A Journey to comprehensible User Behavior Models

SSP 2020, Online

Reiner Jung & Lars Jürgensen

12th November 2020
iObserve - Overview

[Heinrich et al. 2015]
Operator in the Loop

Custom Workloads
user behavior
workload intensities

Operator-in-the-Loop Adaptation to involve operator adaptation decisions

Plan
Privacy-driven Planning
to establish and maintain privacy policy compliance
Design-space Exploration
for cost- and performance-constraint planning

Monitor
Descriptive architectural run-time models

Analyze

Execute
Prescriptive architectural run-time models

Reiner Jung
What we wanted to do?

[Vögele et al. 2018]
iObserve Behavior Analysis

aggregate entry events

compute behavior graph

collect all graphs in a time interval

convert graphs to matrix/markov chain and convert to vector

cluster vectors

Session Collector

Graph Creator

Graph Collector

Computing Behavior Vectors

Clustering

Entry Events

Sessions

Graphs

Collection of Graphs

Collection of Vectors

Centroids for Behaviors

[Jung et al. 2017]
Experiment Setup

<<service>> Workload Driver

<<service>> JPetStore

<<service>> Behavior Analysis

<<tool>> Ideal Behaviors

Randomized Behavior Model Workload

Fixed Behavior Model Workload

Ideal Behaviors

Behavior Model

Model Compare

Reiner Jung
Page and Transition Graph

Catalog.index

2

Catalog.viewCategory

CATS : 9

Catalog.viewProduct

9

FL-DSH-01 : 8

Account.login

Account.signonForm

EST-14 : 8
EST-15 : 1

Cart.addltemToCart

Order.newOrderForm

Order.index
Experiment Execution

- <<service>> Workload Driver
- JPetStore
- Behavior Analysis
- <service> Workload Driver
- Workload Driver
- Fixed Behavior Model Workload
- Ideal Behaviors
- <<tool>> Model Compare
- Behavior Model
- Randomized Behavior Model Workload
- Behavior Model
What we found out?

Clustering matrices did not work

- X-Means
- EM
- Hierarchical clustering
- Similarity of matrices somewhat better
Why does it not work?

- Too many dimensions: $dimensions = node^2$
- Other attributes cannot be included
- Also: Attributes at nodes insufficient
Now What?
New Approach

Entry Events → Session Collector → Graph Creator → Graph Collector
aggregate entry events
compute behavior graph
collect all graphs in a time interval

Collection of Graphs

Generate M-Tree → Optics Algorithm → Extract Clusters

M-Tree

Graph Edit Distance

[Jürgensen 2019]
New Graph Model

REPTILES : 1
BIRDS : 4

Catalog.index

Catalog.viewCategory

Catalog.search

Catalog.viewProduct

Catalog.viewItem

FI-FW-02 : 1
RP-SN-01 : 3
AV-SB-02 : 3
K9-BD-01 : 1
AV-CB-01 : 1

EST-21 : 1
EST-19 : 2
EST-7 : 1
EST-12 : 2
EST-18 : 1
Graph Edit Distance - Results

reachability-distance

MinPts=10

cluster-ordering of the objects

[Jürgensen 2019]
JIRA Data

Issues

1. Issue memory exhausted
2. Some race condition errors
3. Huge user behavior graphs

What went wrong?

- **Shop System (JPetStore)**
  - User pursue single tasks
  - Short sessions
- **Ticket System (JIRA)**
  - User perform multiple tasks
  - Long sessions
**JIRA Data**

**Issues**
1. Issue memory exhausted
2. Some race condition errors
3. Huge user behavior graphs

**What went wrong?**

![Diagram of data processing](image)

- **Session Collector**
  - Entry Events
- **Graph Creator**
  - Sessions
- **Graph Collector**
  - Graphs
- **Generate M-Tree**
  - Collection of Graphs
  - M-Tree
- **Optics Algorithm**
  - OPTiCS Plot
- **Extract Clusters**

**Graph Edit Distance**

C1

Session Collector
Graph Creator
Graph Collector
Entry Events
Sessions
Graphs

*Reiner Jung*
Conclusion

Summary

- **Graph Edit Distance** + **OPTICS** works
- **Attributes for edges** works
- **User sessions** are not a good idea

Potential Solutions

- Mark nodes as **end of an action**
- Split user sessions with graph pattern analysis

What will happen next

- Implement different pattern approaches
- Tune graph clustering
- Move more results to Kieker

Reiner Jung

