

Abstract (English)

Congestion is a major and increasingly limiting factor for mobility on motorway networks. Using floating car data, congestion states were usually predicted for road segments that were identified by an additional map-matching tool. Not using a map-matcher, 12 billion floating car data observations were cell-wise grouped into two direction classes representing two directions of a motorway. To the best knowledge of the author, the direction distinctive grid-based approach for assigning floating car data to motorway segments is proposed for the first time in this study. The well-known random forest classification algorithm was utilised for developing and forecasting models for single segments and the segment's collective of 1,000 motorway segments and 45 million observations. Evaluation was based on the metrics F_1 -score, misclassification rate, and Bookmaker Informedness. Heuristics based on the average velocity of all motorists at specific points on the motorways served as ground truth for forecasting a segment's congestion state into one of the two classes: *free-flowing* and *congestion*. Whole grid forecasting models delivered better results in comparison to single segment models for four highly congested motorway segments. Major influential factors for the five-minute forecast of the segment collective were features regarding the velocity and the traffic count. Whole grid models are seemingly capable of adding value to the congestion state forecast in 5, 10, 20, 30, and 60 minutes in the future in the whole of North Rhine-Westphalia by considerably exceeding the F_1 and Bookmaker Informedness baseline scores for the 1,000 sample segments. The computational effort was more than 30% lower when using the direction distinctive grid-based approach in comparison to a map-matcher approach for assigning road segments to 2.6 million floating car data observations.

Keywords: traffic congestion, FCD, grid, direction-distinctive, segment, ITS, RF

Abstract (German)

Stau ist ein starker und limitierender Faktor für die Mobilität auf Autobahnen. Bei der Nutzung von floating car data werden üblicherweise Stauvorhersagen auf Straßensegmenten gemacht, welche durch einen sogenannten map-matcher mit den floating car data-Punkten verknüpft werden. Ein map-matcher wurde in dieser Arbeit nicht genutzt, sondern die 12 Billionen gruppierte floating car data-Punkte wurden richtungsunterscheidenden Zellen in einem Netz zugeordnet, welche zwei Richtungen einer Autobahn darstellen. Nach bestem Wissen der Autorin wurde der Ansatz zum ersten Mal in dieser Arbeit vorgeschlagen. Der bekannte random forest Klassifizierungsalgorithmus wurde angewandt, um Vorhersagemodelle für einzelne Segmente und für alle 1000 Segmente im Verbund mit 45 Millionen Beobachtungen zu entwickeln. Die Evaluationsmetriken F_1 -score, misclassification rate und Bookmaker Informedness wurden zur Evaluierung herangezogen. Auf der mittleren Geschwindigkeit aller Fahrzeuge an spezifischen Punkten basierende Heuristiken dienen als ground truth Stau-Labels mit den Klassen *frei-fließender Verkehr* und *Stau*. Für vier einzelne Segmente mit viel Stau führten auf 700 verschiedenen Segmenten basierende Vorhersagemodelle zu besseren Ergebnissen als Modelle für einzelne Autobahn-Segmente. Im Modell mit den vielen verschiedenen Segmenten waren Geschwindigkeits- und Fahrzeuganzahl-Variablen die stärksten Einflussfaktoren. Solche Modelle scheinen einen zusätzlichen Wert für Stauvorhersagemodelle für bis zu 60 Minuten in der Zukunft generieren zu können, da sie für den betrachteten Datenkorpus erheblich besser waren als baseline-Modelle des F_1 -scores und der Bookmaker Informedness. Der Rechenaufwand war bei dem Ansatz von richtungsunterscheidenden Zellen in einem Netz zur Generierung von Segmenten deutlich geringer als bei dem genutzten map-matcher Ansatz bei 2,6 Millionen floating car data Beobachtungen.

Schlagwörter: Stau, FCD, Netz, richtungsunterscheidend, Segment, ITS, RF