Event time analysis via point processes and machine learning

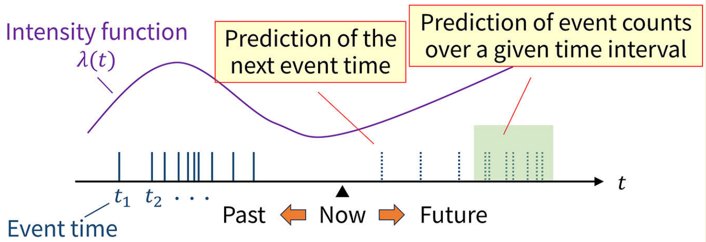
Fast and accurate event timing prediction

Abstract

If the timing of future events can be predicted, risks can be mitigated through proactive preparation, or opportunities can be fully leveraged. In this study, we present a method for efficiently predicting future event times by leveraging point processes and machine learning. While existing approaches achieve high predictive accuracy, they often require substantial computational cost during training. To address this issue, we replace the conventional log-likelihood objective with the least squares contrast for point processes, enabling up to several hundred-fold speedups in training while maintaining comparable predictive performance. This improvement makes our method scalable to large datasets. As machine learning-based event prediction becomes increasingly accurate with the growing volume of data, the associated computational burden also continues to rise. Our approach, which efficiently handles large-scale event series data, aims to support proactive decision-making by accurately predicting when events such as equipment failures or demand fluctuations will occur.

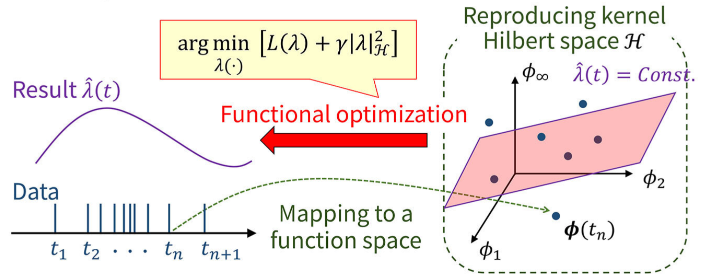
Event time series analysis

- Data that record when and where events occur are referred to as event time series data.
  e.g.) equipment failures, accidents, infections, SNS posts.
- Point processes provide a standard framework for analyzing such data. By learning the intensity function from data, they enable a wide range of analyses and predictions regarding the timing of future events.



Intensity function learning with kernel methods

Kernel methods are a class of machine learning techniques that enable flexible function learning from data without requiring a predefined functional form.



Challenge

For intensity function learning, the log-likelihood has been widely adopted as the loss function $L(\lambda)$ in kernel methods. However, this approach is computationally expensive, making it difficult to scale to large datasets.

Proposed method: Introducing least squares loss

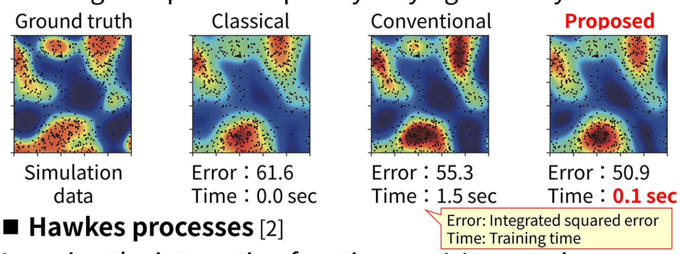
We propose fast kernel methods for estimating intensity functions that replaces the conventional log-likelihood objective with the least squares loss for point processes.

Loss func.	Negative log-likelihood	Least squares loss
	$\int_{\mathcal{T}} \lambda(t)dt - \sum_{n=1}^{N} \log(\lambda(t_n))$	$\int_{\mathcal{T}} \lambda(t)^2dt - 2\sum_{n=1}^{N} \lambda(t_n)$

Conventional methods require solving an optimization problem whose complexity can scale cubically with the size. In contrast, our method eliminates the need for such costly optimization, resulting in significant speedups.

■ Non-homogeneous Poisson processes [1]

Learning the spatio-temporally varying intensity function

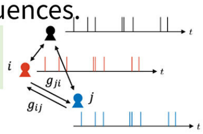


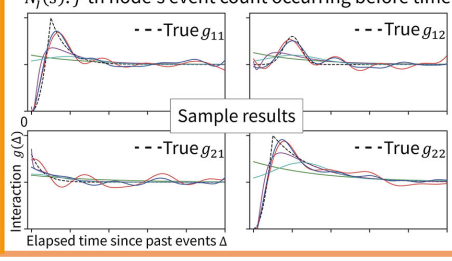
■ Hawkes processes [2]

Learning the interaction functions $g_{ij}(\Delta)$ governing dependencies among multiple event sequences.

$$\lambda_i(t) = \mu_i + \sum_{j\in\mathcal{U}} \int_{\mathcal{T}} g_{ij}(t-s)dN_j(s), \quad i \in \mathcal{U} := \{1, \ldots, U\}$$

U: node number, μ_i: i-th node's base intensity,
$N_j(s)$: j-th node's event count occurring before time s





Comparison	Error	Time
Exponential	2.20	503
Gaussian mixture	1.47	37.3
Exponential mixture	0.71	74.7
Conventional kernel method	0.41	5884
Our method	0.50	**16.1**

References

[1] H. Kim, T. Iwata, A. Fujino, "K2IE: Kernel method-based kernel intensity estimators for inhomogeneous Poisson processes," in *Proc. The 42nd International Conference on Machine Learning (ICML)*, 2025.
[2] H. Kim, T. Iwata, "A representer theorem for Hawkes processes via penalized least squares minimization," in *Proc. The 14th International Conference on Learning Representations (ICLR)*, oral, 2026.

Contact

Hideaki Kim, Learning and Intelligent Systems Research Group, Innovative Communication Laboratory