ANOVIZ: A Visual Inspection Tool of Anomalies in Multivariate Time Series

Patara Trirat*, Youngeun Nam*, Taeyoon Kim, Jae-Gil Lee†

School of Computing, KAIST
{patara.t, youngeun.nam, tykimseoul, jaegil}@kaist.ac.kr

Abstract

This paper presents **ANOVIZ**, a novel visualization tool of anomalies in multivariate time series, to support domain experts and data scientists in understanding anomalous instances in their systems. ANOVIZ provides an overall summary of time series as well as detailed visualizations of relevant detected anomalies in both query and stream modes, rendering near real-time visual analysis available. Here, we show that ANOVIZ streamlines the process of finding a potential cause of an anomaly with a deeper analysis of anomalous instances, giving explainability to any anomaly detector.

Introduction

Motivation Time-series anomaly detection (TSAD) identifies abnormal events of interest that need to be alerted for further investigation or anomalies that need to be removed to enhance downstream task performance. TSAD is a fundamental task in data mining with various applications, e.g., cloud service monitoring and fraud detection. Even though numerous methods and tools (Blázquez-García et al. 2021; Choi et al. 2021) have been proposed for TSAD using statistical or machine learning techniques, no attention has been given to a tool for visual inspection of predicted anomalies, which limits a more profound understanding of the detection results. Moreover, to properly deal with the anomalies in practice, domain experts (i.e., users) prefer to know which sensor/variable is the source of anomalies in the system rather than a simple quantitative performance of a detector. Therefore, we propose **ANOVIZ**—a novel visual inspection tool for anomalies in time series—to facilitate more in-depth analysis in finding potential causes and developing a deeper understanding of anomalous instances.

Approach We propose ANOVIZ as an example-based explanation (Liao, Gruen, and Miller 2020). Besides a system-wide summary of anomalies along with the variable-wise contribution scores, ANOVIZ visualizes relevant patterns of a specific variable and all variables enabling inspection of anomalous instances. It also provides a violin chart to locate the predicted anomalies in each variable's data distribution and visualizes adjacent anomaly scores to show the potential abnormality around the detected anomaly. As a result, users can effortlessly identify the most potential culprit(s) and further analyze the corresponding anomalous pattern.

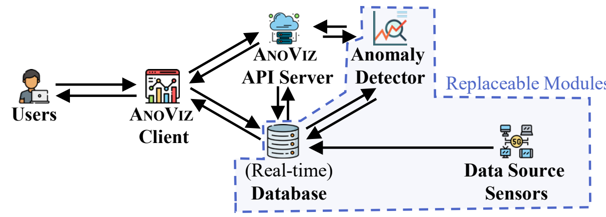

Figure 1: Architectural overview of ANOVIZ.

Significance The necessities of ANOVIZ are as follows. First, anomalies are mostly domain-specific, meaning that the constituents of an anomaly can differ considerably from one domain to another. ANOVIZ, as a domain-agnostic visualizer, helps users find the potential sources of anomalies in any domain given a database and detector. Second, there are point, collective, and contextual anomalies in time series, which means that it is not straightforward to interpret and reason about them. Besides, anomalies are likely to change over time due to the nature of the time series. Thus, tools that can constantly offer visual analysis of the detection results with insights about anomalous instances are imperative for users to understand each domain's anomalies and examine whether the current detector is of acceptable quality.

Novelty The key novelty of ANOVIZ resides in its unique focus on supporting users with visual inspection ability and a more informed explanation of anomalies in multivariate time series. Unlike previous work focusing on finding an optimal anomaly detector based on quantitative metrics, rarely with a simple visualization (Patel et al. 2022; Lai et al. 2021; Khelifati et al. 2021; Li et al. 2020; Eichmann et al. 2019) or locating potential root causes without visualization of relevant anomalies (Georgieva et al. 2022), ANOVIZ allows users to quickly specify the potential culprits and appropriately handle them based on the provided relevant insights. Furthermore, in addition to the query mode used by most existing work, our work also delivers a stream mode, enabling users to continuously analyze the results in near real-time.

*These authors contributed equally.
†Jae-Gil Lee is the corresponding author.

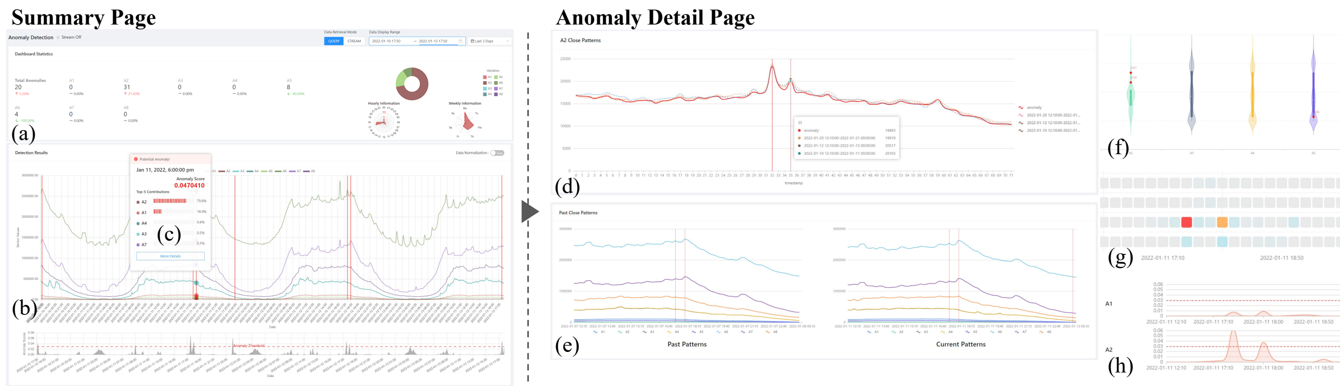

Figure 2: Screenshot of ANOVIZ's graphical user interfaces. Best viewed in electronic format.

System Overview of ANOVIZ

Figure 1 illustrates an overview of **ANOVIZ**[1] consisting of *three* core components: ANOVIZ Client, ANOVIZ API Server, and Replaceable Modules, as described below.

ANOVIZ Client

ANOVIZ Client handles user interfaces and interactions with the following core features. For demonstration, we build the client interfaces with Reactjs, Ant Design, and Chartjs.

Dashboard Statistics Once the detection results become available in the database, as shown in Figure 2(a), the ANOVIZ client will display the statistics of the anomalies detected so far with corresponding visualizations illustrating the distribution of these anomalies based on variables and time of occurrences. This summary helps users obtain an overview of anomalies in their systems.

Detection Results As shown in Figure 2(b), the client displays the raw multivariate time series and potential anomalies, along with anomaly scores and an optimal threshold selected during training. Here, in the *query* mode, users can specify any time range of interest for further analysis, while in the *stream* mode, users are allowed to select only recent l hours or days. These two modes facilitate the process of analysis in different scenarios. For example, users can select the query mode when they want to investigate past events and the stream mode when they want to catch or monitor a future anomaly in near real-time. Also, users can hover to see the values of variables. In the case of anomalous instances, as in Figure 2(c), users can hover to see the contribution score of each variable and click for more details.

Anomaly Details As in Figure 2, after clicking for more details, users can observe a variety of visualizations, including (d) an anomalous close pattern chart, (e) a past close pattern chart, (f) a violin chart, (g) a heatmap, and (h) raw anomaly scores. Here, the *anomalous close pattern chart* compares multiple time frames of a specific variable, and the *past close pattern chart* compares the most similar pattern against the current anomaly time frame, considering all variables; the former assists users in focusing on the anomalous behavior of a *specific* variable, while the latter in recognizing the anomalous trends of *all* variables. The *violin chart* shows the distributions of the variables' values and the position of the anomalous instance. Then, the *heatmap* and *raw anomaly scores* display the anomaly scores of surrounding timestamps based on the anomaly as the referenced point.

[1]Online demonstration is available at https://time-cad.web.app.

ANOVIZ API Server

ANOVIZ API Server, implemented by the Flask framework, controls the access to computations for visualizations to be displayed and application-specific replaceable modules.

Computation API This element computes the relevant values used for visualizations. It receives requests from the user interaction on the client side and connects to the database to get the related data for further computations.

Detector API This API is responsible for monitoring the incoming data stream. Once the initial data is enough for anomaly detection, it will trigger the anomaly detector to obtain results and then save them to the database. If ANOVIZ is in the stream mode, the results will be displayed instantly.

Replaceable Modules

Replaceable modules handle access to a database, a detection model, and data source sensors. As the domain-agnostic tool, ANOVIZ can connect with any (real-time) database and anomaly detector via database and detector adaptors. Here, we implement the database with Cloud Firestore to support real-time manipulations and employ our own TensorFlow-based anomaly detection model for demonstration purposes. Last, the data stream is produced through a simulation.

Conclusion and Future Work

ANOVIZ is a novel visual inspection tool of anomalies in multivariate time series that offers explainability to an anomaly detector through a rich set of user interfaces and visualizations while simultaneously allowing users to validate the detection quality. In future work, we expect to extend ANOVIZ with real-time troubleshooting systems and active learning techniques so that the users can solve the problems and improve the detection quality at the same time.

Acknowledgments

This work was supported by Mobile eXperience Business, Samsung Electronics Co., Ltd. (Real-time Service Incident Prediction Development).

References

Blázquez-García, A.; Conde, A.; Mori, U.; and Lozano, J. A. 2021. A Review on Outlier/Anomaly Detection in Time Series Data. *ACM Computing Surveys (CSUR)*, 54(3): 1–33.

Choi, K.; Yi, J.; Park, C.; and Yoon, S. 2021. Deep Learning for Anomaly Detection in Time-Series Data: Review, Analysis, and Guidelines. *IEEE Access*, 9: 120043–120065.

Eichmann, P.; Solleza, F.; Tatbul, N.; and Zdonik, S. 2019. Visual Exploration of Time Series Anomalies with MetroViz. In *Proceedings of the 2019 International Conference on Management of Data (SIGMOD)*, 1901–1904.

Georgieva, L.; Giurgiu, I.; Monney, S.; Pozidis, H.; Potocnik, V.; and Gusat, M. 2022. CCA: An ML Pipeline for Cloud Anomaly Troubleshooting. In *Proceedings of the 36th AAAI Conference on Artificial Intelligence (AAAI)*, 13167–13169.

Khelifati, A.; Khayati, M.; Cudré-Mauroux, P.; Hänni, A.; Liu, Q.; and Hauswirth, M. 2021. VADETIS: An Explainable Evaluator for Anomaly Detection Techniques. In *Proceedings of the 37th IEEE International Conference on Data Engineering (ICDE)*, 2661–2664.

Lai, K.; Zha, D.; Wang, G.; Xu, J.; Zhao, Y.; Kumar, D.; Chen, Y.; Zumkhawaka, P.; Wan, M.; Martinez, D.; and Hu, X. 2021. TODS: An Automated Time Series Outlier Detection System. In *Proceedings of the 35th AAAI Conference on Artificial Intelligence (AAAI)*, 16060–16062.

Li, Y.; Zha, D.; Venugopal, P. K.; Zou, N.; and Hu, X. 2020. PyODDS: An End-to-end Outlier Detection System with Automated Machine Learning. In *Companion Proceedings of The Web Conference 2020*, 153–157.

Liao, Q. V.; Gruen, D.; and Miller, S. 2020. Questioning the AI: Informing Design Practices for Explainable AI User Experiences. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–15.

Patel, D.; Ganapavarapu, G.; Jayaraman, S.; Lin, S.; Bhamidipaty, A.; and Kalagnanam, J. 2022. AnomalyKiTS: Anomaly Detection Toolkit for Time Series. In *Proceedings of the 36th AAAI Conference on Artificial Intelligence (AAAI)*, 13209–13211.