Context-Aware Location Privacy-Preserving System for Multi-modal Transportation

Huyen-Trang Le[1], Mariana Cunha[2], João P. Vilela[2], Nazim Agoulmine[1]
[1]University of Paris-Saclay, [2]University of Porto

10TH JUNIOR CONFERENCE ON DATA SCIENCES AND ENGINEERING

Introduction and Motivation

- Location-based services continuously collect mobility data, often without user awareness, which raises major privacy concerns.
- Sensitive places such as homes and workplaces can be inferred from mobility patterns and misused by untrustworthy entities.
- Advances in Artificial Intelligence (AI) also increase the risks of predicting destinations and transport modes.
- Current Privacy-Preserving Mechanisms (PPMs) are rarely adapted to user contexts, especially in transportation scenarios.
- We propose a context-aware system that configures different Location PPMs (LPPMs) according to user-specific transport modes.

Methodology and Experimental Setup

A. Privacy Mechanism Pipeline

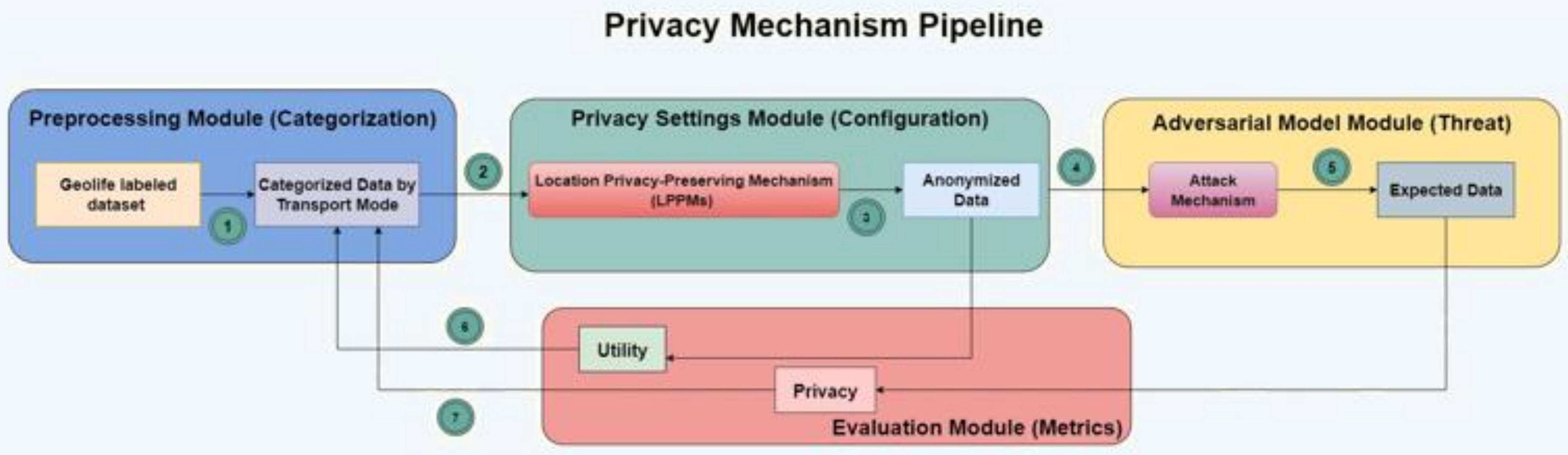


Fig. 1: Example of a typical pipeline for evaluating a privacy mechanism.

- **Preprocessing Module (Categorization)** first partitions Geolife dataset by transportation mode so that mechanism parameters can be adapted to user context and movement characteristics.
- **Privacy Settings Module (Configuration)** anonymizes the preprocessed traces using four baseline mechanisms: Planar Laplace Geo-Indistinguishability (Geo-Ind) [3], Adaptive Geo-Ind [4], Clustering Geo-Ind [5], and Velocity-Aware Geo-Ind (VA-GI) [6]. Each mechanism perturbs the original locations by adding calibrated noise.
- **Adversarial Model Module (Threat)** estimates the original traces based on the anonymized outputs to evaluate the resilience of PPMs against attacks. However, attack mechanisms were not executed in these initial experiments.
- **Evaluation Module (Metrics)** assesses the mechanisms performance using a set of privacy and utility metrics [1] — for example, inferred location accuracy, average quality loss, and task-specific utility measures — to quantify the privacy–utility trade-offs.

B. Configuration Variables and Implementation Notes per LPPM

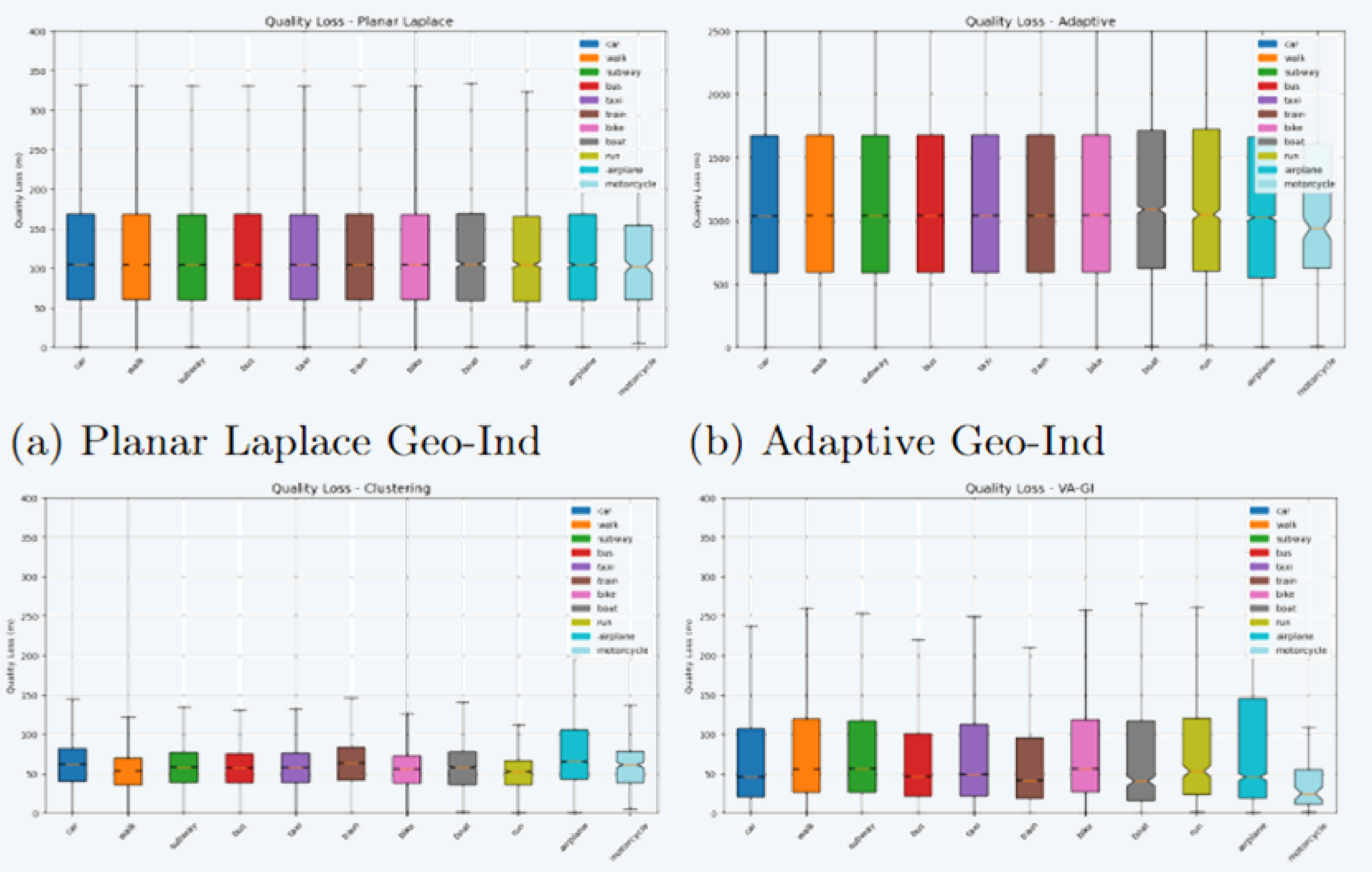
LPPM	Configuration Variables	Implementation Notes
Planar Laplace Geo-Ind	$\epsilon = 0.016$	Fixed ϵ.
Adaptive Geo-Ind	$ws = 2, \Delta_1 = 124.29$, $\Delta_2 = 428.56, \epsilon = 0.016$	ϵ adapted to the context.
Clustering Geo-Ind	$\epsilon = 0.016, r = \ln(4)/\epsilon$	Fixed ϵ, but LPPM adapted to the context.
VA-GI	$m = 10, \epsilon = 0.016$	ϵ adapted to the context.

C. Dataset and Privacy Toolkit

- These experiments resort to the **Geolife Dataset**, a dataset of **real-world trajectories labeled by transport mode**, that represents a solid basis for analyzing mobility patterns.
- The implementation and evaluation process is standardized by **Privkit** [2], an open-source privacy toolkit that enables testing and configuring PPMs, as well as evaluating mechanisms in terms of privacy and utility.

Preliminary Results

- **Planar Laplace Geo-Ind** achieves an average quality loss of approximately 120 m, consistent across different transportation modes.
- **Clustering Geo-Ind** reduces distortion to around 90 m; particularly effective for low and medium velocities.
- **Adaptive Geo-Ind** exhibits the poorest performance, with quality loss reaching up to 1500 m.
- **VA-GI** provides the best performance, with quality loss often below 50 m and consistent across all modes.



(a) Planar Laplace Geo-Ind (b) Adaptive Geo-Ind

(c) Clustering Geo-Ind (d) VA-GI

Fig. 2: Comparison of the quality loss of LPPMs based on the transport mode.

Conclusion

- In the experiments, Clustering Geo-Ind and VA-GI are the LPPMs that achieve the best privacy-utility trade-off.
- These findings indicate that adapting privacy settings to user context significantly improves outcomes.
- Our future work will mainly focus on integrating adversarial models and exploring predictive AI-based re-identification attacks, and corresponding defenses.

References

[1] Mendes, R., Cunha, M., & Vilela, J. P. (2020). Impact of frequency of location reports on the privacy level of geo-indistinguishability. *Proceedings on Privacy Enhancing Technologies Symposium*, 2020(2), 379–396.

[2] Cunha, M., et al. (2024). Privkit: A toolkit of privacy-preserving mechanisms for heterogeneous data types. In *Proceedings of the 14th ACM CODASPY* (pp. 319–324).

[3] Andrés, M., Bordenabe, N., Chatzikokolakis, K., & Palamidessi, C. (2013). Geo-indistinguishability: Differential privacy for location-based systems. In *Proceedings of the ACM SIGSAC Conference on CCS* (pp. 901–914). ACM.

[4] Al-Dhubhani, R., & Cazalas, J. M. (2018). An adaptive geo-indistinguishability mechanism for continuous LBS queries. *Wireless Networks*, 24(8), 3221–3239.

[5] Cunha, M., Mendes, R., & Vilela, J. P. (2019). Clustering geo-indistinguishability for privacy of continuous location traces. In *4th International Conference on Computing, Communications and Security (ICCCS)* (pp. 1–8). IEEE.

[6] Mendes, R., Cunha, M., & Vilela, J. P. (2023). Velocity-aware geo-indistinguishability. In *Proceedings of the 13th ACM CODASPY* (pp. 141–152).