Model Problem (CrowdNav) and Framework (RTX) for Self-Adaptation Based on Big Data Analytics (Artifact)

Sanny Schmid[1], Ilias Gerostathopoulos[2], Christian Prehofer[3], and Tomas Bures[4]

1 **Fakultät für Informatik, Technische Universität München, Munich, Germany**
schmidsa@in.tum.de

2 **Fakultät für Informatik, Technische Universität München, Munich, Germany**
http://orcid.org/orcid.org/0000-0001-9333-7101
gerostat@in.tum.de

3 **Fakultät für Informatik, Technische Universität München, Munich, Germany and fortiss GmbF, Munich, Germany**
prehofer@in.tum.de

5 **Charles University in Prague, Faculty of Mathematics and Physics, Prague, Czech Republic**
bures@d3s.mff.cuni.cz

Abstract

This artifact supports our research in self-adaptation in large-scale software-intensive distributed systems. The main problem in making such systems self-adaptive is that their adaptation needs to consider the current situation in the whole system. However, developing a complete and accurate model of such systems at design time is very challenging. We are instead investigating a novel approach where the system model consists only of the essential input and output parameters and Big Data analytics is used to guide self-adaptation based on a continuous stream of operational data. In this artifact, we provide a concrete model problem that can be used as a case study for evaluating different self-adaptation techniques pertinent to complex large-scale distributed systems. We also provide an extensible tool-based framework for endorsing an arbitrary system with self-adaptation based on analysis of operational data coming from the system. The model problem (CrowdNav) and the framework (RTX) have been packaged together in this artifact, but can also work independently.

1998 ACM Subject Classification C.1.3 Other Architecture Styles–Adaptable Architectures

Keywords and phrases self-adaptation; Big Data analytics; model problem, tool, framework

Digital Object Identifier 10.4230/DARTS.3.1.5

Related Article Sanny Schmid, Ilias Gerostathopoulos, Christian Prehofer, and Tomas Bures, "Self-Adaptation Based on Big Data Analytics: A Model Problem and Tool", in Proceedings of the 12th International Symposium on Software Engineering for Adaptive and Self-Managing Systems (SEAMS 2017).
http://dx.doi.org/10.1109/SEAMS.2017.20

Related Conference 12th International Symposium on Software Engineering for Adaptive and Self-Managing Systems (SEAMS 2017), May 22–23, 2017, Buenos Aires, Argentina

1 Scope

This artifact aims to support research in self-adaptation of large-scale software-intensive systems. In particular, it provides a model problem and a tool-based framework for conducting experimental research in self-adaptation based on Big Data analytics, a recently proposed approach that suggests analyzing large streams of operational data from a system using Big Data techniques to plan and enact changes to the system. The artifact is intended to be used mainly by researchers in the self-adaptive systems (SEAMS) community.

2 Content

The artifact package includes:

- **CrowdNav**—a traffic simulation based on SUMO [1] and TraCI (a Python interface to SUMO) that implements a custom router that can be configured using Kafka [2] messages or a local JSON-based confiuration file. Configuration changes are applied on the fly (while the simulation is running). Also, runtime data is send to a Kafka queue to allow stream processing (e.g. using RTX); data is also logged locally to CSV file.
- **RTX**—a tool-based framework that allows for self-adaptation based on analysis of real time (streaming) data. RTX is particularly useful in analyzing operational data in a Big Data environment. RTX is written in Python. In its current version it internally uses Kafka and Spark Streaming [3]; however, it can be extended to employ other Big Data tools such as Flink.

3 Getting the artifact

The artifact endorsed by the Artifact Evaluation Committee is available free of charge on the Dagstuhl Research Online Publication Server (DROPS). In addition, the artifact is also available at: `https://github.com/Starofall/CrowdNav` and `https://github.com/Starofall/RTX`.

4 Tested platforms

- Both CrowdNav and RTX have been tested and work on Linux, MacOS & Windows.
- The main requirement for RTX is a platform that supports the Python libraries “numpy” and “scipy”; we provide a setup script to install them on Windows.
- To use Spark in RTX the system should have at least 4GB of RAM.
- Both CrowdNav and RTX can be run as Docker containers; Dockerfiles are provided.

5 License

Both CrowdNav and RTX are available under the MIT license.

6 MD5 sum of the artifact

9c66bfa099604fd15388c982ddf7d53b

7 Size of the artifact

3.46 GiB

References

1 Daniel Krajzewicz, Jakob Erdmann, Michael Behrisch, and Laura Bieker. Recent Development and Applications of SUMO - Simulation of Urban MObility. *International Journal On Advances in Systems and Measurements*, 5(3&4):128–138, December 2012. OCLC: 255207349.

2 Jay Kreps, Neha Narkhede, Jun Rao, and others. Kafka: A distributed messaging system for log processing. In *Proceedings of the 6th International Workshop on Networking Meets Databases (NetDB'11)*, pages 1–7, 2011. URL: http://people.csail.mit.edu/matei/courses/2015/6.S897/readings/kafka.pdf.

3 Matei Zaharia, Tathagata Das, Haoyuan Li, Timothy Hunter, Scott Shenker, and Ion Stoica. Discretized streams: fault-tolerant streaming computation at scale. pages 423–438. ACM Press, 2013. URL: http://dl.acm.org/citation.cfm?doid=2517349.2522737, doi:10.1145/2517349.2522737.