Learning with Temporal Point Processes

Manuel Gomez Rodriguez [1] **Isabel Valera** [2]

Abstract

In recent years, there has been an increasing number of machine learning models, inference methods and control algorithms using temporal point processes. They have been particularly popular for understanding, predicting, and enhancing the functioning of social and information systems, where they have achieved unprecedented performance. This tutorial aims to introduce the machine learning community at large to temporal point processes. In the first of the tutorial, we will first provide an introduction to the basic theory of temporal point processes, then revisit several types of points processes, and finally introduce advanced concepts such as marks and dynamical systems with jumps. In the second and third parts of the tutorial, we will explain how temporal point processes have been used in developing a variety of recent machine learning models and control algorithms, respectively. Therein, we will revisit recent advances related to, *e.g.*, deep learning, Bayesian nonparametrics, causality, stochastic optimal control and reinforcement learning. In each of the above parts, we will highlight open problems and future research to facilitate further research in temporal point processes within the machine learning community.

1. Outline of the tutorial

The 2-hour tutorial will be divided into three parts. The first part will give a brief introduction to the theory of temporal point processes, starting from the basics. The second part will go through a variety of machine learning models and efficient inference algorithms which utilize temporal point processes in a wide range of application domains. The third part will introduce several recent control algorithms control using temporal point processes and their connections to dynamical systems, stochastic optimal control and reinforcement learning.

1.1. Introduction (30 minutes)

We will first provide an introduction to the basic theory of temporal point processes (Daley & Vere-Jones, 2007), starting from the type of data they are a natural fit for. Building on the basic theory, we will revisit several types of *canonical* point processes, such as Poisson processes, self-exciting point processes (*e.g.*, Hawkes processes (Hawkes, 1971)) or survival processes, each especially well fitted for a set of specific application domains. Then, we will present the notion of marks (or marked temporal point processes), a powerful mechanism which have been used to design complex models with random variables whose distribution depend on temporal point processes. Finally, we will introduce an alternative representation of marked temporal point processes as dynamical systems with jumps, which have been recently used to develop highly expressive models as well as sophisticated control algorithms.

1.2. Models (60 minutes)

We will elaborate on several specific examples of temporal point process models and efficient inference algorithms proposed in recent years, which leverage the concepts introduced in the first part of the tutorial as well as recent advances in machine learning.

More specifically, in terms of methodology, we will introduce the Dirichlet-Hawkes process and the hierarchical Dirichlet-Hawkes process (Du et al., 2015; Mavroforakis et al., 2017; Xu & Zha, 2017), recurrent temporal point processes (Dai et al., 2016; Mei & Eisner, 2017; Jing & Smola, 2017; Trivedi et al., 2017), deep generative models of temporal point processes (Xiao et al., 2017a; 2018), models based on dynamical systems and stochastic differential equations (SDEs) with jumps (De et al., 2016; Lee et al., 2016; Farajtabar et al., 2017a), and causal inference with temporal point processes (Li et al., 2017; Xu et al., 2016; Kusmierczyk & Gomez-Rodriguez, 2018).

In terms of applications, we will provide a wide variety of examples in social and information networks (Gomez-Rodriguez et al., 2011; Du et al., 2012; Iwata et al., 2013;

[1]Max Planck Institute for Software Systems, Kaiserslautern, Germany [2]Max Planck Institute for Intelligent Systems, Tübingen, Germany. Correspondence to: Manuel Gomez Rodriguez <manuelgr@mpi-sws.org>.

Proceedings of the 35^{th} International Conference on Machine Learning, Stockholm, Sweden, PMLR 80, 2018.

Du et al., 2015; De et al., 2016; Farajtabar et al., 2017a; Rizoiu et al., 2017; Mei & Eisner, 2017; Wang et al., 2017c;d), knowledge representation (Mavroforakis et al., 2017; Tabibian et al., 2017b; Trivedi et al., 2017), recommendation systems (Dai et al., 2016; Wang et al., 2016; Jing & Smola, 2017). human mobility (Jankowiak & Gomez-Rodriguez, 2017; Wang et al., 2017a), financial applications (Achab et al., 2017; Linderman & Adams, 2014), crime prediction (Egesdal et al., 2010; Mohler et al., 2013) health (Alaa et al., 2017; Rizoiu et al., 2018; Xiao et al., 2017b), and biology (Gupta et al., 2017).

1.3. Control (30 minutes)

We will first elaborate on the key innovation shared by most recent control algorithms leveraging temporal point processes—the control policy is an intensity function, often stochastic, which characterizes a temporal point process used as control signal. Then, we will present a variety of offline and online control algorithms based on, *e.g.*, convex optimization (Farajtabar et al., 2014; 2016), stochastic optimal control of dynamical systems with jumps (Karimi et al., 2016; Tabibian et al., 2017a; Kim et al., 2018; Zarezade et al., 2018; Wang et al., 2017b), and reinforcement learning (Farajtabar et al., 2017b). Finally, we will point out open problems and directions for future work.

2. Past tutorials on the topic

In recent years, we presented three closely related tutorials in top conferences in data mining, the Web, and AI:

(a) "Diffusion in Social and Information Networks: Problems, Models and Machine Learning Methods" at KDD '15. The tutorial slides are available at `http://learning.mpi-sws.org/kdd-2015-tutorial/`[1].

(b) "Diffusion in Social and Information Networks: Research Problems, Probabilistic Models and Machine Learning Methods" at WWW '15. The tutorial slides are available at `http://learning.mpi-sws.org/www-2015-tutorial/`.

(c) "Machine Learning for Dynamic Social Network Analysis" at IJCAI ' 17. The tutorial slides are available at `http://learning.mpi-sws.org/ijcai-2017-tutorial/`.

However, our tutorial will differ from the above in two key aspects, which we believe will raise the interest of the machine learning community at large in temporal point processes. First, it will highlight connections between recent models and methods based on temporal point process and other techniques in the machine learning literature, *e.g.*, Bayesian nonparametrics, deep neural networks, dynamical systems, stochastic optimal control, and reinforcement learning. Second, it will not be restricted to applications in social and information systems.

[1]The video recordings are available at `https://youtu.be/jCIDFbMjz3g?list=PLn0nrSd4xjjaNzvUtxHzU64xTz4Y_XNK9` (first part) and `https://youtu.be/TMnUp-87cpk?list=PLn0nrSd4xjjaNzvUtxHzU64xTz4Y_XNK9` (second part).

3. Goals

In our tutorial, we aim to provide an in-depth introduction to the theoretical framework of temporal point processes, which underpins a variety of recent machine learning models, inference methods and control algorithms. The participants will learn about the most recent trends as well as open problems and directions for future work.

This tutorial aims to introduce the machine learning community at large to temporal point processes, which have been proven to be very effective in an increasing range of applications but they are not yet very well-known in the community. Our tutorial relates to several areas within the ICML community such as, *e.g.*, probabilistic and graphical models, time series analysis, deep learning, optimal control, reinforcement learning, computational social science, and novel applications in machine learning.

4. Target Audience and Prerequisites

This tutorial is meant for a broad audience at ICML, including students and researchers specifically interested in temporal point process models and machine learning methods in the context of computational social science, social and information networks, and the Web. We expect the tutorial to be well attended, with at least 50-60 researchers

No specific knowledge will be required beyond basic probability and core machine learning methods; the tutorial is self-contained and most of the foundational concepts are introduced during the tutorial.

5. Tutors' bio and expertise

Manuel Gomez Rodriguez will present "Representation" (30 minutes) and "Control" (30 minutes) and Isabel Valera will present "Models" (60 minutes).

Manuel Gomez Rodriguez is a tenure-track independent research group leader at the Max Planck for Software Systems. Manuel develops machine learning and large-scale data mining methods for the analysis, modeling, and control of large real-world social and information systems. He is particularly interested in problems motivated by the Web and social media and has received several recognitions for his research, including an Outstanding Paper Award at NIPS´13 and Best Research Paper Honorable Mentions at KDD´10 and WWW´17. Manuel holds a PhD in Electrical Engineering from Stanford University and a BS in Electrical Engineering from Carlos III University in Madrid (Spain), and has been a Barrie de la Maza Fellow and a Caja Madrid Fellow.

Isabel Valera is TBD.

References

Achab, Massil, Bacry, Emmanuel, Gaïffas, Stéphane, Mastromatteo, Iacopo, and Muzy, Jean-Francois. Uncovering causality from multivariate hawkes integrated cumulants. In *34th International Conference on Machine Learning*, 2017.

Alaa, Ahmed M, Hu, Scott, and van der Schaar, Mihaela. Learning from clinical judgments: Semi-markov-modulated marked hawkes processes for risk prognosis. In *34th International Conference on Machine Learning*, 2017.

Dai, Hanjun, Wang, Yichen, Trivedi, Rakshit, and Song, Le. Recurrent coevolutionary feature embedding processes for recommendation. *arXiv preprint arXiv:1609.03675*, 2016.

Daley, Daryl J and Vere-Jones, David. *An introduction to the theory of point processes*. Springer Science & Business Media, 2007.

De, Abir, Valera, Isabel, Ganguly, Niloy, Bhattacharya, Sourangshu, and Rodriguez, Manuel Gomez. Learning and forecasting opinion dynamics in social networks. *Advances in Neural Information Processing Systems*, 2016.

Du, N., Song, L., Smola, A., and Yuan, M. Learning Networks of Heterogeneous Influence. In *NIPS*, 2012.

Du, Nan, Farajtabar, Mehrdad, Ahmed, Amr, Smola, Alexander J, and Song, Le. Dirichlet-hawkes processes with applications to clustering continuous-time document streams. In *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2015.

Egesdal, Mike, Fathauer, Chris, Louie, Kym, Neuman, Jeremy, Mohler, George, and Lewis, Erik. Statistical and stochastic modeling of gang rivalries in los angeles. *SIAM Undergraduate Research Online*, 3:72–94, 2010.

Farajtabar, M., Du, N., Gomez-Rodriguez, M., Valera, I., Zha, H., and Song, L. Shaping social activity by incentivizing users. In *NIPS*, 2014.

Farajtabar, M., Wang, Y., Gomez-Rodriguez, M., Li, S., Zha, H., and Song, L. Coevolve: A joint point process model for information diffusion and network evolution. *Journal of Machine Learning Research*, 2017a.

Farajtabar, Mehrdad, Ye, Xiaojing, Harati, Sahar, Song, Le, and Zha, Hongyuan. Multistage campaigning in social networks. In *Advances in Neural Information Processing Systems*, 2016.

Farajtabar, Mehrdad, Yang, Jiachen, Ye, Xiaojing, Xu, Huan, Trivedi, Rakshit, Khalil, Elias, Li, Shuang, Song, Le, and Zha, Hongyuan. Fake news mitigation via point process based intervention. *Advances in Neural Information Processing Systems*, 2017b.

Gomez-Rodriguez, M., Balduzzi, D., and Schölkopf, B. Uncovering the Temporal Dynamics of Diffusion Networks. In *28th International Conference on Machine Learning*, 2011.

Gupta, Amrita, Farajtabar, Mehrdad, Dilkina, Bistra, and Zha, Hongyuan. Hawkes processes for invasive species modeling and management. *arXiv preprint arXiv:1712.04386*, 2017.

Hawkes, Alan G. Spectra of some self-exciting and mutually exciting point processes. *Biometrika*, 58(1):83–90, 1971.

Iwata, Tomoharu, Shah, Amar, and Ghahramani, Zoubin. Discovering latent influence in online social activities via shared cascade poisson processes. In *KDD*, 2013.

Jankowiak, Martin and Gomez-Rodriguez, Manuel. Uncovering the spatiotemporal patterns of collective social activity. In *2017 SIAM International Conference on Data Mining*, 2017.

Jing, How and Smola, Alexander J. Neural survival recommender. In *Proceedings of the Tenth ACM International Conference on Web Search and Data Mining*, pp. 515–524, 2017.

Karimi, M., Tavakoli, E., Farajtabar, M., Song, L., and Gomez-Rodriguez, M. Smart Broadcasting: Do you want to be seen? In *22nd ACM SIGKDD International Conference on Knowledge Discovery in Data Mining*, 2016.

Kim, J., Tabibian, B., Oh, A., Schoelkopf, B., and Gomez-Rodriguez, M. Leveraging the crowd to detect and reduce the spread of fake news and misinformation. In *11th ACM International Conference on Web Search and Data Mining*, 2018.

Kusmierczyk, Tomasz and Gomez-Rodriguez, Manuel. On the causal effect of badges. In *Proceedings of the 27th International World Wide Web Conference*, 2018.

Lee, Young, Lim, Kar Wai, and Ong, Cheng Soon. Hawkes processes with stochastic excitations. In *International Conference on Machine Learning*, 2016.

Li, Shuang, Xie, Yao, Farajtabar, Mehrdad, Verma, Apurv, and Song, Le. Detecting changes in dynamic events over networks. *IEEE Transactions on Signal and Information Processing over Networks*, 3(2):346–359, 2017.

Linderman, S. and Adams, R. Discovering latent network structure in point process data. In *31st International Conference on Machine Learning*, 2014.

Mavroforakis, C., Valera, I., and Gomez-Rodriguez, M. Modeling the dynamics of online learning activity. In *Proceedings of the 26th International World Wide Web Conference*, 2017.

Mei, Hongyuan and Eisner, Jason M. The neural hawkes process: A neurally self-modulating multivariate point process. In *Advances in Neural Information Processing Systems*, 2017.

Mohler, George et al. Modeling and estimation of multi-source clustering in crime and security data. *The Annals of Applied Statistics*, 7(3):1525–1539, 2013.

Rizoiu, Marian-Andrei, Xie, Lexing, Sanner, Scott, Cebrian, Manuel, Yu, Honglin, and Van Hentenryck, Pascal. Expecting to be hip: Hawkes intensity processes for social media popularity. In *26th International Conference on World Wide Web*, 2017.

Rizoiu, Marian-Andrei, Mishra, Swapnil, Kong, Quyu, Carman, Mark, and Xie, Lexing. Sir-hawkes: on the relationship between epidemic models and hawkes point processes. In *Proceedings of the 27th International World Wide Web Conference*, 2018.

Tabibian, Behzad, Upadhyay, Utkarsh, De, Abir, Zarezade, Ali, Schoelkopf, Bernhard, and Gomez-Rodriguez, Manuel. Optimizing human learning. *arXiv preprint arXiv:1712.01856*, 2017a.

Tabibian, Behzad, Valera, Isabel, Farajtabar, Mehrdad, Song, Le, Schoelkopf, Bernhard, and Gomez-Rodriguez, Manuel. Distilling information reliability and source trustworthiness from digital traces. In *26th International World Wide Web Conference*, 2017b.

Trivedi, Rakshit, Dai, Hanjun, Wang, Yichen, and Song, Le. Know-evolve: Deep temporal reasoning for dynamic knowledge graphs. In *34th International Conference on Machine Learning*, 2017.

Wang, Pengfei, Fu, Yanjie, Liu, Guannan, Hu, Wenqing, and Aggarwal, Charu. Human mobility synchronization and trip purpose detection with mixture of hawkes processes. In *23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2017a.

Wang, Y., Grady, W., Theodorou, E., and Song, L. Variational policy for guiding point processes. In *34th International Conference on Machine Learning*, 2017b.

Wang, Y., Ye, X., Zha, H., and Song, L. Predicting user activity level in point process models with mass transport equation. In *Advances in Neural Information Processing Systems*, 2017c.

Wang, Yichen, Du, Nan, Trivedi, Rakshit, and Song, Le. Coevolutionary latent feature processes for continuous-time user-item interactions. In *Advances in Neural Information Processing Systems*, 2016.

Wang, Yichen, Ye, Xiaojing, Zhou, Haomin, Zha, Hongyuan, and Song, Le. Linking micro event history to macro prediction in point process models. In *Artificial Intelligence and Statistics*, 2017d.

Xiao, M., Farajtabar, M., Ye, X., Yan, J., Song, L., and Zha, H. Learning conditional generative models for temporal point processes. In *32nd AAAI Conference on Artificial Intelligence*, 2018.

Xiao, Shuai, Farajtabar, Mehrdad, Ye, Xiaojing, Yan, Junchi, Song, Le, and Zha, Hongyuan. Wasserstein learning of deep generative point process models. In *Advances in Neural Information Processing Systems*, 2017a.

Xiao, Shuai, Yan, Junchi, Farajtabar, Mehrdad, Song, Le, Yang, Xiaokang, and Zha, Hongyuan. Joint modeling of event sequence and time series with attentional twin recurrent neural networks. *arXiv preprint arXiv:1703.08524*, 2017b.

Xu, Hongteng and Zha, Hongyuan. A dirichlet mixture model of hawkes processes for event sequence clustering. In *Advances in Neural Information Processing Systems*, 2017.

Xu, Hongteng, Farajtabar, Mehrdad, and Zha, Hongyuan. Learning granger causality for hawkes processes. In *33rd International Conference on Machine Learning*, 2016.

Zarezade, A., De, A., Upadhyay, U., Rabiee, H., and Gomez-Rodriguez, M. Steering social activity: A stochastic optimal control point of view. *Journal of Machine Learning Research*, 2018.