

9th SIGKDD International Workshop on Mining and Learning from Time Series

Sanjay Purushotham
University of Maryland,
Baltimore County
Baltimore, Maryland, USA
psanjay@umbc.edu

Dongjin Song
University of Connecticut
Storrs, Connecticut, USA
dongjin.song@uconn.edu

Qingsong Wen
Alibaba Group (U.S.) Inc
Bellevue, Washington
USA
qingsong.wen@alibaba-inc.com

Jun Huan
Amazon Web Services
Sunnyvale, California, USA
lukehuan@amazon.com

Cong Shen
University of Virginia
Charlottesville, Virginia
USA
cs7dt@virginia.edu

Yuriy Nevmyvaka
Morgan Stanley
New York, New York, USA
yuriy.nevmyvaka@morganstanley.com

Abstract

Time series data has become pervasive across domains such as finance, transportation, retail, entertainment, and healthcare. This shift towards continuous monitoring and recording, fueled by advancements in sensing technologies, necessitates the development of new tools and solutions. Despite extensive study, the importance of time series analysis continues to increase. However, modern time series data present challenges to existing techniques, including irregular sampling and spatiotemporal structures. Time series mining research is both challenging and rewarding as it connects diverse disciplines and requires interdisciplinary solutions. The goals of this workshop are to (1) highlight the significant challenges that underpin learning and mining from time series data (e.g., irregular sampling, spatiotemporal structure, uncertainty quantification), (2) discuss recent algorithmic, theoretical, statistical, or systems-based developments for tackling these problems, and (3) to synergize the research activities and discuss both new and open problems in time series analysis and mining. In summary, our workshop will focus on both the theoretical and practical aspects of time series data analysis and will provide a platform for researchers and practitioners from academia and industry to discuss potential research directions and critical technical issues and present solutions to tackle related issues in practical applications. We will invite researchers and practitioners from the related areas of AI, machine learning, data science, statistics, and many others to contribute to this workshop.

CCS Concepts

• **Mathematics of computing** → **Time series analysis**; • **Information systems** → **Data mining**; • **Computing methodologies** → **Machine learning**.

Keywords

time-series analysis, temporal data mining, deep forecasting

ACM Reference Format:

Sanjay Purushotham, Dongjin Song, Qingsong Wen, Jun Huan, Cong Shen, and Yuriy Nevmyvaka. 2023. 9th SIGKDD International Workshop on Mining and Learning from Time Series. In *Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD '23)*, August 6–10, 2023, Long Beach, CA, USA. ACM, New York, NY, USA, 2 pages. <https://doi.org/10.1145/3580305.3599214>

1 Introduction

Time series analysis has attracted innovative approaches from various disciplines, including data mining, machine learning, statistical learning, dynamical systems, and the physical sciences. This interdisciplinary nature has resulted in wide-ranging applications and significant impact, making it an ideal subject for sharing new ideas at the SIGKDD conference. The SIGKDD Workshop on Mining and Learning from Time Series (MiLeTS) has been the sole dedicated workshop on time series mining in the past eight KDD conferences. The first seven workshops, held from 2015 to 2021, were highly successful, culminating in the 8th SIGKDD MiLeTS workshop in Washington DC during SIGKDD 2022. This workshop received praise for its ability to bring together researchers from academia and industry to address the challenges posed by large and complex time series data. With SIGKDD 2023 approaching, it presents the perfect opportunity and venue to host this multidisciplinary workshop on ubiquitous temporal data mining.

2 Objective and Topics of Interests

The proliferation of new sensing technologies, such as wearable sensors, mobile phones, and satellites, coupled with the increasing affordability and efficiency of storage, has resulted in an unprecedented surge of time series data across various domains such as biology, health and medicine, climate and weather, road traffic, astronomy, and energy. The sheer volume and complexity of this data pose significant challenges to existing methods, including state-of-the-art approaches. The primary focus of this workshop is to foster collaboration among researchers in this field, aiming to address both novel and existing problems in time series analysis and mining. Solutions to these challenges may encompass algorithmic, theoretical, statistical, or systems-based approaches. The workshop aims to serve as a platform for exchanging ideas and exploring

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

KDD '23, August 6–10, 2023, Long Beach, CA, USA

© 2023 Copyright held by the owner/author(s).

ACM ISBN 979-8-4007-0103-0/23/08.

<https://doi.org/10.1145/3580305.3599214>

innovative solutions to effectively navigate the evolving landscape of time series data.

Topics of interest include but are not limited to:

- Time series pattern mining and detection, representation, searching and indexing, classification, clustering, prediction, forecasting, and rule mining.
- Multivariate time series that are high-dimensional, have heterogeneous variables, or other atypical properties.
- Time series with special structure: spatiotemporal (e.g., traffic speeds at different locations), relational (e.g., patients with similar diseases), hierarchical, etc.
- Time series with sparse or irregular sampling, missing values at and not at random, and special types of measurement noise or bias.
- Time series analysis using less traditional approaches, such as deep learning and subspace clustering.
- Privacy preserving time series mining and learning.
- Online, high-speed learning and mining from streaming time series.
- Uncertain time series mining.
- Applications to high impact or relatively new time series domains, such as health and medicine, road traffic, seismology, AIOps, AIoT, and air quality.

In addition to soliciting traditional Research Track papers on the above or related topics, we also aim to encourage the submission of one additional type of paper describing open problems. These papers do not necessarily need to propose new methods; instead, they should describe and analyze a new or unresolved time series problem that presents notable challenges to existing techniques. While these papers are not expected to solve the problem, they should provide a comprehensive and detailed description, along with a thorough empirical investigation that demonstrates the insufficiency of state-of-the-art methods.

Reviewing: The review process is single-round and double-blind. Accepted papers will be presented as posters during the workshop and list on the website¹ (non-archival/without proceedings). Besides, a small number of accepted papers will be selected to be presented as contributed talks.

3 Workshop Format and Schedule

The schedule for the workshop will include the following key elements:

- **Presentation Sessions:** The presentations will be held from 8:00 AM to 12:00 PM and will be divided into two sessions. There will be a break between each session, allowing ample time for discussion and networking.
- **Keynote Presentations:** Esteemed experts in time series analysis, both from academia and industry, will deliver invited keynote presentations. Each keynote presentation will have a duration of 30-45 minutes.
- **Oral Presentations:** Each accepted oral submission will be allocated 10 minutes for presentation, followed by a question-and-answer session
- **Poster Session:** Depending on the level of interest, we may incorporate an additional session dedicated to demonstrations or

posters during the break. Each accepted paper will have the opportunity to either present a demonstration or showcase their work as a poster during this session.

- **Panel Discussion:** The final session will feature a panel discussion on current and future topics in large-scale time series analysis. The panelists will represent academia, federal agencies, and industry, bringing diverse perspectives to the discussion.

These elements aim to provide a comprehensive and engaging program that encourages interaction, knowledge sharing, and collaboration among participants.

3.1 Target Audience

This workshop appeals to two groups of researchers: methodologists who design algorithms and models for time series analysis, and researchers from application domains generating complex time series data. The former group spans various backgrounds (e.g., signal processing, data mining, machine learning) and is interested in problem-focused workshops. The latter group includes researchers from application domains (e.g. health, climate, traffic, mobile video, etc.) that recognize the value of their data but need help with analysis, making the workshop an opportunity to learn from methodologists. The panel discussion will address open problems in methodologies and applications, benefiting student researchers seeking dissertation topics and expert guidance.

4 Workshop Organizers

4.1 Program Committee

We are very grateful to each of our Program Committee members for their reviews of submissions and crucial contributions. The members of the Program Committee are: Zhengping Che, Dongsheng Luo, Yingtao Luo, Yushan Jiang, Dingsu Wang, Yi-Fan Zhang, Renjie Wu, Yiyuan Yang, Jing Dai, Jiechao Gao, Shima Imani, Feiyang Cai, Yuxuan Liang, Haomin Wen, Minghua Ma, Xiang Li, Lili Meng, Ming Jin, Xingjian Shi, Pengyang Wang, Chen Xu, Dongkuan Xu, An Wang, John Paparrizos.

4.2 Steering Committee

- Eamonn Keogh, University of California Riverside
- Yan Liu, University of Southern California
- Abdullah Mueen, University of New Mexico

4.3 Organization Committee

- Sanjay Purushotham, University of Maryland, Baltimore County
- Jun (Luke) Huan, AWS AI Labs
- Cong Shen, University of Virginia
- Dongjin Song, University of Connecticut
- Qingsong Wen, DAMO Academy, Alibaba Group
- Yuriy Nevmyvaka, Morgan Stanley

Acknowledgments

We would like to extend our sincere appreciation to Morgan Stanley for their generous sponsorship of our workshop.

¹<https://kdd-milets.github.io/milets2023/>