

Exposé

Vorhersage von menschlichen Mobilitätsmustern über mehrere
Städte hinweg
(GISCU^P 2025)

Vorgelegt von: 230646 Ivan Pushkarev, `ivan.pushkarev@tu-dortmund.de`

Betreuer: Mart Hagedoorn, M.Sc. `mart.hagedoorn@tu-dortmund.de`

Studiengang: B. Sc. Informatik

Fakultät für Informatik, TU Dortmund

1 Fragestellung

Die Vorhersage menschlicher Mobilitätsmuster ist ein zentrales Forschungsgebiet mit vielfältigen Anwendungen in der Stadtplanung, im Transportwesen, im Katastrophenmanagement und in der Epidemiemodellierung. Durch die Verfügbarkeit großer Datensätze zur menschlichen Bewegung, die von Mobilgeräten und sozialen Medienplattformen gesammelt werden, konnten komplexe Modelle zur menschlichen Mobilitätsvorhersage entwickelt und getestet werden.

Der Human Mobility Prediction Challenge 2025 (GISCUP 2025) fokussiert sich auf die Vorhersage von Bewegungsmustern in vier japanischen Metropolregionen (Städte A, B, C und D) über einen Zeitraum von 75 Tagen.

Die zentrale Fragestellung des Wettbewerbs lautet: “Können wir die Vorhersagbarkeit menschlicher Mobilität verbessern, indem wir Daten aus verschiedenen Städten nutzen?” Diese Frage ist besonders relevant, da städteübergreifende Transferierbarkeit von Mobilitätsmodellen ein wichtiges und bisher unzureichend erforschtes Gebiet darstellt.

Bei dieser Aufgabe stehen verschiedene Herausforderungen im Vordergrund:

1. Die Datensätze haben unterschiedliche Größen (Stadt A: 100.000 Personen, Stadt B: 25.000, Stadt C: 20.000, Stadt D: 6.000), was die Frage aufwirft, ob datenreiche Städte zur Verbesserung der Vorhersagen in datenarmen Städten beitragen können.
2. Die räumliche Struktur jeder Stadt ist einzigartig, was die direkte Übertragbarkeit von Modellen erschwert.
3. Für die Städte C und D liegen die vorherzusagenden Tage 61-75 zeitlich mehrere Monate nach den Tagen 1-60, was eine zusätzliche zeitliche Herausforderung darstellt.

Hauptfrage:

- Wie lässt sich die Vorhersagbarkeit menschlicher Mobilitätsmuster durch die Nutzung von Daten aus mehreren Städten verbessern?

Teilfragen:

- Welche Mobilitätsmuster sind stadtspezifisch und welche sind stadtübergreifend beobachtbar?
- Wie kann Transfer Learning eingesetzt werden, um Wissen von datenreichen Städten auf datenarme Städte zu übertragen?
- Welche Methoden eignen sich besonders für die langfristige Vorhersage in Szenarien mit zeitlichen Lücken (wie bei Städten C und D)?
- Inwieweit verbessert die Integration von POI-Daten (Points of Interest) die Qualität der Vorhersage?

2 Theoretischer Rahmen

Die Arbeit stützt sich auf verschiedene theoretische Konzepte und Forschungsgebiete:

2.1 Menschliche Mobilitätsmodelle

Die Forschung zur menschlichen Mobilität hat verschiedene Modelle hervorgebracht, von einfachen Random Walks bis hin zu komplexen State-Space-Modellen. Grundlegende Arbeiten wie die von Brockmann et al. (2006) und González et al. (2008) haben gezeigt, dass menschliche Bewegungsmuster trotz ihrer scheinbaren Komplexität skalenfreien Lévy-Flügen mit charakteristischen Verteilungen folgen.

2.2 Maschinelles Lernen für Zeitreihendaten

Die Mobilitätsvorhersage lässt sich als Zeitreihenprognose darstellen. Moderne Deep-Learning-Ansätze wie Transformer-Architekturen (Vaswani et al., 2017) und große Sprachmodelle (LLMs) haben sich als vielversprechend für diese Aufgabe erwiesen, wie die Arbeiten von Tang et al. (2024) und He et al. (2024) zeigen.

2.3 Transfer Learning und Domain Adaptation

Zur Übertragung von Wissen zwischen verschiedenen Städten sind Techniken des Transfer Learnings und der Domain Adaptation relevant. Diese ermöglichen es, Modelle auf datenreichen Quellen zu trainieren und anschließend auf datenarme Zieldomänen anzupassen (Pan & Yang, 2010).

3 Methodische Herangehensweise

3.1 Retrieval-basierte Methode

Als Hauptansatz wird eine retrieval-basierte Methode verfolgt. Dabei werden für einen neuen Tag und Nutzer ähnliche beobachtete Tage und Nutzerprofile aus dem Trainingsdatensatz identifiziert und deren Trajektorien als Basis für die Vorhersage genutzt. Dieser Ansatz hat den Vorteil, dass er weniger anfällig gegenüber exakten Verhaltensprognosen ist und stattdessen auf Ähnlichkeitsmustern basiert, die städteübergreifend gelten können.

Die Implementierung umfasst:

- Entwicklung von Random-Walk-Modellen, die POI-Informationen als Gewichtungsfaktoren für Übergangswahrscheinlichkeiten zwischen Gitterzellen nutzen
- Definition geeigneter Ähnlichkeitsmetriken für den Vergleich von Trajektorien
- Implementierung effizienter Suchalgorithmen zur Identifikation ähnlicher Muster

3.2 Evaluierung und Vergleich

Die entwickelten Methoden werden anhand der vom GISCUP vorgegebenen Metriken evaluiert:

- GEO-BLEU: Ein Ähnlichkeitsmaß für räumliche Sequenzen
- Dynamic Time Warping (DTW): Ein Maß für die Ähnlichkeit von Zeitreihen

Als Baseline-Modelle dienen die im Challenge-Dokument beschriebenen Ansätze (Global Mean, Global Mode, Per-User Mean, Per-User Mode, Unigram und Bigram Modelle).

4 Daten und Materialien

Die Arbeit verwendet den offiziellen GISCUP 2025 Datensatz, der menschliche Mobilitätsmuster in vier japanischen Städten enthält:

- Stadt A: 100.000 Individuen
- Stadt B: 25.000 Individuen
- Stadt C: 20.000 Individuen
- Stadt D: 6.000 Individuen

Jede Stadt ist in ein 200×200 Raster unterteilt, wobei jede Zelle einem Gebiet von 500×500 Metern entspricht. Die Mobilitätsdaten decken einen Zeitraum von 75 Tagen ab und sind in 30-Minuten-Intervalle unterteilt.

Die Daten enthalten zusätzlich POI-Informationen (Points of Interest) für jede Gitterzelle als 85-dimensionalen Vektor, die als ergänzende Informationen genutzt werden können.

5 Erwartete Ergebnisse

Die Arbeit soll folgende Ergebnisse liefern:

1. Eine detaillierte Analyse der stadtübergreifenden Gemeinsamkeiten und Unterschiede in Mobilitätsmustern
2. Ein effektives Modell zur Vorhersage menschlicher Bewegung, das die GEO-BLEU- und DTW-Metriken der Baseline-Modelle übertrifft
3. Erkenntnisse über die Transferierbarkeit von Mobilitätswissen zwischen verschiedenen urbanen Umgebungen

4. Empfehlungen für zukünftige Ansätze im Bereich der stadtübergreifenden Mobilitätsvorhersage

Die erwarteten Erkenntnisse sind nicht nur für den GISCUP 2025 relevant, sondern bieten auch wertvolle Einsichten für praktische Anwendungen in der Stadtplanung, im Katastrophenmanagement und in der Epidemiemodellierung.

6 Zeitplan

Zeitraum	Phase	Aktivitäten
Woche 1-2	Literaturrecherche	Einarbeitung in die Fachliteratur zur Mobilitätsvorhersage und Transfer Learning
Woche 3-4	Datenanalyse	Exploration und Visualisierung der Datensätze, Identifikation wichtiger Muster
Woche 5-10	Modellentwicklung	Implementierung der retrieval-basierten Methode
Woche 11-12	Evaluation	Vergleich der entwickelten Modelle mit Baselines anhand der GISCUP-Metriken
Woche 13-14	Dokumentation	Zusammenfassung der Ergebnisse, Abschluss der Bachelorarbeit

7 Literaturverzeichnis

Brockmann, D., Hufnagel, L., & Geisel, T. (2006). The scaling laws of human travel. *Nature*, 439(7075), 462-465.

González, M. C., Hidalgo, C. A., & Barabási, A. L. (2008). Understanding individual human mobility patterns. *Nature*, 453(7196), 779-782.

Pan, S. J., & Yang, Q. (2010). A survey on transfer learning. *IEEE Transactions on knowledge and data engineering*, 22(10), 1345-1359.

Wikipedia. (2023). Matrix completion. Retrieved from https://en.wikipedia.org/wiki/Matrix_completion

Buchin, K., Hagedoorn, M., & Korn, A. (2024). Discretized Random Walk Models for Efficient Movement Interpolation. In *The 32nd ACM International Conference on Advances in Geographic Information Systems*.

He, H., Luo, H., & Wang, Q. R. (2024). ST-MoE-BERT: A Spatial-Temporal Mixture-of-Experts Framework for Long-Term Cross-City Mobility Prediction. In *2nd ACM SIG-SPATIAL International Workshop on the Human Mobility Prediction Challenge*.

Tang, P., Yang, C., Xing, T., Xu, X., Jiang, R., & Sezaki, K. (2024). Instruction-Tuning Llama-3-8B Excels in City-Scale Mobility Prediction. In *2nd ACM SIGSPATIAL International Workshop on the Human Mobility Prediction Challenge*.

Terashima, H., Takagi, S., Tamura, N., Shoji, K., Hossain, T., Katayama, S., Urano, K., Yonezawa, T., & Kawaguchi, N. (2024). Time-series Stay Frequency for Multi-City Next Location Prediction using Multiple BERTs. In *2nd ACM SIGSPATIAL International Workshop on the Human Mobility Prediction Challenge*.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). Attention is all you need. *Advances in neural information processing systems*, 30.