Evaluating Privacy-Utility Trade-offs in Differentially Private Mobility Data Analysis

Syed Ibrahim Khalil

Context and Why This Matters

The Problem Statement

"How do we publish useful mobility statistics while protecting individuals?"

- Evaluated across **three distinct datasets** exhibiting fundamentally different user behavior concentration.
- **One key parameter in focus:** M-clipping (the absolute maximum trips permitted per user).

Because datasets differ structurally, some have uniformly balanced users, while others are dominated by heavy contributors. The exact same privacy setup behaves radically differently across contexts.

Worst-Case Math: Sensitivity and Noise

- **User-level adjacency:** DP mathematically protects the shift of *one user in* vs. *one user out*.
- In spatial analyses, each trip contributes to an origin and a destination tile.

SPATIAL SENSITIVITY

$$\Delta f = 2M$$

LAPLACE SCALE

$$b = \frac{\Delta f}{\varepsilon} = \frac{2M}{\varepsilon}$$

Concrete Example (GeoLife):

Max trips per user: **2,153**

Worst-case spatial impact *without clipping*: up to **4,306 records** (2 × 2,153)

With M=10: strictly bounded to 20 records

Conclusion: If you increase **M**, the required noise increases linearly.

Worst-Case User Removal Scenario

Differential Privacy noise calibration is driven entirely by the absolute worst-case possibility.

1. Remove one regular, low-activity user: **small output change.**
2. Remove one power user: **very large output change.**
3. DP noise **must be calibrated to case 2.**

M-clipping explicitly limits how much any single person can move the published statistics.

Dataset User-Contribution Profiles (Output-Backed)

Dataset	Records	Users	High-Activity Users	High-Activity Share	Threshold	Max Trips/User
Berlin	2,834,268	378,759	62,496	**30.5%**	9	16
Madrid	443,481	75,207	8,410	**23.3%**	9	16
GeoLife	37,340	182	19	**62.2%**	400	2,153

Concentration Ratios (High-Activity Users / Total Users)

- Berlin: **16.5%**
- Madrid: **11.2%**
- GeoLife: **10.4%**

Why "Dense" vs "Sparse" Is Not Raw Row Count

Dense/sparse refers strictly to the **signal strength in analysis cells**, not the total rows alone.

Berlin

Has significantly more overall rows, but features many weak/sparse cells for key spatial tasks.

Madrid

Active locations show highly concentrated counts. Relative DP noise impact is extremely low.

Measured M-Clipping Impact: Berlin

Representative high-activity locations (Mean Absolute Error):

Location ID	M=10, ε=0.5	M=10, ε=1.0	M=15, ε=0.5	M=15, ε=1.0
110100330	15.67	14.53	**19.93**	20.40
110100410	13.67	10.33	**43.27**	26.60
110101210	16.40	6.20	**81.93**	4.07
110109820	16.07	6.40	**36.13**	10.93
110110310	23.60	10.87	**30.67**	18.80

Error rises strongly when M increases in sparse contexts.

Measured M-Clipping Impact: Madrid

Representative high-density locations (Mean Absolute Error):

Location ID	ε=1.0, M=10	ε=1.0, M=20	ε=1.0, M=30	ε=2.0, M=10	ε=2.0, M=20
005-003	738.47	738.33	**742.53**	729.00	718.40
005-020	1331.20	1331.80	**1336.80**	1321.70	1322.90
005-025	774.07	766.27	**761.00**	775.33	773.27
006-006A	702.27	719.87	**722.47**	708.47	723.13

Changing M from 10 to 30 often changes values only slightly for high-density locations.

GeoLife as the Power-User Stress Test

- **182** users total
- **19** high-activity users
- **62.2%** of all trips from these 19 users
- Top trip counts: **2,153, 2,024, 809**

> This is exactly the case where user-level DP without clipping becomes unstable for publication utility, because worst-case influence is too high.

M-Clipping Preprocessing Effect (Run-Time)

At **M=10** (from run logs):

Berlin

1,414,920

Unique trips selected

Dataframe shape: 2,829,840

Madrid

222,383

Unique trips selected

Dataframe shape: 442,805

GeoLife

1,559

Unique trips selected

Dataframe shape: 3,118

What Happens If a Power User Is Removed?

Scenario explanation (user-level adjacency):

1. Compare two adjacent datasets: one includes the user, one excludes the same user
2. Query delta can be very large without clipping
3. DP must protect this maximum delta
4. Therefore noise calibration is driven by this worst case

Publication Implication: Why This Matters

For publication pipelines:

- M is a publication design control, not just a training hyperparameter.
- **Too high M:** stronger worst-case influence, more noise, weaker utility.
- **Too low M:** lower noise but possible representation bias (over-truncation).

Operational recommendation:

- Publish with documented tuple: *(privacy unit, M, eps, delta, query type)*
- Include concentration diagnostics (high-activity user share).
- Report expected utility uncertainty by analysis category.

Practical M-Selection Protocol

1. Compute user contribution distribution first
2. Identify concentration risk (share of trips from high-activity users)
3. Sweep M values on representative analyses
4. Choose smallest M that preserves required utility
5. Lock M and publish with full metadata

Mechanism Choice vs Contribution Choice

Per-query mechanism difference (Laplace vs Gaussian): empirically smaller differences

Contribution bounding effect (M choice): large and context-dependent

Mechanism selection is secondary when user concentration risk is high.

Thank you.