New Trends in Time-Series Anomaly Detection

Paul Boniol
LIPADE, Université Paris Cité
boniol.paul@gmail.com

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

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

ABSTRACT

Anomaly detection is an important problem in data analytics with applications in many domains. In recent years, there has been an increasing interest in anomaly detection tasks applied to time series. In this tutorial, we take a holistic view on anomaly detection in time series, starting from the core definitions and taxonomies related to time series and anomaly types, to an extensive description of the anomaly detection methods proposed by different communities in the literature. Then, we discuss shortcomings in traditional evaluation measures. Finally, we present new solutions to assess the quality of anomaly detection approaches and new benchmarks capturing diverse domains and applications.

1 INTRODUCTION

A wide range of sensing, networking, storage, and processing solutions enable the collection of enormous amounts of measurements over time [45, 53, 59]. The recording of these measurements results in an ordered sequence of real-valued data points commonly referred to as *time series*. Analytical tasks over time series data are becoming increasingly important in virtually every domain [6, 20, 30, 31, 35, 46, 48–51, 54], including astronomy [28], biology [5], energy sciences [3], environmental sciences [23], medicine [55], and social sciences [18].

Anomaly detection has received ample academic and industrial attention [21, 44]. Moreover, as illustrated in Figure 1, anomaly detection applied to time series (compared to other data types) is attracting more interest lately. As commonly defined in the literature [7, 27], *anomalies* refer to data points (single points or group of 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 [1]: (i) noise or erroneous data (e.g., broken sensors), or (ii) actual data of interest (e.g., anomalous behavior of the measured system). In both cases, detecting such cases is crucial for many applications [2, 26].

In recent years, many research works have appeared in this area of time-series anomaly detection. Multiple surveys and experimental benchmarks have been written to summarize and analyze the state-of-the-art proposed methods [8, 29, 47, 52, 57]. Such surveys and benchmarks provide a holistic view of anomaly detection methods and how they perform on benchmarks.

Therefore, based on these recent works, this tutorial provides a comprehensive view of the task of anomaly detection in time series. We start from terminology and definitions for time series, anomaly, and method types to appropriate accuracy evaluation measures. Our goal is three-fold: (i) introduce the motivations and the challenges related to the anomaly detection task in time series by describing a taxonomy of time series and anomaly type usually considered in the literature, (ii) describe the category of anomaly detection methods proposed in the literature as well as their comparisons on recently proposed benchmarks, (iii) discuss

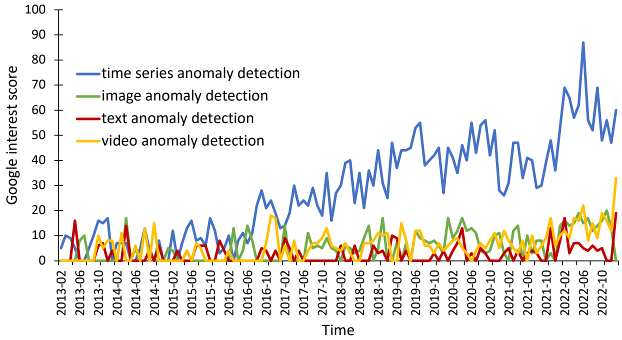

Figure 1: Evolution of the Google interest score for anomaly detection over time series, images, text, and video.

the different solutions to evaluate anomaly detection methods and assess the robustness of the evaluation measures. Finally, we discuss the open research problems and opportunities for anomaly detection in time series.

2 TIME SERIES ANOMALY DETECTION

In this tutorial, we will go through the problem of anomaly detection in time series, starting from very basic definitions of time series and anomaly to analyzing the appropriate evaluation measures to use for the specific case of anomaly detection.

2.1 Introduction, Motivation, and Foundations

We will start by discussing examples of scientific domains and industrial applications that rely on large time series collections and need to perform anomaly detection. We illustrate this variety of domains by showing concrete time series with real anomalies from multiple domains. Figure 2 displays multiple of examples of time series and anomalies from various domains. In Figure 2, ECG [41], MITDB [42], SVDB [24] are ElectroCardioGram (ECG) in which anomalies are either irregular heartbeats or arrhythmia. Moreover, SensorScope [61] is a collection of environmental data, such as temperature, humidity, and solar radiation, collected from a typical tiered sensor measurement system, and Daphnet [4] contains the annotated readings of 3 acceleration sensors at the hip and leg of Parkinson's disease patients that experience freezing of gait (FoG) during walking tasks.

While these previously described examples illustrate well the variety of domains and application in the scope of time series anomaly detection, we then introduce some foundational aspects of time series and anomaly detection. We first introduce the different types of time series. Specifically, we define a **univariate** time series as an ordered sequence of real values on a single dimension. In this case, a subsequence can be represented as a vector. Then, we define a **multivariate** time series as either a set of ordered sequences of real values (with each ordered sequence having the same length) or an ordered sequence of vectors composed of real values. In this specific case, a subsequence is a matrix in which each line corresponds to a subsequence of one single dimension. Moreover, a core characteristic of time series is their evolution with time. Therefore, we define **static** time series as sequences with a fixed length. In this case, one does

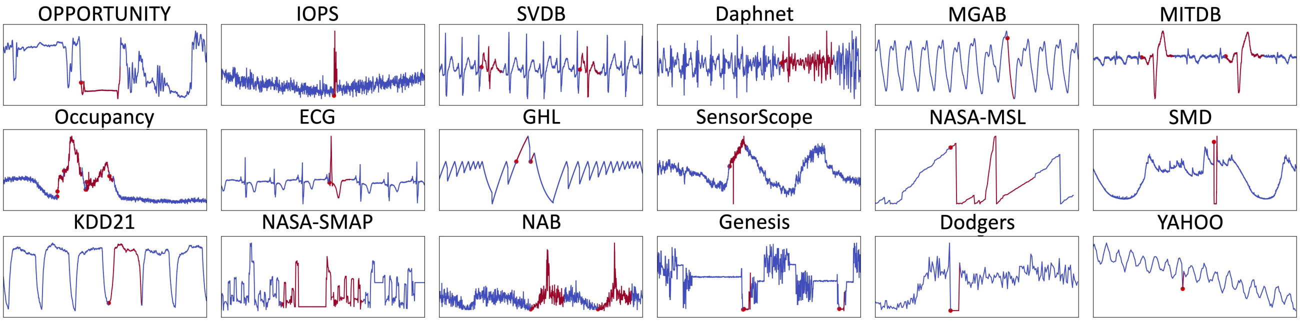

Figure 2: Example of time series from different domains with anomalies (in red) of different types [52].

not expect more values to be added and can analyze points or subsequences simultaneously. We also define **streaming** time series as sequences with an infinite length with new points or subsequences continuously arriving. In this case, models need to be updated dynamically as new points arriving. Finally, the distribution of the values of a time series might change over time. In this case, we make a distinction between **stationary** (i.e., with a constant distribution of values over time) and **non-stationary** (i.e., with a changing distribution of values over time) time series. Moreover, even though the distribution of values is constant, the normal behavior (i.e., the subsequence representing the normal and recurrent behavior, such as a normal heartbeat) might change over time. In this case, we talk about **single normality** and **multiple normalities** time series.

We then introduce the different types of anomalies. There are three types of time-series anomalies: *point*, *contextual*, and *collective* anomalies. We describe them in the following section.

2.1.1 **Point Anomaly**. The first category, *point* anomalies, refers to data points that deviate remarkably from the rest of the data. The IOPS dataset in Figure 2 depicts an an example of *point* anomaly and Figure 3(a.1.1) illustrates a synthetic example.

2.1.2 **Contextual Anomaly**. Then, *contextual* anomalies refer to data points within the expected range of the distribution (in contrast to point anomalies) but deviate from the expected data distribution, given a specific context (e.g., a window). The YAHOO dataset in Figure 2 depicts an an example of *contextual* anomaly and Figure 3(a.1.2) illustrates a synthetic example.

2.1.3 **Collective Anomaly**. With *collective* anomalies, we refer to sequences of points that do not repeat a typical (previously observed) pattern. The SVDB dataset in Figure 2 depicts an an example of *collective* anomaly and Figure 3(a.2) illustrates a synthetic example.

The first two categories, namely, point and contextual anomalies, are referred to as *point-based* anomalies, whereas *collective* anomalies are referred to as *subsequence* anomalies.

2.1.4 **Single vs. Multiple Anomalies**. Then, on top of these categories, the combination of them also matters. First, we need to differentiate time series containing *single* anomalies from time series containing *multiple* anomalies. Last, the *multiple* time series category has to be divided into two other categories, namely time series containing *multiple different* and *multiple similar* anomalies. For instance, the MITDB and Genesis time series in Figure 2 contains multiple similar anomalies. As described in the next section, some methods that are based on nearest-neighbor distance might be affected by this distinction. Figure 3 illustrates all the aforementioned anomaly types and combinations. For all these definitions and taxonomies, we will provide explicit examples.

2.2 Existing Methods and Benchmarks

We will then dive into the existing anomaly detection methods proposed in the literature. Due to the large variety of applications, domains, and anomaly types, every year, a vast number of papers appear in the literature proposing new methods for anomaly detection in time series, and it is beyond our scope to cover extensively here. In this tutorial, we will only briefly summarize popular categories of methods, and we refer the attendees to three recent survey papers for detailed coverage of methods [8, 16, 25].

We will first mention the three main categories of methods based on the external knowledge provided to them. First, **unsupervised** methods take the time series as input only and are not provided by any information relative to the normal or abnormal behavior. Then, **semi-supervised** methods take as input time series without any anomalies. In this case, the model is 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.

Then, we will describe the following categories of methods (refer to Figure 4):

2.2.1 **Distance-based Methods**. The first family of method is distance-based appraoches. These methods focus on the analysis of subsequences for the purpose of detecting anomalies in time series, mainly by utilizing distances to a given model. For instance, discord-based approaches use the nearest neighbor distances among subsequences [19, 22, 33, 34, 37, 38, 58, 62]. As another example, recent methods in this category first cluster data to obtain the normal behavior and compute the distance to this normal behavior to detect anomalies [9–15].

2.2.2 **Density-based Methods**. Second, instead of measuring nearest neighbor distances, density-based methods focus on detecting globally normal or isolated behaviors. General-purpose multi-dimensional point outlier methods have been proposed in this category [17, 36, 39], with Isolation Forest [36] working particularly well when extended for subsequences [11]. The latter aims to isolate instances (or time series and subsequences in our specific case) by building random splitting trees. The longer the depth of the tree, the more splits were necessary to isolate a given instance. Thus, an anomaly score can be computed based on the depth of the trees.

2.2.3 **Forecasting-based Methods**. Third, forecasting-based methods, such as recurrent neural network-based [40] or convolutional network-based approach [43], have been proposed for this task. Such methods use the past values as input, predict the following one, and use the forecasting error as an anomaly score.

2.2.4 **Reconstruction-based Methods**. Last, reconstruction-based methods, such as the AutoEncoder-based approach [56],

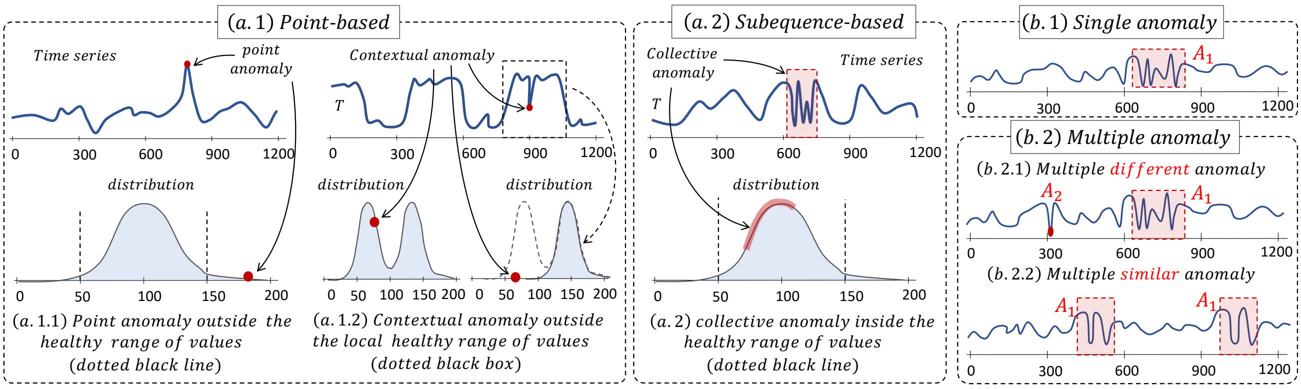

Figure 3: Synthetic examples of different types of anomalies in time series.

are trained to reconstruct the time series and use the reconstruction error as an anomaly score. As both forecasting and reconstruction-based categories detect anomalies using prediction errors (either forecasting or reconstruction error), we can group them into **prediction-based methods**.

Finally, we will discuss recent benchmarks proposed for anomaly detection in time series task [29, 32, 47, 52, 57]. Such benchmarks provide a large 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. Figure 2 illustrates several time series from the TSB-UAD benchmark [52] that contains 18 different datasets from different domains and applications, with a total of more than 2000 time series.

2.3 Evaluating Anomaly Detection

In the last part of this tutorial, we will focus on evaluating anomaly detection accuracy. Indeed, the choice of measure to quantify the quality of methods may also significantly bias the experimental outcome. Thus, 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 [60] that aim to overcome shortcomings of traditional measures when evaluating time series containing subsequence anomalies. We will discuss AUC (i.e., Area Under the Curve) measures that, contrary to previously mentioned measures, eliminate the need to define a threshold. We will finally discuss VUS [47] (i.e., Volume Under the Surface) measures that provide more robustness for time series.

We will conclude this section on the complex question of evaluating evaluation measures. We will mainly discuss the results and the methodology proposed in a recent benchmark for evaluation measures [47, 52]. This benchmark proposes an experimental evaluation of all the aforementioned measures by comparing them with regard to (i) robustness (of labeling misalignment, anomaly score noise, and normal/abnormal ratio), (ii) separability (i.e., the ability to observe a significant difference between anomaly scores of accurate and inaccurate methods), and (iii) consistency (i.e., the ability of a measure to provide similar accuracy values for a same method applied to two similar time series). In this tutorial, we will discuss the conclusion of such an experimental evaluation, but also its limitations and challenges.

2.4 Challenges and Conclusions

We will conclude this tutorial by summarizing the main insights obtained from recent benchmarks on the performances anomaly detection methods [52] and the proper evaluation measures to evaluate them [47]. We will discuss the new problems that these insights opened. Finally, we will elaborate on new ideas and research directions (such as ensembling solutions and model selection methods) that could solve the new open problems.

Acknowledgments: Work supported in part by Meta Research.

REFERENCES

[1] Charu C Aggarwal. 2017. An introduction to outlier analysis. In *Outlier analysis*. Springer, 1–34.
[2] Jerome Antoni and Pietro Borghesani. 2019. A statistical methodology for the design of condition indicators. *Mechanical Systems and Signal Processing* (2019).
[3] Martin Bach-Andersen, Bo Rømer-Odgaard, and Ole Winther. 2017. Flexible non-linear predictive models for large-scale wind turbine diagnostics. *Wind Energy* 20, 5 (2017), 753–764.
[4] Marc Bachlin, Meir Plotnik, Daniel Roggen, Inbal Maidan, Jeffrey M. Hausdorff, Nir Giladi, and Gerhard Troster. 2010. Wearable Assistant for Parkinson's Disease Patients With the Freezing of Gait Symptom. *IEEE Transactions on Information Technology in Biomedicine* 14, 2 (2010), 436–446. https://doi.org/10.1109/TITB.2009.2036165
[5] Ziv Bar-Joseph, Georg K Gerber, David K Gifford, Tommi S Jaakkola, and Itamar Simon. 2003. Continuous representations of time-series gene expression data. *Journal of Computational Biology* 10, 3-4 (2003), 341–356.
[6] 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.
[7] Vic Barnett and Toby Lewis. 1984. Outliers in statistical data. *Wiley Series in Probability and Mathematical Statistics. Applied Probability and Statistics* (1984).
[8] 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).
[9] Paul Boniol, Michele Linardi, Federico Roncallo, and Themis Palpanas. 2020. Automated Anomaly Detection in Large Sequences. In *ICDE*.
[10] Paul Boniol, Michele Linardi, Federico Roncallo, Themis Palpanas, Mohammed Meftah, and Emmanuel Remy. 2021. Unsupervised and scalable subsequence anomaly detection in large data series. *The VLDB Journal* 30, 6 (2021).
[11] Paul Boniol and Themis Palpanas. 2020. Series2Graph: Graph-based Subsequence Anomaly Detection for Time Series. *PVLDB* 13, 11 (2020).
[12] Paul Boniol, Themis Palpanas, Mohammed Meftah, and Emmanuel Remy. 2020. GraphAn: Graph-based Subsequence Anomaly Detection. *Proc. VLDB Endow.* 13, 12 (2020), 2941–2944.
[13] 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. *Proceedings of the VLDB Endowment* 15, 12 (2022), 3702–3705.
[14] Paul Boniol, John Paparrizos, Themis Palpanas, and Michael J Franklin. 2021. SAND in action: subsequence anomaly detection for streams. *Proceedings of the VLDB Endowment* 14, 12 (2021).
[15] Paul Boniol, John Paparrizos, Themis Palpanas, and Michael J Franklin. 2021. SAND: streaming subsequence anomaly detection. *Proceedings of the VLDB Endowment* 14, 10 (2021), 1717–1729.
[16] Mohammad Braei and Sebastian Wagner. 2020. Anomaly detection in univariate time-series: A survey on the state-of-the-art. *arXiv:2004.00433* (2020).

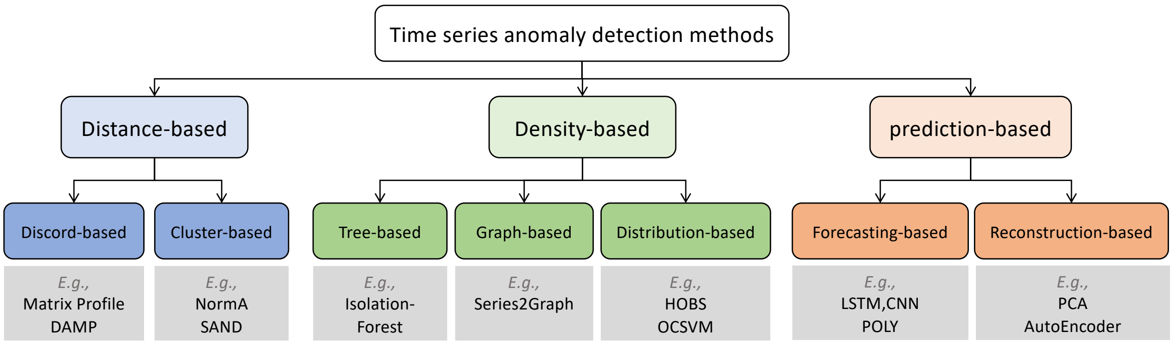

Figure 4: Taxonomy of anomaly detection methods.

[17] Markus M. Breunig, Hans-Peter Kriegel, Raymond T. Ng, and Jörg Sander. 2000. LOF: Identifying Density-based Local Outliers. In *SIGMOD*.

[18] Peter J Brockwell and Richard A Davis. 2016. *Introduction to time series and forecasting*. springer.

[19] Yingyi Bu, Oscar Tat-Wing Leung, Ada Wai-Chee Fu, Eamonn J. Keogh, Jian Pei, and Sam Meshkin. 2007. WAT: Finding Top-K Discords in Time Series Database. In *SIAM*.

[20] 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.

[21] Anthony J Fox. 1972. Outliers in time series. *Journal of the Royal Statistical Society: Series B (Methodological)* 34, 3 (1972), 350–363.

[22] Ada Wai-Chee Fu, Oscar Tat-Wing Leung, Eamonn J. Keogh, and Jessica Lin. 2006. Finding Time Series Discords Based on Haar Transform. In *ADMA*.

[23] Steve Goddard, Sherri K Harms, Stephen E Reichenbach, Tsegaye Tadesse, and William J Waltman. 2003. Geospatial decision support for drought risk management. *Commun. ACM* 46, 1 (2003), 35–37.

[24] Scott David Greenwald. 1990. *Improved detection and classification of arrhythmias in noise-corrupted electrocardiograms using contextual information.* Thesis. Massachusetts Institute of Technology. https://dspace.mit.edu/handle/1721.1/29206 Accepted: 2005-10-07T20:45:22Z.

[25] Manish Gupta, Jing Gao, Charu C Aggarwal, and Jiawei Han. 2013. Outlier detection for temporal data: A survey. *IEEE Transactions on Knowledge and data Engineering* 26, 9 (2013), 2250–2267.

[26] 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*.

[27] Douglas M Hawkins. 1980. *Identification of outliers*. Vol. 11. Springer.

[28] Pablo Huijse, Pablo A Estevez, Pavlos Protopapas, Jose C Principe, and Pablo Zegers. 2014. Computational intelligence challenges and applications on large-scale astronomical time series databases. *IEEE Computational Intelligence Magazine* 9, 3 (2014), 27–39.

[29] 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.

[30] 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. *Proceedings of the VLDB Endowment* 13, 6 (2020), 925–938.

[31] 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.

[32] 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/.

[33] Eamonn Keogh, Stefano Lonardi, Chotirat Ann Ratanamahatana, Li Wei, Sang-Hee Lee, and John Handley. 2007. Compression-based data mining of sequential data. *Data Mining and Knowledge Discovery* (2007).

[34] Michele Linardi, Yan Zhu, Themis Palpanas, and Eamonn J. Keogh. 2020. Matrix Profile Goes MAD: Variable-Length Motif And Discord Discovery in Data Series. In *DAMI*.

[35] Chunwei Liu, Hao Jiang, John Paparrizos, and Aaron J Elmore. 2021. Decomposed bounded floats for fast compression and queries. *Proceedings of the VLDB Endowment* 14, 11 (2021), 2586–2598.

[36] Fei Tony Liu, Kai Ming Ting, and Zhi-Hua Zhou. 2008. Isolation Forest. In *2008 Eighth IEEE International Conference on Data Mining*. 413–422. ISSN: 2374-8486.

[37] Yubao Liu, Xiuwei Chen, and Fei Wang. 2009. Efficient Detection of Discords for Time Series Stream. *Advances in Data and Web Management* (2009).

[38] Yue Lu, Renjie Wu, Abdullah Mueen, Maria A. Zuluaga, and Eamonn Keogh. 2022. Matrix Profile XXIV: Scaling Time Series Anomaly Detection to Trillions of Datapoints and Ultra-Fast Arriving Data Streams. In *SIGKDD*.

[39] Haoran Ma, Benyamin Ghojogh, Maria N. Samad, Dongyu Zheng, and Mark Crowley. 2020. Isolation Mondrian Forest for Batch and Online Anomaly Detection. arXiv:2003.03692 [cs.LG]

[40] Pankaj Malhotra, Lovekesh Vig, Gautam Shroff, and Puneet Agarwal. 2015. Long Short Term Memory Networks for Anomaly Detection in Time Series. (2015).

[41] G.B. Moody and R.G. Mark. 2001. The impact of the MIT-BIH Arrhythmia Database. *IEEE Engineering in Medicine and Biology Magazine* 20, 3 (2001), 45–50. https://doi.org/10.1109/51.932724

[42] George B Moody and Roger G Mark. 1992. MIT-BIH Arrhythmia Database. https://doi.org/10.13026/C2F305

[43] 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.

[44] ES Page. 1957. On problems in which a change in a parameter occurs at an unknown point. *Biometrika* 44, 1/2 (1957), 248–252.

[45] 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. https://doi.org/10.1145/3377391.3377400

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

[47] 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).

[48] 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.

[49] John Paparrizos and Michael J Franklin. 2019. Grail: efficient time-series representation learning. *Proceedings of the VLDB Endowment* 12, 11 (2019).

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

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

[52] 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).

[53] 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*.

[54] 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.

[55] Joshua S Richman and J Randall Moorman. 2000. Physiological time-series analysis using approximate entropy and sample entropy. *American Journal of Physiology-Heart and Circulatory Physiology* 278, 6 (2000), H2039–H2049.

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

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

[58] Pavel Senin, Jessica Lin, Xing Wang, Tim Oates, Sunil Gandhi, Arnold P. Boedihardjo, Crystal Chen, and Susan Frankenstein. 2015. Time series anomaly discovery with grammar-based compression. In *EDBT*.

[59] 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. (2023).

[60] Nesime Tatbul, Tae Jun Lee, Stan Zdonik, Mejbah Alam, and Justin Gottschlich. 2018. Precision and Recall for Time Series. In *Advances in Neural Information Processing Systems*, Vol. 31. Curran Associates, Inc.

[61] Yuan Yao, Abhishek Sharma, Leana Golubchik, and Ramesh Govindan. 2010. Online anomaly detection for sensor systems: A simple and efficient approach. *Performance Evaluation* 67, 11 (2010), 1059–1075. https://doi.org/10.1016/j.peva.2010.08.018 Performance 2010.

[62] Chin-Chia Michael Yeh, Yan Zhu, Liudmila Ulanova, Nurjahan Begum, Yifei Ding, Hoang Anh Dau, Diego Furtado Silva, Abdullah Mueen, and Eamonn J. Keogh. 2016. Matrix Profile I: All Pairs Similarity Joins for Time Series. In *ICDM*.