# Time-Series Anomaly Detection: Overview and New Trends

Qinghua Liu*
The Ohio State University
liu.11085@osu.edu

Paul Boniol*
Inria, DI ENS, PSL, CNRS
paul.boniol@inria.fr

Themis Palpanas
Université Paris Cité; IUF
themis@mi.parisdescartes.fr

John Paparrizos
The Ohio State University
paparrizos.1@osu.edu

## ABSTRACT

Anomaly detection is a fundamental data analytics task across scientific fields and industries. In recent years, an increasing interest has been shown in the application of anomaly detection techniques to time series. In this tutorial, we take a holistic view of anomaly detection in time series and comprehensively cover detection algorithms ranging from the 1980s to the most current state-of-the-art techniques. Importantly, the scope of this tutorial extends beyond algorithmic discussion, delving into the latest advancements in benchmarking and evaluation measures for this area. In particular, our interactive systems enable the exploration of detection algorithms and benchmarking results, thereby promoting user comprehension. Driven by the absence of a one-size-fits-all anomaly detector for various time series domains and applications, we review recent advancements in automated solutions and propose a new taxonomy to motivate further research.

**PVLDB Reference Format:**
Qinghua Liu, Paul Boniol, Themis Palpanas, and John Paparrizos. Time-Series Anomaly Detection: Overview and New Trends. PVLDB, 17(12): 4229 - 4232, 2024.
doi:10.14778/3685800.3685842

## 1 INTRODUCTION

A wide range of sensing, networking, and processing solutions enable the collection of large amounts of data over time [21, 22, 25–28, 34, 35, 38, 43]. Analytical tasks over such ordered sequences of real-valued data, known as *time series*, are becoming increasingly important in virtually every domain [3, 14–16, 31, 39–41, 44–49].

Anomaly detection has received ample academic and industrial attention over the past decades. *Anomalies* refer to data points that do not conform to some notion of normality or an expected behavior based on previously observed data. In practice, anomalies can correspond to erroneous data (e.g., broken sensors), or data of interest (e.g., anomalous behavior of the measured system) [1]. Detecting such cases is crucial for many applications [19]. Moreover, as illustrated in Figure 1, anomaly detection applied to time series (compared to other data types) is attracting more interest lately. Multiple surveys, benchmarks, and experimental studies summarize and analyze the state-of-the-art methods [4, 9, 10, 42, 52], exploring different aspects of the problem.

**Tutorial Overview.** As the interest in Time-Series Anomaly Detection (TSAD) expands across scientific fields and industries, this tutorial provides a comprehensive view of this task. We start with definitions for time series and anomalies. Our goal is threefold: (i) introduce the motivation related to the anomaly detection task in time series by describing different types of time-series anomaly and the taxonomy of anomaly detection algorithms proposed by different communities; (ii) describe recently proposed benchmarks and experimental evaluations along with the discussion on the challenges and problems inherent in the evaluation process; (iii) explore the latest trends in automated anomaly detection solutions for time series, providing a taxonomy and insights for each methodology.

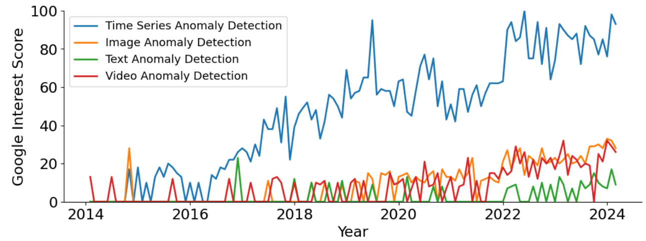



**Figure 1: Evolution of the Google interest score for anomaly detection over time series, images, text, and video.**

This tutorial features (i) an extensive review of literature spanning several decades, from early methods of the 1980s to the latest state-of-the-art approaches, and developed in different communities, from data management to machine learning; (ii) insights into the recent advancements in automated solutions, enriching the research landscape with novel perspectives; (iii) interactive systems for exploring a method's computation steps and experimental evaluation, making complex concepts engaging. We hope this tutorial will equip the audience with a fundamental comprehension of TSAD and inspire them to work in this dynamic and evolving area.

**Relation to Previous Tutorials.** Recent tutorials related to anomaly detection have focused (i) on specific types of methods (such as deep-learning methods [13]) ignoring several of the state-of-the-art methods, (ii) on specific types of temporal data (such as spatiotemporal data [56]), or (iii) on a much more general topic (only briefly discussing anomaly detection as a subpart of time series analysis [54]). Earlier versions of this tutorial [7, 8] focused on describing the state-of-the-art time-series anomaly detection methods and their operation.

In contrast, this tutorial undergoes significant expansion, offering a detailed explanation of recently proposed methodologies with interactive exploration for the audience. Beyond the introduction of anomaly detection techniques, this tutorial adopts a comprehensive perspective on the field, with a focus on evaluation measures and benchmark analysis. A brand new section has been introduced, outlining automated solutions with a new taxonomy and opening up new research opportunities. To our knowledge, this is the first tutorial that (i) extensively covers the time-series anomaly detection landscape; (ii) incorporates the latest developments in automated solutions for anomaly detection in time series; and (iii) elaborates

---

*These authors contributed equally to this work.



Proceedings of the VLDB Endowment, Vol. 17, No. 12 ISSN 2150-8097.
doi:10.14778/3685800.3685842

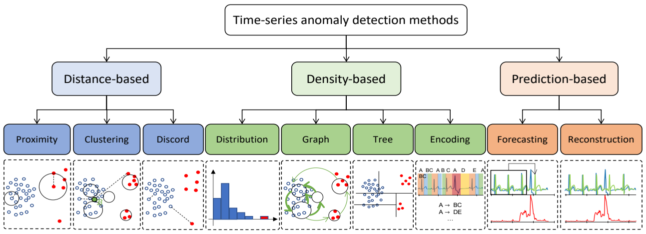



**Figure 2: Process-centric anomaly detection taxonomy.**

on concepts through the support of experimental studies and use of interactive demonstration systems.

**Audience and Expected Background.** This tutorial is for researchers and data analysts and will focus on recent advances in TSAD. The tutorial aims to initiate new collaborations between members of the data management community and data science practitioners in various domains and increase the interest in anomaly detection for time series. We will present material that includes the necessary background to follow the entire presentation and technical details of the current state-of-the-art solutions for anomaly detection. We will also discuss the drawbacks of existing solutions and the open problems. Both experts and newcomers in the area will be able to follow the material and benefit from the tutorial.

## 2 TUTORIAL SCOPE

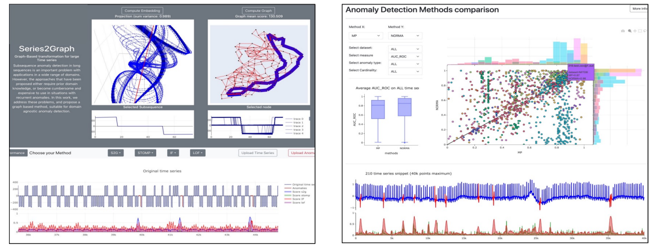
In this **1.5 hour lecture-style tutorial**, we will go through the problem of anomaly detection in time series, starting from fundamental definitions of time series and anomaly to open problems and opportunities that arise with the advent of automated solutions.

### 2.1 Introduction, Motivation and Foundations

We will start by discussing examples of scientific and industrial applications that rely on the effectiveness of time-series anomaly detectors. We will illustrate this variety of domains by showing concrete time series with real anomalies from multiple domains.

**Type of Time Series.** We then introduce the different types of time series. Specifically, we define the time series as an ordered sequence of real values on one (for *univariate* time series) or multiple dimensions (for *multivariate* time series). Moreover, a core characteristic of time series is their evolution with time. Therefore, we define *static* and *streaming* time series as sequences with a fixed length or continuously arriving subsequences. Finally, the normal behavior (i.e., subsequences representing the normal and recurrent behavior) might change over time. In this case, we differentiate *single* and *multiple normalities* time series.

**Type of Anomalies.** We then introduce the different types of anomalies. The first two categories, *point* and *contextual* anomalies, refer to data points deviating remarkably from the rest of the data globally or given a specific context, respectively. The third category, *Collective* anomalies, corresponds to sequences of points that do not repeat a typical (previously observed) pattern. Then, on top of these categories, their combination also matters. For instance, we need to differentiate time series containing *single* anomalies from time series containing *multiple* anomalies (either *similar* or *different*). As will described later, some methods that are based on nearest-neighbor distance might be affected by this distinction. For all these definitions and classes, we will provide explicit examples.




(a) GraphAn system (b) Theseus system

**Figure 3: Example of interactive systems that (a) allow the user to dive into computational steps [5], and (b) depicts experimental results [6].**

### 2.2 Taxonomy of Anomaly Detection Methods

We will dive into the different anomaly detection methods proposed in the literature. As many papers appear every year proposing new methods for anomaly detection in time series based on different applications [4, 18], it is beyond our scope to cover all proposed methods extensively here. In this tutorial, we will summarize the popular categories of methods.

**Classification by Inputs.** We will first mention the three categories of methods based on the external knowledge provided to them. First, *unsupervised* methods take time series as input and are not provided by any other information. Then, *semi-supervised* methods take as input time series without any anomalies and are trained on normal data only. Finally, *supervised* methods take as input separately both normal and abnormal data. Thus, the model is trained to discriminate the anomalies from the normality.

**Classification by Methodologies.** Then, we will describe the following categories of methods as depicted in Figure 2. First, *distance-based* approaches analyze subsequences by utilizing distances to a given model to detect anomalies. Second, *density-based* methods focus on detecting recurring or isolated behaviors by evaluating the density of the points or subsequences space. Third, *prediction-based* methods consist of two main categories: forecasting-based and reconstruction-based methods. The former, such as recurrent or convolutional neural network-based [30, 32], use the past values as input, predict the following, and use the forecasting error as an anomaly score. The latter, such as the autoencoder-based approach [50], are trained to reconstruct the time series and use the reconstruction error as an anomaly score.

### 2.3 Evaluating Anomaly Detection Methods

After describing various anomaly detectors, we will focus on how to evaluate them. The choice of benchmarks and accuracy measures may significantly bias the evaluation.

**Evaluation Measures.** We will start by describing the evaluation measures. Briefly, we will first discuss traditional measures, such as Precision, Recall, and F-score, that assess the methods by assuming each time-series point can be marked as an anomaly or not (e.g., by a threshold on an anomaly score). We will then discuss range-based variants [53] that aim to overcome shortcomings of traditional measures when evaluating time series containing subsequence anomalies. We will discuss Area Under the Curve (AUC) measures that, contrary to the previous measures, eliminate the need to define a threshold. We will finally discuss Volume Under the Surface (VUS) [36] measures, which provide more robustness.

**Benchmark Study.** Then, we will discuss recent benchmarks proposed for anomaly detection in time series task [20, 23, 42, 51]. Such benchmarks provide an extensive collection of time series from various domains and evaluate multiple methods belonging to the aforementioned categories. In this tutorial, we will discuss the results and conclusions of these recent benchmarks, as well as the criticisms that have been expressed about the characteristics and suitability of some datasets for this task [55].

**Interactive Exploration and Interpret Ability.** In contrast to previous tutorials, this tutorial employs interactive exploration to assist in understanding the concepts. It is important for users to overcome the lack of interpretability that anomaly detection methods can have [30, 50]. This is becoming possible by recently proposed prototype systems [5, 6, 11], which enable users to interactively explore methods and their inner workings and thus, better understand the different computation steps. In this tutorial, we will discuss and demonstrate recently proposed systems for anomaly detection, along with systems that allow the user to explore large-scale experimental evaluation studies. As depicted in Figure 3, we will use such systems during the tutorial to explain interactively the anomaly detection methods and the experimental conclusions.

## 2.4 New Trends and Opportunities

Recent benchmarks and evaluation studies (described in the previous sections) have demonstrated that no overall best anomaly detection methods exist when applied on very heterogeneous time series (i.e., coming from very different domains). The primary question that arises is: How can we *automatically* identify the most accurate anomaly detector in the time series dataset? In this tutorial, we will outline a taxonomy of the methodologies proposed in this field, providing detailed descriptions for each group. We will also highlight several promising directions for future research.

**Automated Solution Taxonomy.** We will begin by outlining the pipeline of automated solutions and then delve into different categories of methods. The works in this field can be classified into two main categories as depicted in Figure 4. *Model Selection* refers to identifying the best model and its corresponding hyperparameters from a predefined candidate model set. Subsequently, the selected model is utilized for anomaly detection. In this category, internal evaluation methods evaluate the effectiveness of a model by using surrogate metrics for anomaly detection, independent of external data such as ground truth labels for anomalies [17, 29]. In addition, meta-learning-based methods leverage the knowledge of the performance of various anomaly detectors on historical labeled datasets to enable the automatic model selection for new datasets [33, 52]. On the other hand, *Model Generation* entails the construction of a completely new model based on the candidate model set. This newly generated model then can operate independently as an anomaly detector. In this category, ensembling-based methods involve constructing models that aggregate predictions from the model set [2, 58]. Moreover, pseudo-label-based methods generate pseudo-labels to transform the unsupervised anomaly detection problem into a supervised framework [12, 57].

**Towards Reliable Benchmark Practice.** In light of recent criticisms regarding dataset flaws, such as mislabeling and bias [55], alongside flaws in some evaluation measures where random anomaly scores sometimes surpass SOTA methods [24], the need for a

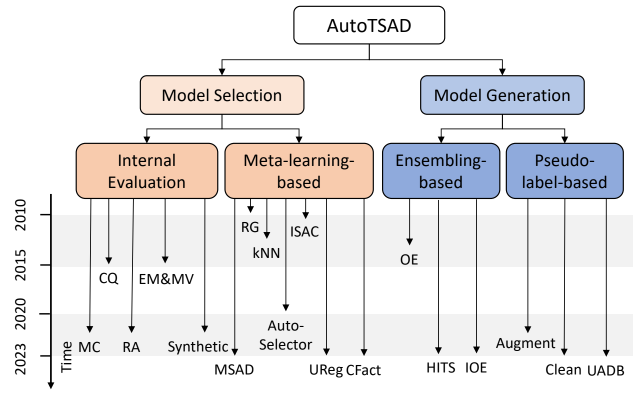



**Figure 4: A taxonomy and chronicle of automated solutions for anomaly detection in time series.**

reliable benchmark is paramount. A benchmark established on an unreliable testbed creates an illusion of progress. However, designing and agreeing on a comprehensive benchmark is still a question that the community needs to address. The burgeoning volume of data from the Internet of Things (IoT), which generates substantial time series and anomalous instances, underscores the importance of this task and presents opportunities for leveraging this wealth of data. We will outline recent efforts and discussion [37, 42] towards advancing unbiased and rigorous benchmarking practices.

**Towards Automated Anomaly Detection.** Given the absence of labeled data and one-fits-all anomaly detectors in anomaly detection, achieving optimal performance requires in-depth knowledge of the myriad of methods. This necessity drives data analysts into an exhaustive and time-consuming trial-and-error procedure process, making selecting suitable anomaly detectors cumbersome. Following the discussion of automated solution taxonomy, we will discuss prospective research avenues [11, 52], including addressing domain generalization problems, tackling challenges inherent in online settings and resource constraints, and advancing the development of incremental automated solutions. This field represents a promising area of research, and we hope our tutorial can serve as a catalyst to steer research efforts towards this direction.

## 2.5 Conclusions

We will conclude this tutorial by summarizing the main insights obtained from recent benchmarks of various anomaly detectors and automated solutions. Finally, we will summarize the new trends and challenges with the advent of automated solutions.

# 3 PRESENTERS

**Qinghua Liu** is a Ph.D. student at The Ohio State University and a member of The DATUM Lab. His research interests include time-series analysis and anomaly detection. He received the B.Eng. degree in Electronic Engineering from Tianjin University in 2022. Previously, he worked at SRIBD, CASIA and JD.com Inc.

**Paul Boniol** is a researcher at Inria, member of the VALDA project-team. Previously, he worked at ENS Paris-Saclay (Centre Borelli), Université Paris Cité, EDF Research lab, and Ecole Polytechnique. His research interests lie between data analytics, machine learning, and time-series analysis. His Ph.D. thesis focused on subsequence anomaly detection and time-series classification. His work has been published in the top data management and analytics venues.

**Themis Palpanas** is an elected Senior Member of the French University Insitute (IUF), and Distinguished Professor of computer science at Universite Paris Cite (France). He has authored 14 patents, received 3 best paper awards and the IBM SUR award, has been Program Chair for VLDB 2025 and IEEE BigData 2023, General Chair for VLDB 2013, and has served Editor in Chief for BDR. He has been working in the fields of Data Series Management and Analytics for more than 15 years, and has developed several of the state of the art techniques. He has delivered 19 tutorials in top conferences.

**John Paparrizos** is an assistant professor at The Ohio State University, leading The DATUM Lab. His research focuses on adaptive solutions for data-intensive and machine-learning applications. His doctoral work was recognized at the 2019 ACM SIGKDD Doctoral Dissertation Award competition. He has also received the inaugural ACM SIGMOD Research Highlight Award, a NetApp Faculty Award, and the 2023 IEEE TCDE Rising Star Award. His ideas have been widely adopted across scientific areas, Fortune 100-500 companies (e.g., Exelon and Nokia), and organizations such as ESA.

# REFERENCES


[1] Charu C Aggarwal. 2017. An introduction to outlier analysis. In *Outlier analysis*. 1–34.

[2] Charu C Aggarwal and Saket Sathe. 2015. Theoretical foundations and algorithms for outlier ensembles. *Acm sigkdd explorations newsletter* 17, 1 (2015), 24–47.

[3] Mohini Bariya, Alexandra von Meier, John Paparrizos, and Michael J Franklin. 2021. k-ShapeStream: Probabilistic Streaming Clustering for Electric Grid Events. In *2021 IEEE Madrid PowerTech*. IEEE, 1–6.

[4] Ane Blázquez-García, Angel Conde, Usue Mori, and Jose A Lozano. 2021. A Review on outlier/Anomaly Detection in Time Series Data. *ACM CSUR* 54, 3 (2021).

[5] Paul Boniol, Themis Palpanas, Mohammed Meftah, and Emmanuel Remy. 2020. GraphAn: Graph-based Subsequence Anomaly Detection. *Proc. VLDB Endow.* 13, 12 (2020), 2941–2944.

[6] Paul Boniol, John Paparrizos, Yuhao Kang, Themis Palpanas, Ruey S Tsay, Aaron J Elmore, and Michael J Franklin. 2022. Theseus: navigating the labyrinth of time-series anomaly detection. *Proc. VLDB Endow* 15, 12 (2022), 3702–3705.

[7] Paul Boniol, John Paparrizos, and Themis Palpanas. 2023. New Trends in Time-Series Anomaly Detection. In *EDBT*.

[8] Paul Boniol, John Paparrizos, and Themis Palpanas. 2024. An Interactive Dive into Time-Series Anomaly Detection. In *2024 IEEE 40th International Conference on Data Engineering (ICDE)*.

[9] Paul Boniol, John Paparrizos, Themis Palpanas, and Michael J Franklin. 2021. SAND in action: subsequence anomaly detection for streams. *Proc. VLDB Endow* 14, 12 (2021).

[10] Paul Boniol, John Paparrizos, Themis Palpanas, and Michael J Franklin. 2021. SAND: streaming subsequence anomaly detection. *Proc. VLDB Endow* 14, 10 (2021), 1717–1729.

[11] Paul Boniol, Emmanouil Sylligardos, John Paparrizos, Panos Trahanias, and Themis Palpanas. 2024. ADecimo: Model Selection for Time Series Anomaly Detection. In *2024 IEEE 40th International Conference on Data Engineering (ICDE)*.

[12] Lei Cao, Yizhou Yan, Yu Wang, Samuel Madden, and Elke A Rundensteiner. 2023. AutoOD: Automatic Outlier Detection. *Proceedings of the ACM on Management of Data* 1, 1 (2023), 1–27.

[13] Raghavendra Chalapathy, Nguyen Lu Dang Khoa, and Sanjay Chawla. 2020. Robust Deep Learning Methods for Anomaly Detection. In *SIGKDD (KDD '20)*.

[14] Adam Dziedzic, John Paparrizos, Sanjay Krishnan, Aaron Elmore, and Michael Franklin. 2019. Band-limited training and inference for convolutional neural networks. In *International Conference on Machine Learning*. PMLR, 1745–1754.

[15] Jens E d'Hondt, Odysseas Papapetrou, and John Paparrizos. 2024. Beyond the Dimensions: A Structured Evaluation of Multivariate Time Series Distance Measures. In *2024 IEEE 40th International Conference on Data Engineering Workshops (ICDEW)*. IEEE, 107–112.

[16] Rahul Goel, Sandeep Soni, Naman Goyal, John Paparrizos, Hanna Wallach, Fernando Diaz, and Jacob Eisenstein. 2016. The social dynamics of language change in online networks. In *Social Informatics: 8th International Conference, SocInfo 2016, Bellevue, WA, USA, November 11-14, 2016, Proceedings, Part I 8*. Springer, 41–57.

[17] Mononito Goswami, Cristian Ignacio Challu, Laurent Callot, Lenon Minorics, and Andrey Kan. 2022. Unsupervised Model Selection for Time Series Anomaly Detection. In *The Eleventh International Conference on Learning Representations*.

[18] Manish Gupta, Jing Gao, Charu C Aggarwal, and Jiawei Han. 2013. Outlier detection for temporal data: A survey. *TKDE* 26, 9 (2013), 2250–2267.

[19] Medina Hadjem, Farid Naït-Abdesselam, and Ashfaq A. Khokhar. 2016. ST-segment and T-wave anomalies prediction in an ECG data using RUSBoost. In *Healthcom*.

[20] Vincent Jacob, Fei Song, Arnaud Stiegler, Bijan Rad, Yanlei Diao, and Nesime Tatbul. 2021. Exathlon: A Benchmark for Explainable Anomaly Detection over Time Series. *Proc. VLDB Endow.* 14, 11 (oct 2021), 2613–2626.

[21] Hao Jiang, Chunwei Liu, Qi Jin, John Paparrizos, and Aaron J Elmore. 2020. Pids: attribute decomposition for improved compression and query performance in columnar storage. *Proc. VLDB Endow* 13, 6 (2020), 925–938.

[22] Hao Jiang, Chunwei Liu, John Paparrizos, Andrew A Chien, Jihong Ma, and Aaron J Elmore. 2021. Good to the Last Bit: Data-Driven Encoding with CodecDB. In *SIGMOD*. 843–856.

[23] E. Keogh, T. Dutta Roy, U. Naik, and A Agrawal. 2021. Multi-dataset Time-Series Anomaly Detection Competition 2021. https://compete.hexagon-ml.com/practice/competition/39/.

[24] Siwon Kim, Kukjin Choi, Hyun-Soo Choi, Byunghan Lee, and Sungroh Yoon. 2022. Towards a rigorous evaluation of time-series anomaly detection. In *AAAI*, Vol. 36. 7194–7201.

[25] Sanjay Krishnan, Aaron J Elmore, Michael Franklin, John Paparrizos, Zechao Shang, Adam Dziedzic, and Rui Liu. 2019. Artificial intelligence in resource-constrained and shared environments. *ACM SIGOPS Operating Systems Review* 53, 1 (2019), 1–6.

[26] Chunwei Liu, Hao Jiang, John Paparrizos, and Aaron J Elmore. 2021. Decomposed bounded floats for fast compression and queries. *Proc. VLDB Endow* 14, 11 (2021), 2586–2598.

[27] Chunwei Liu, John Paparrizos, and Aaron J Elmore. 2024. AdaEdge: A Dynamic Compression Selection Framework for Resource Constrained Devices. In *2024 IEEE 40th International Conference on Data Engineering (ICDE)*. IEEE, 1506–1519.

[28] Shinan Liu, Tarun Mangla, Ted Shaowang, Jinjin Zhao, John Paparrizos, Sanjay Krishnan, and Nick Feamster. 2023. Amir: Active multimodal interaction recognition from video and network traffic in connected environments. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 7, 1 (2023), 1–26.

[29] Martin Q Ma, Yue Zhao, Xiaorong Zhang, and Leman Akoglu. 2023. The need for unsupervised outlier model selection: A review and evaluation of internal evaluation strategies. *ACM SIGKDD Explorations Newsletter* 25, 1 (2023).

[30] Pankaj Malhotra, Lovekesh Vig, Gautam Shroff, Puneet Agarwal, et al. 2015. Long Short Term Memory Networks for Anomaly Detection in Time Series.. In *Esann*, Vol. 2015. 89.

[31] Kathy McKeown, Hal Daume III, Snigdha Chaturvedi, John Paparrizos, Kapil Thadani, Pablo Barrio, Or Biran, Suvarna Bothe, Michael Collins, Kenneth R Fleischmann, et al. 2016. Predicting the impact of scientific concepts using full-text features. *Journal of the Association for Information Science and Technology* 67, 11 (2016), 2684–2696.

[32] M. Munir, S. A. Siddiqui, A. Dengel, and S. Ahmed. 2019. DeepAnT: A Deep Learning Approach for Unsupervised Anomaly Detection in Time Series. *IEEE Access* 7 (2019), 1991–2005.

[33] Jose Manuel Navarro, Alexis Huet, and Dario Rossi. 2023. Meta-Learning for Fast Model Recommendation in Unsupervised Multivariate Time Series Anomaly Detection. In *AutoML Conference 2023*.

[34] Themis Palpanas and Volker Beckmann. 2019. Report on the First and Second Interdisciplinary Time Series Analysis Workshop (ITISA). *SIGMOD Rec.* 48, 3 (2019), 36–40.

[35] Ioannis Paparrizos. 2018. *Fast, Scalable, and Accurate Algorithms for Time-Series Analysis*. Ph.D. Dissertation. Columbia University.

[36] John Paparrizos, Paul Boniol, Themis Palpanas, Ruey S. Tsay, Aaron Elmore, and Michael J. Franklin. 2022. Volume under the Surface: A New Accuracy Evaluation Measure for Time-Series Anomaly Detection. *Proc. VLDB Endow.* 15, 11 (2022).

[37] John Paparrizos, Paul Boniol, Themis Palpanas, Ruey S. Tsay, Aaron Elmore, and Michael J Franklin. 2022. Volume Under the Surface: A New Accuracy Evaluation Measure for Time-Series Anomaly Detection. *Technical Report LIPADE-TR-N7, Université Paris Cité* (2022).

[38] John Paparrizos, Ikraduya Edian, Chunwei Liu, Aaron J Elmore, and Michael J Franklin. 2022. Fast adaptive similarity search through variance-aware quantization. In *ICDE*. IEEE, 2969–2983.

[39] John Paparrizos and Michael J Franklin. 2019. Grail: efficient time-series representation learning. *Proc. VLDB Endow* 12, 11 (2019).

[40] John Paparrizos and Luis Gravano. 2015. k-shape: Efficient and accurate clustering of time series. In *SIGMOD*. 1855–1870.

[41] John Paparrizos and Luis Gravano. 2017. Fast and accurate time-series clustering. *TODS* 42, 2 (2017).

[42] John Paparrizos, Yuhao Kang, Paul Boniol, Ruey S. Tsay, Themis Palpanas, and Michael J. Franklin. 2022. TSB-UAD: An End-to-End Benchmark Suite for Univariate Time-Series Anomaly Detection. *Proc. VLDB Endow.* 15, 8 (2022).

[43] John Paparrizos, Chunwei Liu, Bruno Barbarioli, Johnny Hwang, Ikraduya Edian, Aaron J Elmore, Michael J Franklin, and Sanjay Krishnan. 2021. VergeDB: A Database for IoT Analytics on Edge Devices.. In *CIDR*.

[44] John Paparrizos, Chunwei Liu, Aaron J Elmore, and Michael J Franklin. 2020. Debunking four long-standing misconceptions of time-series distance measures. In *SIGMOD*. 1887–1905.

[45] John Paparrizos, Chunwei Liu, Aaron J Elmore, and Michael J Franklin. 2023. Querying Time-Series Data: A Comprehensive Comparison of Distance Measures. *IEEE Data Eng. Bull.* 46, 3 (2023), 69–88.

[46] John Paparrizos and Sai Prasanna Teja Reddy. 2023. Odyssey: An engine enabling the time-series clustering journey. *Proceedings of the VLDB Endowment* 16, 12 (2023), 4066–4069.

[47] John Paparrizos, Ryen W White, and Eric Horvitz. 2016. Detecting devastating diseases in search logs. In *SIGKDD*. 559–568.

[48] John Paparrizos, Ryen W White, and Eric Horvitz. 2016. Screening for pancreatic adenocarcinoma using signals from web search logs: Feasibility study and results. *Journal of oncology practice* 12, 8 (2016), 737–744.

[49] John Paparrizos, Kaize Wu, Aaron Elmore, Christos Faloutsos, and Michael J Franklin. 2023. Accelerating Similarity Search for Elastic Measures: A Study and New Generalization of Lower Bounding Distances. *Proceedings of the VLDB Endowment* 16, 8 (2023), 2019–2032.

[50] Mayu Sakurada and Takehisa Yairi. 2014. Anomaly Detection Using Autoencoders with Nonlinear Dimensionality Reduction. In *MLSDA*.

[51] Sebastian Schmidl, Phillip Wenig, and Thorsten Papenbrock. 2022. Anomaly Detection in Time Series: A Comprehensive Evaluation. *Proc. VLDB Endow.* 15, 9 (2022), 1779–1797.

[52] Emmanouil Sylligardos, Paul Boniol, John Paparrizos, Panos Trahanias, and Themis Palpanas. 2023. Choose wisely: An extensive evaluation of model selection for anomaly detection in time series. *Proceedings of the VLDB Endowment* 16, 11 (2023), 3418–3432.

[53] Nesime Tatbul, Tae Jun Lee, Stan Zdonik, Mejbah Alam, and Justin Gottschlich. 2018. Precision and Recall for Time Series. In *NeurIPS*, Vol. 31.

[54] Qingsong Wen, Linxiao Yang, Tian Zhou, and Liang Sun. 2022. Robust Time Series Analysis and Applications: An Industrial Perspective. In *SIGKDD*.

[55] Renjie Wu and Eamonn J. Keogh. 2023. Current Time Series Anomaly Detection Benchmarks are Flawed and are Creating the Illusion of Progress. *IEEE Trans. Knowl. Data Eng.* 35, 3 (2023), 2421–2429. https://doi.org/10.1109/TKDE.2021.3112126

[56] Guang Yang, Ninad Kulkarni, Paavani Dua, Dipika Khullar, and Alex Anto Chirayath. 2022. Anomaly Detection for Spatiotemporal Data in Action. In *SIGKDD*.

[57] Hangting Ye, Zhining Liu, Xinyi Shen, Wei Cao, Shun Zheng, Xiaofan Gui, Huishuai Zhang, Yi Chang, and Jiang Bian. 2023. UADB: Unsupervised Anomaly Detection Booster. In *2023 IEEE 39th International Conference on Data Engineering (ICDE)*. IEEE Computer Society, 2593–2606.

[58] Arthur Zimek, Ricardo JGB Campello, and Jörg Sander. 2014. Ensembles for unsupervised outlier detection: challenges and research questions a position paper. *Acm Sigkdd Explorations Newsletter* 15, 1 (2014), 11–22.