Guest Editorial Robust Learning of Spatio-Temporal Point Processes: Modeling, Algorithm, and Applications

TEMPORAL data are ubiquitous in real-world applications, and they can be generally divided into two categories: 1) synchronous temporal data which are basically equivalent to time series data; and 2) the asynchronous data which are often in the form of event data with a time stamp in continuous time-space. In fact, the event data are often converted to the time series by aggregating the event count in equal time intervals in many previous approaches. While it is often of one's greater interest to directly establish models based on the raw event data whose time stamps carry useful information, especially for those time-sensitive tasks, ranging from earthquake prediction, crime analysis, to infectious disease diffusion forecasting, etc. Developing the spatio-temporal point process and the related applications is the theme of this Special Issue, which treats an event as a point in the spatio-temporal space, with possibly extra attributes. The model captures the instantaneous happening rate of the events and their potential dependency. The derived use cases often refer to future events prediction, and causality estimation.

This Special Issue presents 19 original articles covering topics of machine learning of temporal point process, especially for robust models with algorithm design and applications in different domains. All the papers have undergone through a rigorous review process. We try to categorize these works into three aspects, based on their relevance and contribution to the temporal point process by different means.

The first category includes articles with clear methodological novelty for temporal data learning, including both with explicit temporal point process models and those not.

In [A1], Yang et al. identify an important problem in practice that many temporal data can be missing in segments, and they then present a generative adversarial network-based approach for this task. Specifically, they deal with the time series-formatted temporal data, while it remains open for how to fulfill this data imputation goal for more complex event data in point process, as the recovery needs to deal with the time stamp in the continuous time domain (or in addition, types and other attributes) of events rather than only the value.

In [A2], Zhang et al. show how to use the thermodynamic expressions for motif representation and detection from time-evolving complex networks, with potential applications to dynamic stock correlation analysis. A partition function is devised to capture the motif, with an analogy to the cluster expansion in statistical physics.

In [A3], Chapfuwa et al. propose a neural network-based approach for time-to-event prediction, whereby the event data actually forms a temporal point process in the continuous time domain. This article perfectly fits well with the Special Issue and the targeted problem is important especially for neural models which are often thought lacking enough physical meaning or rigorous mathematical treatment, especially when the occurrence probability needs to be carefully considered.

In [A4], Wu et al. propose a Graph-Biased Temporal Point Process which leverages the structural information from graph representation learning. The authors attempt to model direct influence between nodes and indirect influence from event history, respectively, and integrate the learned node embedding vector into the embedded event history as side information.

The second category focuses on the domain-specific method for temporal data learning, especially with interesting and important applications.

In [A5], Cheng et al. address a practical problem of sequential delinquent event prediction for networked-guarantee loans which is a good testbed for temporal point process-related methods. Specifically, the authors propose a temporal default prediction model that integrates a temporal guarantee network, loan behavior sequences, and small- and medium-sized enterprises (SME) features in recurrent attention mechanisms. Meanwhile, it enjoys good interpretability to discover the risk patterns. This work is also one of the few which considers both the temporal network structure and loan behavior sequences in a unified graph-based neural attention network for delinquent loan event prediction.

In [A6], He et al. refer to a real-world anomaly detection system for cloud computing, whereby the event logs are transformed to time series for processing with a long short-term memory (LSTM), and the topology of the network system is encoded by a graph neural network (GNN). As a result, they propose the graphLSTM approach to combine both the topology and temporal information for more comprehensive modeling of the multidimensional temporal data. They show a successful deployment of their methods as well as make their used real-world data publicly available.

In [A7], the authors developed a tailored model for the problem of social link inference in a given location-based social network, whereby the location information and especially the user trajectories are particularly considered and modeled. The work devises a novel multiview matching network (MVMN)

Digital Object Identifier 10.1109/TNNLS.2023.3258199

2162-237X © 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See https://www.ieee.org/publications/rights/index.html for more information.

module by regarding each of the spatial, temporal, and social factors of user pairs as different views for user link prediction. The experimental results demonstrate the proposed model consistently outperforms strong competitors proposed for this problem and related ones, with ablation studies verifying the rationality of the model design.

In [A8], Lou et al. built a probabilistic regularized extreme learning adaptive neuro-fuzzy inference system to capture the inner correlations among sequential traffic flow data and improve the accuracy of short-term traffic flow forecasting. The model adopts a novel objective function that minimizes both the mean and the variance of the model bias. The experiments conducted on real-world data collected from six roads in a Minnesotan city show the proposed method could raise the prediction efficiency compared with existing models.

In [A9], Li et al. proposed a general asynchronous spatio-temporal spike metric which considers both spatio-temporal structural properties and polarity attributes for event cameras. This work serves as a major step toward building an effective spike metric for event-based signal processing applied to neuromorphic engineering, spatio-temporal data compression, and machine learning to event-based vision.

In [A10], Zhang presented a novel adversarial transformer to generate music pieces with high musicality. This novel method could compute effective rewards for each generated time step in the long sequence, and the author devises the global and local objective functions combined with a scheduled procedure to guide the discriminator training. The experimental results demonstrate that the proposed model can generate long music pieces with improved quality compared to the original music transformers.

In [A11], Gao et al. raised an end-to-end method to better understand the underlying human mobility and improve the modeling of the POIs' transitional distribution in human moving patterns. In the proposed model, a trip encoder embeds the contextual route into a latent variable with a recurrent neural network, and a trip decoder reconstructed this route conditioned on an optimized latent space. The authors also define an Adversarial Net composed of a generator and a critic, which generates a representation for a given query and uses the critic to distinguish the trip representation generated from TE and the query representation obtained from Adversarial Net.

In [A12], Wang et al. proposed a novel point process-based framework for robust multiple object tracking (MOT). The proposed framework can effectively predict and mask out the noisy and confusing detection results before associating the objects into trajectories, and thus suppress their negative influences on final tracking results. In particular, the authors formulated noisy and confusing detection results as a sequence of events and applied a neural network-based spatio-temporal point process to model the dynamics of such events. The experimental results show that the proposed model captures both temporal and spatial evolution of noisy and confusing detection results, which improves the robustness of state-of-the-art MOT methods.

In [A13], Liang and Zhang proposed a model, called ARNPP-GAT, for the check-in inference problem, which aims to jointly infer the next check-in location (Where) and time (When) for a target user (Who) based on the user mobility data. The proposed model combines an attention-based recurrent neural point process (ARNPP) with a graph attention network (GAT). The GAT captures the social relations among different users, and the ARNPP models the dynamics of the users' sequential behaviors. The authors learned the model in a multitask learning framework, considering the impacts of social relations among different users on the sequential behaviors of individuals. Experiments on two real-world datasets demonstrate the effectiveness of the ARNPP-GAT model and its learning algorithm.

In [A14], Bacciu et al. discuss a neural point process approach for health and behavioral data, which comprises both sparse events coming from user subjective declarations and fast-flowing time series from wearable sensors. The authors validate different neural architectures and assess the effect of including input sources of different nature. The experimental results imply the potential of neural point processes in both future event type and user interaction time prediction.

In [A15], Bai et al. proposed a novel framework to measure the similarity between time-varying financial networks. For a time-varying financial network, whose vertices are the time series of stocks and edges are two stocks' Pearson correlation coefficients at different timestamps, the authors first defined and computed a commute time matrix. Based on the commute time matrix, they identify a reliable set of dominant correlated time series as well as an associated dominant probability distribution of stocks belonging to this set. Each network is represented as a discrete dominant Shannon entropy time series, based on the dominant probability distribution. Given arbitrary two networks, the authors applied the classic dynamic time warping method to the corresponding discrete dominant Shannon entropy time series and derived the similarity between the two networks accordingly.

The last group of works is a bit miscellaneous which loosely relate to temporal point process and their applications. While we believe these works can be of help in expanding the applicability of the temporal point process to more real-world settings where the event information needs to be extracted from other supporting machine learning techniques. They may shed light on the area of the point process when more information, e.g., multimodal data, is introduced to develop more comprehensive models including both TPP learning and other information extraction modules. The point is that for real-world especially open-world applications, TPP alone is not the panacea, instead more related technologies need to be combined, and it is practically needed to combine a few representative works into this Special Issue.

In [A16], Ma et al. are concerned with the problem of decorrelating Non-Gaussian Neutral Vector Variables. The authors claimed that the article has connections with the point process with its points distributed in the specific plane with L_1 constraints. Then the authors empirically evaluate a suite of decorrelating methods for performance comparison. Despite its seemingly loose relevance to the topic of the special issue, it does also well show some exploitative area with potential connection to point process, which may consequently prompts the research with more expanded scope and impact to the whole machine learning and data mining society.

In [A17], Jin et al. deal with multimodal learning of image/video texts for retrieval, and the presented neural hashing techniques can be a general approach for efficient multimodal feature extraction. In fact, the events of a point process are often extracted from the raw texts, and video data rather than ideally defined and recorded in the raw data.

In [A18], Zhang et al. developed a new model for weakly supervised action recognition and localization. The proposed model decomposes visual features into domain-adaptable features and domain-specific ones and transfers reliable knowledge between trimmed and untrimmed action videos in a bidirectional way, which only requires video-level annotations. The authors implemented this model with an auto-encoding architecture and learned it to facilitate effective action classification and temporal localization. This work may inspire the weakly supervised learning of TPP models.

In [A19], Qiu et al. proposed an effective method for mining negative sequential patterns (NSP) from discrete symbolic sequences, which loosens some inflexible constraints in NSP mining and captures nonoccurring behaviors with low computational complexity. The authors provided a new definition of negative containment with loose constraints and developed an efficient algorithm, called NegI-NSP, to identify significant NSPs. The methodology in this work has the potential to the nonoccurring event modeling challenge in the study of TPP.

Junchi Yan
Department of Computer Science and Engineering
Shanghai Jiao Tong University
Shanghai 200240, China

Hongteng Xu
Gaoling School of Artificial Intelligence
Renmin University of China
Beijing 100872, China

Liangda Li
Newsbreak
Mountain View, CA 94043 USA

Mehrdad Farajtab
Apple
Redmond, WA 98052 USA

Xiaokang Yang
Institute of Artificial Intelligence
Shanghai Jiao Tong University
Shanghai 200240, China

Appendix: Related Articles

[A1] S. Yang, M. Dong, Y. Wang, and C. Xu, "Adversarial recurrent time series imputation," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, Aug. 4, 2020, doi: 10.1109/TNNLS.2020.3010524.

[A2] Z. Zhang, D. Chen, L. Bai, J. Wang, and E. R. Hancock, "Graph motif entropy for understanding time-evolving networks," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, Oct. 13, 2023, doi: 10.1109/TNNLS.2020.3027426.

[A3] P. Chapfuwa et al., "Calibration and uncertainty in neural time-to-event modeling," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, Oct. 29, 2020, doi: 10.1109/TNNLS.2020.3029631.

[A4] W. Wu, H. Liu, X. Zhang, Y. Liu, and H. Zha, "Modeling event propagation via graph biased temporal point process," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, Jul. 10, 2020, doi: 10.1109/TNNLS.2020.3004626.

[A5] D. Cheng, Z. Niu, and L. Zhang, "Delinquent events prediction in temporal networked-guarantee loans," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, Oct. 13, 2020, doi: 10.1109/TNNLS.2020.3027346.

[A6] Z. He et al., "A spatio-temporal deep learning approach for unsupervised anomaly detection in cloud systems," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, Oct. 16, 2020, doi: 10.1109/TNNLS.2020.3027736.

[A7] W. Zhang, X. Lai, and J. Wang, "Social link inference via multi-view matching network from spatio-temporal trajectories,"

[A8] J. Lou, Y. Jiang, Q. Shen, R. Wang, and Z. Li, "Probabilistic regularized extreme learning for robust modeling of traffic flow forecasting," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, Oct. 16, 2020, doi: 10.1109/TNNLS.2020.3027822.

[A9] J. Li, Y. Fu, S. Dong, Z. Yu, T. Huang, and Y. Tian, "Asynchronous spatiotemporal spike metric for event cameras," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, Mar. 8, 2021, doi: 10.1109/TNNLS.2021.3061122.

[A10] N. Zhang, "Learning adversarial transformer for symbolic music generation," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, Jul. 2, 2020, doi: 10.1109/TNNLS.2020.2990746.

[A11] Q. Gao, F. Zhou, K. Zhang, F. Zhang, and G. Trajcevski, "Adversarial human trajectory learning for trip recommendation," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, Feb. 23, 2021, doi: 10.1109/TNNLS.2021.3058102.

[A12] T. Wang et al., "Spatio-temporal point process for multiple object tracking," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, Jun. 8, 2020, doi: 10.1109/TNNLS.2020.2997006.

[A13] W. Liang and W. Zhang, "Learning social relations and spatiotemporal trajectories for next check-in inference," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, Oct. 20, 2020, doi: 10.1109/TNNLS.2020.3016737.

[A14] D. Bacciu, D. Morelli, and V. Pandelea, "Modeling mood polarity and declaration occurrence by neural temporal point processes," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, May 13, 2022, doi: 10.1109/TNNLS.2022.3172871.

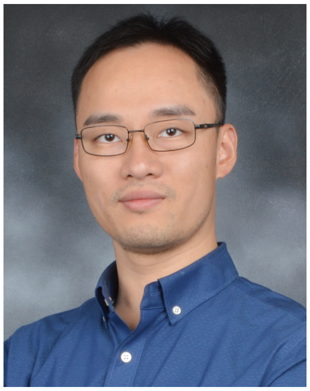
[A15] L. Bai, L. Cui, Z. Zhang, L. Xu, Y. Wang, and E. R. Hancock, "Entropic dynamic time warping kernels for co-evolving financial time series analysis," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, Jul. 21, 2023, doi: 10.1109/TNNLS.2020.3006738.

[A16] Z. Ma et al., "On the comparisons of decorrelation approaches for non-Gaussian neutral vector variables," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, Mar. 31, 2020, doi: 10.1109/TNNLS.2020.2978858.

[A17] L. Jin, Z. Li, and J. Tang, "Deep semantic multimodal hashing network for scalable image-text and video-text retrievals," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, Jun. 5, 2020, doi: 10.1109/TNNLS.2020.2997020.

[A18] X.-Y. Zhang, C. Li, H. Shi, X. Zhu, P. Li, and J. Dong, "AdapNet: Adaptability decomposing encoder–decoder network for weakly supervised action recognition and localization," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, Jan. 23, 2020, doi: 10.1109/TNNLS.2019.2962815.

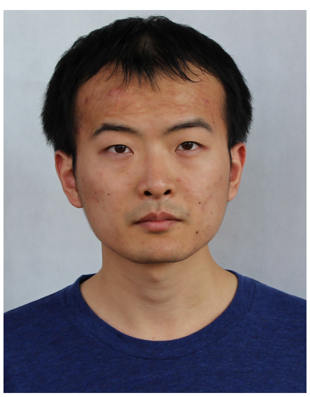
[A19] P. Qiu, Y. Gong, Y. Zhao, L. Cao, C. Zhang, and X. Dong, "An efficient method for modeling nonoccurring behaviors by negative sequential patterns with loose constraints," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, Mar. 17, 2021, doi: 10.1109/TNNLS.2021.3063162.



Junchi Yan (Senior Member, IEEE) was a Senior Research Staff Member with IBM Research, where he started his career in April 2011. He is currently an Associate Professor with the Department of Computer Science and Engineering and the AI Institute, Shanghai Jiao Tong University, Shanghai, China. Before that, he was a Senior Research Staff Member with IBM Research, Shanghai, China, where he started his career in April 2011. His research interests include machine learning and computer vision. He regularly serves as the Senior PC/Area Chair for Neural Information Processing Systems (NeurIPS), International Conference on Machine Learning (ICML), IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), AAAI Conference on Artificial Intelligence (AAAI), and International Joint Conferences on Artificial Intelligence (IJCAI), and as an Associate Editor for *Pattern Recognition* journal.

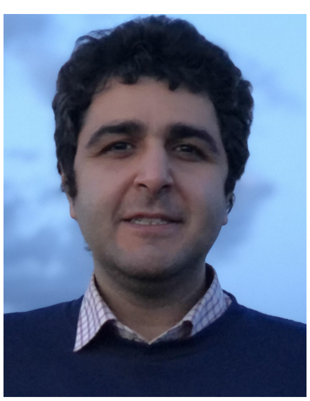
Hongteng Xu (Member, IEEE) received the Ph.D. degree from the School of Electrical and Computer Engineering, Georgia Institute of Technology (Georgia Tech), Atlanta, GA, USA, in 2017.

From 2018 to 2020, he was a Senior Research Scientist with Infinia ML Inc., Durham, NC, USA. During the same period, he was a Visiting Faculty Member with the Department of Electrical and Computer Engineering, Duke University, Durham. He is currently an Associate Professor (Tenure-Track) with the Gaoling School of Artificial Intelligence, Renmin University of China, Beijing, China. His research interests include machine learning and its applications, especially optimal transport theory, sequential data modeling and analysis, deep learning techniques, and their applications in computer vision and data mining.



Liangda Li received the B.S. degree (ACM Hons.) from the School of Computer Science, Shanghai Jiao Tong University, Shanghai, China, in 2010, and the Ph.D. degree from the School of Computer Science, Georgia Institute of Technology (Georgia Tech), Atlanta, GA, USA, in 2015.

He is currently a Research Scientist with NewsBreak, Mountain View, CA, USA. Before that, he was a Senior Research Scientist with Yahoo Lab. His research interests include machine learning and its applications in information retrieval and social networks.



Mehrdad Farajtab received the B.Sc. degree in software engineering and the M.Sc. degree in artificial intelligence from the Department of Computer Engineering, Sharif University of Technology, Tehran, Iran, in 2009 and 2011, respectively, and the Ph.D. degree in computational science and engineering from the Georgia Institute of Technology (Georgia Tech), Atlanta, GA, USA, in 2018, under the supervision of Hongyuan Zha and Le Song.

He used to work on modeling and optimization of events data, stochastic point processes, and dynamics of and on the networks when he was with Georgia Tech. He is currently a Senior Research Scientist with Google DeepMind, Mountain View, CA, USA.

Xiaokang Yang (Fellow, IEEE) received the B.S. degree from Xiamen University, Xiamen, China, in 1994, the M.S. degree from the Chinese Academy of Sciences, Beijing, China, in 1997, and the Ph.D. degree from Shanghai Jiao Tong University, Shanghai, China, in 2000.

He is currently a Distinguished Professor and the Associate Director of the AI Institute, Shanghai Jiao Tong University. His research interests include image processing and computer vision.

Dr. Yang serves as an Associate Editor for IEEE TRANSACTIONS ON MULTIMEDIA.