
                "values": [
                  "DB"
                ]
              },
              "txValue": {
                "op": "gt",
                "values": [
                  "1000"
                ]
              }
            },
            "minInstance": 10,
            "maxInstance": null
          }
        ]
      }
    }
  },
  "result": {
    "message": {
      "topic": "fraud-alert",
      "qos": 100
    }
  }
}
```

VTechnik TVue – Automatic pattern discovery

- TVue also supports ad-hoc pattern queries and directed pattern discovery
- Ad-hoq queries:
  - Most common sequences leading up to a known pattern
  - Most common sequences that failed to complete a pattern
  - Population comparisons across patterns and time
- Automated pattern discovery:
  - TVue can discovery patterns based on any “fitness function”. The fitness function can measure any combination of event sequence or attribute values. This can be used to find previously unknown event sequences that maximize the “pattern fitness” - for example complex payment flows accumulating to or draining from an account.

VTechnik TVue – Pattern match response

- When TVue detects a pattern match it can...
- Send an alert message:
  - To a Kafka topic, or
  - An AQMP exchange
  - Execute a REST API call to a remote endpoint
- Re-direct the event traffic:
  - Rather than returning the message to Mojaloop to continue processing, TVue can re-direct the message to an Exceptions Queue

VTechnik TVue – Integration

- TVue integrates via one or more message streams
  - Kafka
  - Kinesis
  - AMQP
- Other options for acquiring event data are:
  - Most SQL databases
  - CSV files
  - S3 (data lake) or any Apache Drill source

VTechnik TVue – Scalability & Performance

- TVue scales with 99.97% linearity
  - Works with any size cluster
  - Many smaller instances will perform better than fewer large instances
  - Can be deployed in AWS, Azure, datacenter
- TVue can inspect approx. 100,000 tx/sec
  - 10 instances AWS (medium), 70MM events
  - 1.3% (fraud) match rate
  - Kafka integration