Irregular Traffic Detection for Containerized Microservices RPC in the Real World

PI: Hao Chen (Univerisity of California, Davis)

Key Problem

Containerized microservices have been widely deployed. Irregular traffic frequently appears in RPCs among containers. Current RPC security relies on predefined rules and policies. However, it is challenging when it comes to thousands of microservices and massive real-time unstructured data. Due to the excellent performance of machine learning in extracting features from data, we seek to find a machine-learning framework for RPC traffic anomaly detection.

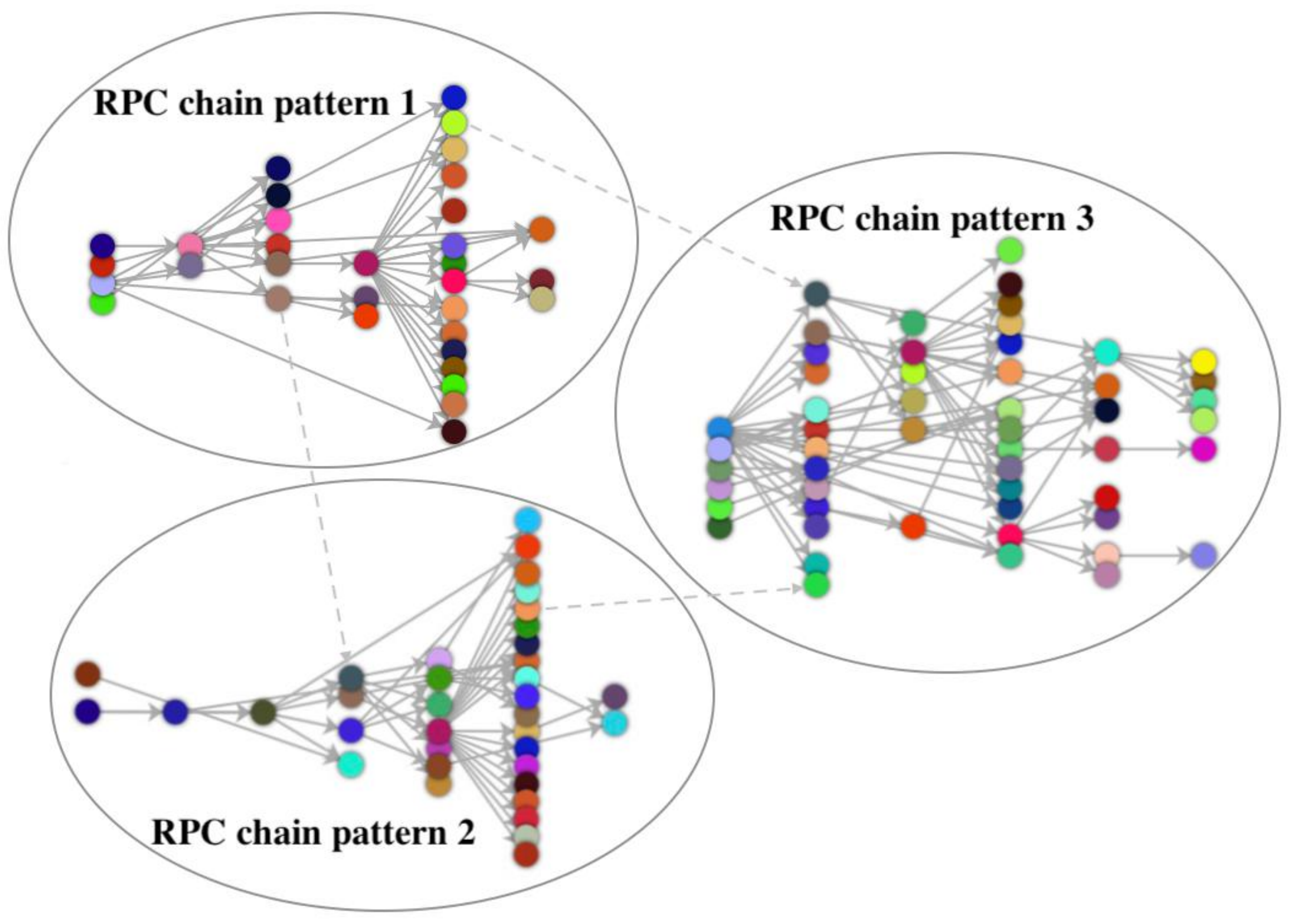

Fig.1 Schematic diagram of the RPC chain pattern clustering. Each circle is an RPC chain pattern. Intra-pattern dependencies are strong while inter-pattern dependencies are weak.

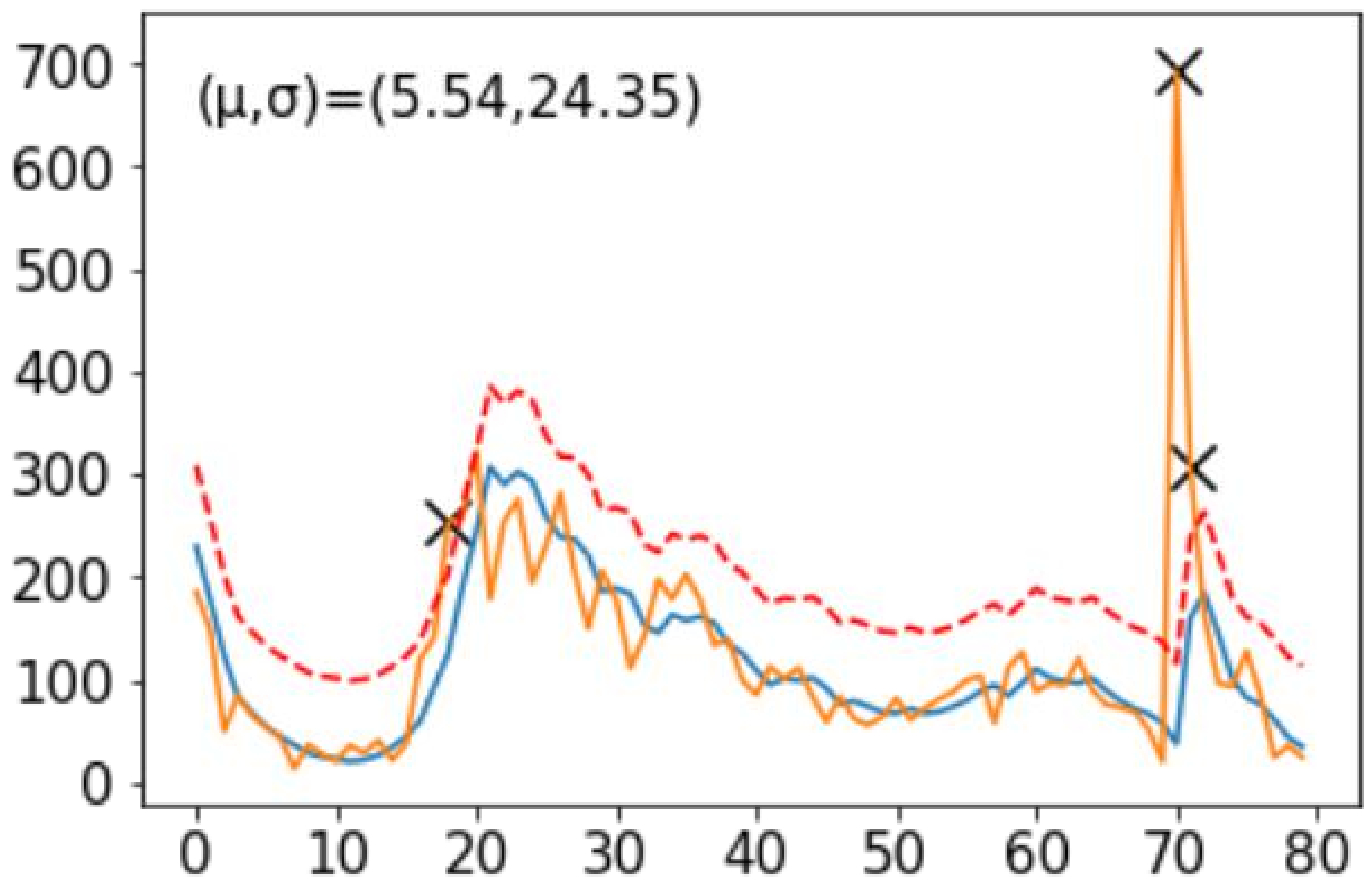

Fig.2 Traffic prediction and anomaly detection for one RPC. The X-axis is the timeline, and the Y-axis is the traffic. (Yellow:real, Blue:predicted, Red:threshold)

Solutions

- We propose a two-phase machine learning framework to track the traffic of each RPC and report irregular points automatically.
- First, we identify RPC chain patterns by density-based clustering and build a graph for each critical pattern.
- Next, we perform anomaly detection by a spatial-temporal graph convolution network of each graph. Each sub-model is lightweight and can be easily retrained.

Scientific Impact

Graph-structured data exist not only in RPCs inside containerized microservices systems but also in other networked systems. Our framework shows a way to leverage machine learning to provide automatic, lightweight, efficient security solutions given graph-structured data.

Preliminary Results

- Our framework can efficiently cluster all RPC chain patterns (Fig. 1).
- The spatial-temporal GCN model can accurately predict future RPC traffic (MAPE: 3.8%) and detect anomaly (Fig. 2).
- We perform two case studies to show the framework's capability of finding real-world security threats.

Broader Impact

Applying containerized microservice architecture is currently a common strategy in the industry for deploying large-scale applications and services. More than 1000 large companies are using Kubernetes. As shown by our real-world evaluations, our approach can be applied to large-scale real-world containerized microservices systems and can achieve promising performance in anomaly detection.