Anomaly detection on the CERN data centre monitoring data

Author: Ing. Antonín Dvořák
Supervisor: Mgr. Alexander Kovalenko, Ph.D.
Co-Supervisor: RNDr. Dagmar Adamová, CSc.
Co-Supervisor: Domenico Giordano, Ph.D.

FACULTY OF INFORMATION TECHNOLOGY CTU IN PRAGUE

Motivation

The CERN data centre is a crucial part of its scientific infrastructure. Over 90% of the computing resources at CERN are provided by OpenStack private cloud. This cloud is composed of 7895 physical servers, which contain 423 000 CPU cores and 1,73 petabytes of RAM.

The cloud infrastructure is continuously monitored by many sensors reporting every few seconds either the performance of the computer hardware or the status of the running applications.

The monitoring system is primarily used by the data centre's operators for:

- the real-time inspection of the data centre status,
- the threshold-based alarming on the infrastructure components,
- the post-mortem analysis of incidents.

In order to maximize the infrastructure availability, operators need to be provided with advanced analytic predictions, that effectively identify real operational issues, while still keeping low the false alarms rate that would otherwise increase the operators' load. In this context, the adoption of anomaly detection approaches leveraging machine learning techniques is considered crucial.

Solution

Data analytic infrastructure has been developed to process the cloud monitoring data. It exploits the ecosystem of big-data tools deployed at CERN (HDFS, Spark, Kubernetes, ElasticSearch, Grafana) to plug-in machine learning algorithms and streamline their predictions into the monitoring dashboards.

In this work, we present a novel anomaly detection solution applied to the performance metrics of a large-scale data centre.

The approach leverages the observation that the multitude of entities – computer servers and services – in a large-scale data centre are organized in aggregates, named clusters. During the daily operation, the monitored metrics of the entities in the same cluster show similar patterns, even if the behavior of each individual entity is largely unpredictable due to the lack of details about the demand of the running applications. In this scenario, the identification of anomalies by analyzing each individual time series is highly ineffective. A more effective approach leverages the deviations of the time series generated by anomalous servers from the values of the rest of the cluster.

We exploit this evidence to solve the anomaly detection problem in the multi-dimensional space of the monitoring time series of the whole cluster.

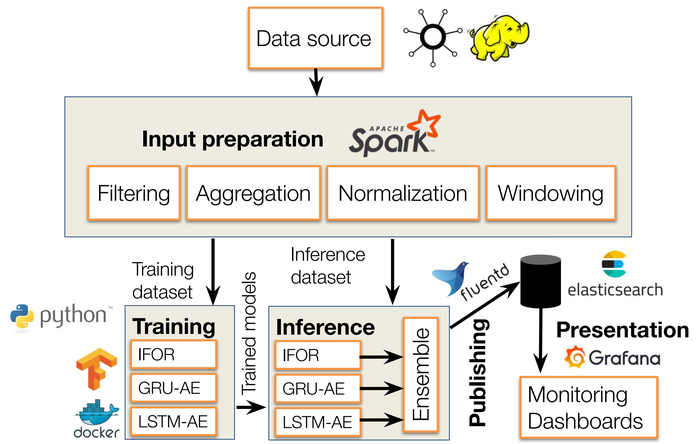
We adopt three unsupervised models, namely:

- Isolation Forest,
- LSTM-autoencoder,
- GRU-autoencoder,
- ensemble strategies.

The selected algorithms are suitable to model the collective behavior of the servers by jointly processing all their time series metrics, after we organize them as a multivariate data structure.

Pipeline

The multi-step pipeline had been developed and covers the whole chain of the Anomaly Detection procedure.



Dataset

The main reason for building the labeled dataset is to benchmark the Anomaly Detection methods under study for our use case. This benchmark gives a fast and objective way to evaluate the Anomaly Detection System.

The hand-annotated dataset is based on data of 40 servers over 2 months, resulting in 11712 samples, from which 228 are anomalous and 11484 are normal. The percentage of anomalies in the whole dataset is therefore equal to 2%.

Results

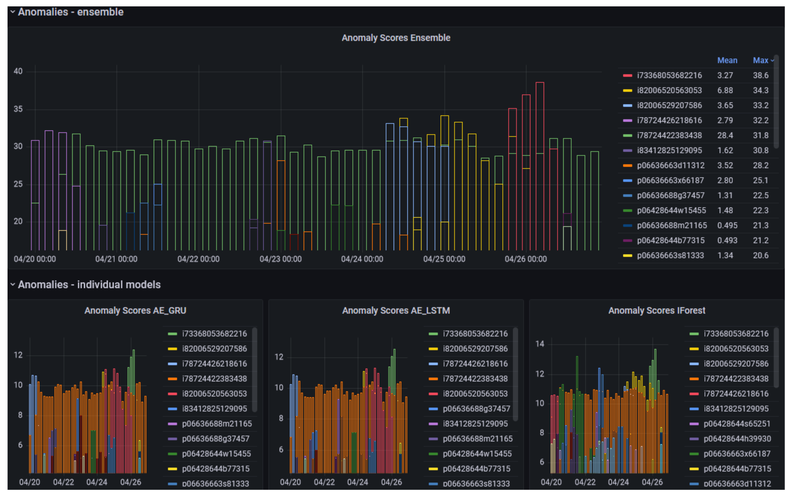
The effectiveness of our approach has been validated over a large amount of monitoring data produced by the cloud computing infrastructure.

A manually-labeled dataset representative of these data was used to benchmark the proposed models against the current CERN alerting system, which relies on threshold-based univariate time series analysis.

We show that both the individual models and the ensemble strategies significantly outperform it in terms of true positive rate, for the given false positive rate required by the data centre's operators.

FPR	True Positive Rates						
	Current System	Individual			Ensemble		
		IFOR	LSTM-AE	GRU-AE	ENS_1	ENS_2	ENS_3
0.001	0.08	0.09	0.45	0.43	0.47	0.45	0.21
0.01	0.14	0.29	0.56	0.58	0.53	0.58	0.59
0.02	0.19	0.57	0.66	0.79	0.61	0.69	0.71
0.04	0.26	0.81	0.81	0.93	0.92	0.89	0.92

Data centre's operators are provided a dashboard displaying the *degree of anomalousness* of servers by individual models and also by the best performing ensemble strategy.



We regard as a great achievement the satisfactory experience reported by the operators after months of extensive usage of our proposed Anomaly Detection System, which also qualitatively confirms the effectiveness of our solution.