Abstract

A recommender system aims to understand the users' inclination towards the different items and provide better experiences by recommending candidate items for future interactions. These personalized recommendations can be of various forms, such as e-commerce products, points-of-interest (POIs), music, social connections, *etc.* Traditional recommendation systems, such as content-based and collaborative filtering models, calculate the similarity between items and users and then recommending similar items to similar users. However, these approaches utilize the user-item interactions in a *static* way, *i.e.*, without any time-evolving features. This assumption significantly limits their applicability in real-world settings, as a notable fraction of data generated via human activities can be represented as a sequence of events over a continuous time. These continuous-time event sequences or CTES [1] are pervasive across a wide range of domains such as online purchases, health records, spatial mobility, social networks, *etc.* Moreover, these sequences can implicitly represent the time-sensitive properties of events, the evolving relationships between events, and the temporal patterns within and across sequences. For example, (i) event sequences derived from the purchases records in e-commerce platforms can help in monitoring the users' evolving preferences towards products; and (ii) sequences derived from spatial mobility of users can help in identifying the geographical preferences of users, their check-in category interests, and the physical activity of the population within the spatial region. Therefore, we represent the user-item interactions as temporal sequences of discrete events, as understanding these patterns is essential to power accurate recommender systems.

With the research directions described in this thesis, we seek to address the critical challenges in designing recommender systems that can understand the dynamics of continuous-time event sequences. We follow a *ground-up* approach, *i.e.*, first, we address the problems that may arise due to the poor quality of CTES data being fed into a recommender system. Later, we handle the task of designing accurate recommender systems. To improve the quality of the CTES data, we address a fundamental problem of overcoming *missing* events in temporal sequences. Moreover, to provide accurate sequence modeling frameworks, we design solutions for points-

[1]We use the acronym CTES to denote a single and well as multiple continuous-time event sequences.

of-interest recommendation, *i.e.*, models that can handle spatial mobility data of users to various POI check-ins and recommend candidate locations for the next check-in. Lastly, we highlight that the capabilities of the proposed models can have applications beyond recommender systems, and we extend their abilities to design solutions for large-scale CTES retrieval and human activity prediction.

To summarize, this thesis includes three directions: (i) Temporal Sequences with Missing Events; (ii) Recommendation in Spatio-Temporal Settings; and (iii) Applications of Modeling Temporal Sequences. In the first part, we present an unsupervised model and inference method for learning neural sequence models in the presence of CTES with missing events. This framework has many downstream applications, such as imputing missing events and forecasting future events. In the second part, we design point-of-interest recommender systems that utilize the geographical features associated with the spatial check-ins to recommend future locations to a user. Here, we propose solutions for *static* POI recommendation, sequential recommendations, and recommendations models that utilize the similarity between physical mobility and the smartphone activities of users. Lastly, in the third part, we highlight the strengths of the proposed frameworks to design solutions for two tasks: (i) retrieval systems, *i.e.*, retrieving relevance sequences for a given query CTES from a large corpus of CTES data; and (ii) understanding that different users take different times to perform similar actions in activity videos. Moreover, in each chapter, we highlight the drawbacks of current deep learning-based models, design better sequence modeling frameworks, and experimentally underline the efficacy of our proposed solutions over the state-of-the-art baselines. Lastly, we report the drawbacks and the possible extensions for every solution proposed in this thesis.

A significant part of this thesis uses the idea of modeling the underlying distribution of CTES via *neural* marked temporal point processes (MTPP). Traditional MTPP models are stochastic processes that utilize a fixed formulation to capture the generative mechanism of a sequence of discrete events localized in continuous time. In contrast, neural MTPP combine the underlying ideas from the point process literature with modern deep learning architectures. The ability of deep-learning models as accurate function approximators has led to a significant gain in the predictive prowess of neural MTPP models. In this thesis, we utilize and present several neural network-based enhancements for the current MTPP frameworks for the aforementioned real-world applications.