

Deep Learning for Urban Planning and Location-Based Services

Accurately attributing user visits to Points of Interest (POIs) is a cornerstone of human mobility analytics, aiding applications in personalized services, marketing and downstream geo-spatial tasks such as next-location prediction and anomaly detection [1]. *POI attribution* maps raw GPS trajectories to semantically meaningful places [7], adding interpretability—e.g., identifying a coffee shop visit at 3 pm is far more useful than recording coordinates $\langle \text{latitude}, \text{longitude} \rangle$ at time t . Yet attribution is difficult: GPS errors (2–20 meters) and dense urban clustering of POIs (often 50+ within 100 meters), render proximity-based heuristics unreliable. Accurate attribution, however, yields fine-grained behavioral insights (e.g., which store in a strip mall was visited), enabling more precise applications, from urban planning [6] to public health, such as identifying potential pandemic hotspots [2]. Conversely, misattributions risk contaminating downstream models, leading them to learn misleading or spurious patterns. Despite this complexity, attribution is often reduced to a simple heuristic: assigning each stay to the nearest POI [4]. While straightforward, this approach overlooks key real-world challenges, including GPS noise, dense urban settings where multiple POIs fall within error bounds, and contextual signals such as visit duration or time of day. More sophisticated methods [5] can improve accuracy by leveraging detailed spatial features like building footprints and hierarchical metadata, but such information is not universally available.

Instead, we propose **POIFormer**, a novel Transformer-based framework for POI attribution that jointly models a diverse set of signals, including spatial proximity, temporal features of the visit (arrival/departure and dwell time), POI semantics, user-specific mobility patterns, and population-level historical trends. A key innovation of **POIFormer** is its explicit incorporation of two dimensions of behavioral context: one capturing individual preferences, and another capturing crowd-level visit patterns. Individual preferences are modeled using a transformer that considers both past and future visits, with the location of the current (target) visit masked. This context enables the transformer to evaluate which nearby POI candidate is most likely given a user’s past and future visits, based on the time of day and duration of the stay of the target visit. Crowd-level historical visit patterns are modeled using the temporal popularity distributions of POIs, estimated via Kernel Density Estimation (KDE). These KDE models capture the joint distribution of location and time (e.g., hour of day) for visits within each POI category. This enables **POIFormer** to probabilistically downweight unlikely POIs—for example, reducing the likelihood of assigning a late-night visit to a coffee shop if historical data shows it is rarely visited at that hour. These KDEs are pre-computed per category facilitating efficient, scalable inference without sacrificing accuracy since they retain the full joint distribution of location and time while avoiding the need for computation at time of inference. Finally, **POIFormer** combines individual and crowd-level scores into a unified likelihood measure, selecting the most probable POI (or set of POI) among nearby candidates.

Furthermore, unlike prior approaches [3, 5], **POIFormer** makes no restrictive assumptions about POI categories, and does not rely on detailed spatial data layers about POIs, thereby enhancing its applicability across diverse geographic and data-constrained contexts. Extensive experimental evaluation on publicly available datasets, one simulated and one derived from real-world mobility traces, demonstrate that **POIFormer** consistently outperforms existing baselines including the current state-of-the-art technique proposed by SafeGraph [5] by a substantial margin, particularly in top-3 and top-5 accuracy.

- [1] Shang-Ling Hsu, Emmanuel Tung, John Krumm, Cyrus Shahabi, and Khuram Shafique. Trajgpt: Controlled synthetic trajectory generation using a multitask transformer-based spatiotemporal model. In *Proceedings of the 32nd ACM International Conference on Advances in Geographic Information Systems*, pages 362–371, 2024.
- [2] Qingchun Li, Liam Bessell, Xin Xiao, Chao Fan, Xinyu Gao, and Ali Mostafavi. Disparate patterns of movements and visits to points of interest located in urban hotspots across us metropolitan cities during covid-19. *Royal Society open science*, 8(1):201209, 2021.
- [3] Kyosuke Nishida, Hiroyuki Toda, Takeshi Kurashima, and Yoshihiko Suhara. Probabilistic identification of visited point-of-interest for personalized automatic check-in. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pages 631–642, 2014.
- [4] Achilleas Psyllidis, Song Gao, Yingjie Hu, Eun-Kyeong Kim, Grant McKenzie, Ross Purves, May Yuan, and Clio Andris. Points of interest (poi): A commentary on the state of the art, challenges, prospects for future. *Computational urban science*, 2(1):20, 2022.
- [5] SafeGraph. Determining points of interest visits from location data:technical guide to visit attribution. <https://www.safegraph.com/guides/visit-attribution-white-paper>, March 2025. Data for Accurate Geospatial Analytics.
- [6] Maria Despoina Siamou, Jialiang Li, John Krumm, Cyrus Shahabi, and Hua Lu. Poly2vec: Polymorphic fourier-based encoding of geospatial objects for geoai applications. In *Forty-second International Conference on Machine Learning*, 2025.
- [7] Yu Zheng and Xiaofang Zhou. *Computing with spatial trajectories*. Springer Science & Business Media, 2011.