
Automatische Verkehrsmittelerkennung aus Sensordaten

Birgit KOHLA

Institut für Straßen- und Verkehrswesen, Technische Universität Graz · birgit.kohla@tugraz.at

Zusammenfassung

Ein Einsatz mobiler Erhebungsgeräte, wie Smartphones und GPS-Logger zur automatisierten Aufzeichnung des Mobilitätsverhaltens verspricht viele Vorteile gegenüber Befragungsmethoden. Das folgende Paper beschäftigt sich mit der automatisierten Erkennung von 8 Verkehrsmittelkategorien aus charakteristischen Bewegungsmustern von Personen unter Nutzung aufgezeichneter Daten tragbarer GPS-Erhebungsgeräte mit eingebauten 3D-Beschleunigungssensoren. Es wird ein Algorithmus zur voll automatisierten Verkehrsmittelerkennung inklusive einer verlässlichen Detektion von Verkehrsmittelwechsellpunkten zur Etappenabgrenzung aus diesen Sensordaten vorgestellt. Bei einer Testanwendung des entwickelten Verfahrens konnte, allein durch Verarbeitung von Beschleunigungs- und GPS-Daten, eine richtige Gesamttrefferquote von 77 % und mehr erreicht werden.

1 Einleitung und Problemstellung

Qualitativ hochwertige Daten zum Mobilitätsverhalten von Personen stellen eine wichtige Grundlage in Verkehrs- und Raumplanung sowie für die Forschung dar. Traditionell werden zur Erhebung des täglichen Mobilitätsverhaltens überwiegend Befragungsmethoden (Fragebogen oder Interview) eingesetzt. Diese Methoden sind aufwendig für die Zielpersonen und es kommt leicht zu inhaltlichen Qualitätseinbußen bei den Ergebnisdaten (vgl. STOPHER et al. 2008). Ein Einsatz mobiler Erhebungsgeräte, wie Smartphones und GPS-Logger zur automatisierten Aufzeichnung des Mobilitätsverhaltens verspricht viele Vorteile gegenüber Befragungsmethoden. Diese Methode wird daher in der Mobilitätsforschung als zukünftige Alternative angesehen. Je nach eingesetztem Erhebungsgerät können Daten unterschiedlicher Sensoren genutzt werden, um das Mobilitätsverhalten von Personen über einen bestimmten Zeitraum zu rekonstruieren.

Eine bisher nicht vollständig gelöste Fragestellung bei technologiegestützten Mobilitätserhebungen ist die automatisierte Identifizierung und Zuordnung genutzter Verkehrsmittel zu einzelnen Wegen bzw. Wegetappen. Es finden sich dazu bereits einige Ansätze in nationaler und internationaler Forschung, die überwiegend auf der Auswertung von Geschwindigkeitsprofilen aus Positionsdaten (GPS) und teilweise auf dem Abgleich von Positionsdaten mit Geodaten des Verkehrsangebots basieren (RUDLOFF & RAY 2010, CHEN et al. 2011, RIESER-SCHÜSSLER et al. 2011, STOPHER et al. 2011). Die Nutzung von Daten aus dreidimensionalen Beschleunigungssensoren wird auch international erst seit Kurzem in wenigen Arbeiten erforscht (TROPED et al. 2008, OLIVER et al. 2010, REDDY et al. 2010, FENG et al. 2011, RIESER-SCHÜSSLER et al. 2011, WIDHALM et al. 2012, YU et al. 2013). Je nach Be-

einträchtigung des Satellitenempfangs kommt es bei der GPS-Positionierung immer wieder zu Datenausfällen. 3D-Beschleunigungssensoren sind unabhängig von Satellitenempfang und liefern daher vollständigere und genauere Daten als die GPS-Technologie. Daher ist diese Datenquelle ein Schwerpunkt dieser Forschungsarbeit.

Das folgende Paper beschäftigt sich mit der automatisierten Erkennung von acht Verkehrsmittelkategorien aus charakteristischen Bewegungsmustern von Personen unter Nutzung aufgezeichneter Daten tragbarer GPS-Erhebungsgeräte mit eingebauten 3D-Beschleunigungssensoren. Es wird ein Algorithmus zur voll automatisierten Verkehrsmittelerkennung inklusive einer verlässlichen Detektion von Verkehrsmittelwechsellpunkten zur Etappenabgrenzung in erhobenen Zeitreihen von GPS- und Beschleunigungsdaten vorgestellt.

2 Sensordaten und Hypothesen

Für die statistische Datenanalyse und die Entwicklung eines Algorithmus zur Verkehrsmittelerkennung stehen empirisch erhobene Referenzdaten aus dem Forschungsprojekt MOBIFIT (HERRY et al. 2011) zur Verfügung. Darin wurden 134 freiwillige Probanden mit passiven Datenloggern ausgestattet, welche GPS-Daten (Zeitreihen mit einer Aufzeichnung pro Sekunde) und Beschleunigungsdaten in drei Achsen (Zeitreihen mit zehn Aufzeichnungen pro Sekunde) aufzeichneten. Personenspezifische Merkmale wurden zusätzlich erhoben und das Mobilitätsverhalten der Probanden in einem persönlichen Interview mithilfe einer kartographischen Darstellung der GPS-Daten rekonstruiert. Für manche Verkehrsmittel mussten für die vorliegende Arbeit zusätzlich gezielte Ergänzungserhebungen durchgeführt werden, um eine ausreichende Datengrundlage zu erhalten. Im Zuge der Datenaufbereitung werden ungenaue und unplausible Daten gefiltert, sodass schließlich 1.713 Wege (das entspricht 327 Stunden Datenmaterial) für die weiteren Schritte zur Verfügung stehen.

Eine statistische Analyse des Zusammenhangs zwischen Sensordaten (aus 3D-Beschleunigungssensoren und GPS Positionierung) bzw. personenbezogenen Daten und gewählter Verkehrsmittelkategorie erfolgt mittels multinomialer logistischer Regressionsmodelle. Für die Modellformulierung werden insgesamt 230 potenzielle Einflussvariable aus den verfügbaren Sensordaten und personenspezifischen Merkmalen abgeleitet und getestet. Folgende Merkmalsgruppen wurden untersucht: Geschwindigkeit in Bewegungsrichtung, Beschleunigung in Bewegungsrichtung, resultierende 3D-Beschleunigung, Änderung der Bewegungsrichtung (Azimut), Umwegfaktor, verkehrsbedingte Halte, Qualität der GPS-Positionierung und Verkehrsmitteloptionen der Nutzer.

3 Modellkalibrierung

Für den Algorithmus zur Verkehrsmittelerkennung werden zwei multimodale logistische Regressionsmodelle zur Ermittlung einer Zugehörigkeitswahrscheinlichkeit für jede Verkehrsmittelkategorie kalibriert. Anhand der Zugehörigkeitswahrscheinlichkeiten erfolgt die Zuordnung von Bewegungsabschnitten zu einer der untersuchten Verkehrsmittelkategorien. Ein Modell wird für die Identifikation von Fußgängern, eines zur Unterscheidung aller acht Verkehrsmittelkategorien optimiert. Die Variablenauswahl erfolgt für beide Modelle schrittweise basierend auf statistischen Qualitätsmerkmalen der getesteten Variablen.

Für das Fußgängermodell werden nach dieser Vorgangsweise drei Variablen ausgewählt, davon zwei Variablen abgeleitet aus Beschleunigungsdaten und eine Variable abgeleitet aus GPS Positionen. Bei Zuordnung zu jener Kategorie mit der maximalen ermittelten Zugehörigkeitswahrscheinlichkeit wird mit diesem Modell für die Modellkalibrierung eine Trefferquote von 96 % erreicht. Der Likelihood-Ratio-Test zeigt mit einem Signifikanzniveau von 0,000 eine gute Modellanpassung. Aus dem Pseudo-R² nach Nagelkerke kann geschlossen werden, dass 88 % der Varianz durch die drei ausgewählten Variablen erklärbar sind.

Für das 8-Verkehrsmittel-Modell werden sieben Variable ausgewählt, davon vier abgeleitet aus Beschleunigungsdaten und drei abgeleitet aus GPS Positionen. Bei Zuordnung zu jener Kategorie mit der maximalen ermittelten Zugehörigkeitswahrscheinlichkeit wird mit diesem Modell für die Modellkalibrierung eine Trefferquote von 85 % erreicht. Der Likelihood-Ratio-Test zeigt auch bei diesem Modell mit einem Signifikanzniveau von 0,000 eine gute Modellanpassung. Aus dem Pseudo-R² nach Nagelkerke kann geschlossen werden, dass 97 % der Varianz durch die ausgewählten Variablen erklärbar sind.

4 Algorithmus zur Verkehrsmittelerkennung

Die beiden Modelle werden in einen Algorithmus implementiert, mit dem Ziel aus aufgezeichneten Datenreihen Wegetappen abzugrenzen und diese einer der folgenden acht Verkehrsmittelkategorien zuzuordnen: zu Fuß, Fahrrad, einspuriger motorisierter Individualverkehr, mehrspuriger motorisierter Individualverkehr, Linienbus, Straßenbahn, U-Bahn, Eisenbahn.

Um Mess- und Speicherfehler zu beheben, werden die aufgezeichneten Sensordaten anhand verschiedener Kriterien vorgefiltert. Anschließend erfolgt eine Unterteilung der Datenreihen in Abschnitte mit und ohne Bewegung. Die Unterteilung in Wegetappen basiert auf der Überlegung, dass mit jedem Verkehrsmittelwechsel eine (zumindest kurze) Etappe zu Fuß verbunden ist. Daher kommt ein zweistufiges Verfahren zur Anwendung (siehe Abbildung 1). In der ersten Stufe wird ein gleitendes Fenster von 20 s zur Unterteilung der Datenreihen und Ermittlung der Variablen für das Fußgängermodell eingesetzt. Anhand der ermittelten Zugehörigkeitswahrscheinlichkeiten aus dem Fußgängermodell erfolgt eine Unterteilung der Datenreihen in Fußgänger und „Nicht-Fußgänger“ und damit eine erste Abgrenzung von Verkehrsmittletappen. Diese abgegrenzten Etappen werden mithilfe des 8-Verkehrsmittel-Modells einer der acht Verkehrsmittelkategorien zugeordnet. Nach der Klassifizierung werden die Etappengrenzen erneut überprüft und aufeinander folgende Etappen derselben Verkehrsmittelkategorie aggregiert.

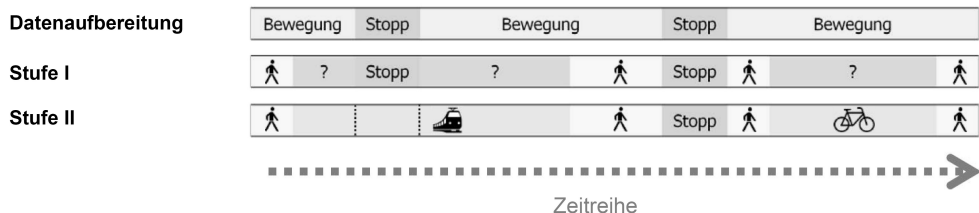


Abb. 1: Schematische Darstellung des Verfahrens zur Verkehrsmittelerkennung

5 Ergebnisse der Testanwendung

Aufgrund der geringen Fallzahlen in einigen Verkehrsmittelkategorien kann der verwendete Datensatz für Modellkalibrierung und Testanwendung nicht geteilt werden und es kommt derselbe Datensatz zum Einsatz. Allerdings wird für die Testanwendung davon ausgegangen, dass Beginn und Ende von Wegetappen unbekannt sind, also ausschließlich Messreihen der Sensoren vorliegen. Das muss vor allem bei einer Beurteilung der Ergebnisse der Testanwendung berücksichtigt werden.

Mit dem entwickelten zwei-stufigen Algorithmus konnte für den gesamten Datensatz bei der Verkehrsmittelkategorisierung eine Gesamttrefferquote von 77 % erreicht werden (siehe Tabelle 1). Die Verkehrsmittel Motorrad und Moped, Eisenbahn, Fahrrad und zu Fuß erreichten vergleichbare Ergebnisse zur Modellkalibrierung, während insbesondere städtische öffentliche Verkehrsmittel schlechtere Ergebnisse brachten. Das Filtern von Daten im Bereich von berichteten Verkehrsmittelwechseln brachte eine weitere Verbesserung der Ergebnisse, was auf Ungenauigkeiten bei der Abgrenzung der Wegetappen in den Referenzdaten schließen lässt.

Tabelle 1: Klassifizierungsergebnisse der Testanwendung für acht Verkehrsmittelkategorien, n = 1.166.221 Sekunden

Verkehrsmittel		Geschätzt									% richtig
Beobachtet	Keine Daten	F	R	M	P	U	S	B	E		
zu Fuß (F)	4.343	295.532	4.902	0	7.986	3.190	27.343	8.237	6.125	84%	
Fahrrad (R)	865	8.501	68.709	779	0	0	0	1.020	180	87%	
MIV einspurig (M)	0	0	540	7.153	0	0	0	0	0	93%	
MIV mehrspurig (P)	6.420	20.130	3.994	19.448	426.177	9.784	41.377	46.307	12.898	73%	
ÖV U-Bahn (U)	334	1.482	0	0	0	7.615	0	0	2.078	68%	
ÖV Straßenbahn (S)	5.254	11.089	1.323	0	1.221	3.232	27.180	1.821	5	59%	
ÖV Bus (B)	245	5.826	1.560	0	85	597	2.700	15.487	2.804	53%	
ÖV Eisenbahn (E)	1.745	675	0	169	1.138	873	1.200	0	36.543	90%	
% richtig		86%	85%	26%	98%	30%	27%	21%	60%	77%	

6 Fazit und Ausblick

Eine detaillierte Beschreibung der entwickelten Modelle, sowie des Algorithmus zur Etapenabgrenzung und Verkehrsmittelidentifikation findet sich bei KOHLA 2013. Es zeigte sich, dass Daten von 3D-Beschleunigungssensoren neben Positionsdaten sehr gut zur Erklärung des genutzten Verkehrsmittels geeignet sind. Variable aus personenbezogenen Merkmalen brachten für beide Modelle keine zusätzliche Verbesserung der Ergebnisse.

Unter Berücksichtigung der vergleichsweise großen Anzahl an zu unterscheidenden Verkehrsmittelkategorien (acht plus Stopp), und der geringen Anzahl an erklärenden Variablen (insgesamt neun) aus nur zwei Datenquellen (3D-Beschleunigung und GPS), liefert das entwickelte Verfahren gute Ergebnisse. Ein Vergleich mit Ergebnissen anderer Autoren ist aufgrund der unterschiedlichen Datengrundlagen und methodischen Vorgangsweisen kaum möglich. Ein Vergleich statistischer Methoden zur Verkehrsmittelklassifizierung wurde von

FENG & TIMMERMANS 2010 durchgeführt. Mit der Anwendung von Bayes'schen Netzen wurden dabei die besten Ergebnisse erzielt. Diese und andere vielversprechende Methoden sollten für die Weiterentwicklung des Verfahrens geprüft werden.

Durch Optimierung einzelner Verfahrensschritte, sowie durch Nutzung zusätzlicher Datenquellen lässt sich ebenfalls eine Verbesserung des entwickelten Verfahrens erwarten. In modernen Smartphones sind bereits vielversprechende Sensoren und Daten verfügbar:

- „Footprint“ Technologie zur kombinierten Positionierung (GPS, GSM, WLAN),
- Gyroskop und Beschleunigungssensor zur Inertialnavigation,
- Magnetometer zur Bestimmung der Orientierung des Geräts und zur Erkennung von Fahrzeugkarosserien,
- Barometer zur Analyse von Höhenprofilen.

Zur Verbesserung der Unterscheidung öffentlicher Verkehrsmittel können Informationen zu räumlicher und zeitlicher Verfügbarkeit (wie Liniennetze oder Fahrplandaten) der Verkehrsmittel genutzt werden. Ansätze dazu finden sich beispielsweise bei CHEN et al. (2011), RIESER-SCHÜSSLER et al. (2011) oder STOPHER et al. (2011). Eine Implementierung dieser Ansätze im Zuge von weiterführender Forschung ist geplant, sodass Trefferquoten über 80 % zu erwarten sind.

Literatur

- CHEN, J., BIERLAIRE, M. & FLÖTTERÖD, G. (2011), Probabilistic multi-modal map matching with rich smartphone data. Paper presented at the 11th Swiss Transport Research Conference, Monte Verità.
- FENG, T., MOISEEVA, A. & TIMMERMANS, H. (2011), Processing of National Travel Survey GPS Pilot Data – A technical report prepared for the Department for Transport. Eindhoven, Netherlands: Eindhoven University of Technology, Department of the built environment.
- FENG, T. & TIMMERMANS, H. (2010), Comparative evaluation of algorithms for GPS data imputation. Paper presented at the 13th WCTR, Rio de Janeiro, Brazil.
- HERRY, M., TOMSCHY, R., SAMMER, G., MESCHIK, M., KOHLA, B., WALLY, R. & FÜRDÖS, A. (2011), MobiFIT – Mobilitätserhebungen basierend auf intelligenten Technologien. Endbericht. Gefördert vom Bundesministerium für Verkehr, Innovation und Technologie, Wien.
- KOHLA, B. (2013), Erkennung von Wegetappen und Verkehrsmitteln für Mobilitätserhebungen mit mobilen Erhebungsgeräten. Doktorarbeit, Institut für Verkehrswesen, Universität für Bodenkultur Wien.
- OLIVER, M., BADLAND, H., MAVOA, S., DUNCAN, M. J. & DUNCAN, S. (2010), Combining GPS, GIS, and Accelerometry: Methodological Issues in the Assessment of Location and Intensity of Travel Behaviors. *Journal of Physical Activity and Health*, 7, 102-108.
- REDDY, S., MUN, M., BURKE, J., ESTRIN, D., HANSEN, M. & SRIVASTAVA, M. (2010), Using mobile phones to determine transportation modes. *ACM Transactions on Sensor Networks*, 6 (2), Article 13.
- RIESER-SCHÜSSLER, N., MONTINI, L. & DOBLER, C. (2011), Improving post-processing routines for GPS observations using prompted-recall data. Paper presented at the 9th International Conference on Survey Methods in Transport, Termas de Puyehue, Chile.

- RUDLOFF, C. & RAY, M. (2010), Detecting travel modes and profiling commuter habits solely based on GPS data. Paper presented at the Transportation Research Board 89th Annual Meeting, Washington DC.
- STOPHER, P. R., KOCKELMAN, K., GREAVES, S. P. & CLIFFORD, E. (2008), Reducing Burden and Sample Sizes in Multiday Household Travel Surveys. *Transportation Research Record*, 2064, 12-18.
- STOPHER, P. R., ZHANG, J. & PRASAD, C. (2011), Evaluating and Improving Software for Identifying Trips, Occupancy, Mode and Purpose from GPS Traces. Paper presented at the 9th International Conference on Survey Methods in Transport, Termas de Puyehue, Chile.
- TROPED, P. J., OLIVEIRA, M. S., MATTHEWS, C. E., CROMLEY, E. K., MELLY, S. J. & CRAIG, B. A. (2008), Prediction of Activity Mode with Global Positioning System and Accelerometer Data. *Medicine and science in sports and exercise. Official Journal of the American College of Sports Medicine*, 40 (5), 972-978.
- WIDHALM, P., NITSCHKE, P. & BRÄNDLE, N. (2012), Transport mode detection with realistic smartphone sensor data. Paper presented at the 21st International Conference on Pattern Recognition (ICPR). Tsukuba, Japan.
- YU, M.-C., YU, T., WANG, S.-C., LIN, C.-J. & CHANG, E. Y. (2013), Low-power and Low-cost Sensor Hub for Transportation-mode Detection (Technical report). HTC and National Taiwan University, Taiwan.