

Fakultät für Maschinenwesen

Datenbasierte Methoden zur Prädiktion des Trips und des Fahrprofils

Dipl.-Ing. Univ. Jürgen Matthias Lohrer

Vollständiger Abdruck der von der

Fakultät für Maschinenwesen

der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktor-Ingenieurs (Dr.-Ing)

genehmigten Dissertation.

Vorsitzende: Prof. Dr.-Ing. Birgit Vogel-Heuser

Prüfende der Dissertation:

1. Prof. Dr.-Ing. Markus Lienkamp

2. Prof. Dr.-Ing. Liqiu Meng

Die Dissertation wurde am **15.02.2018** bei der Technischen Universität München eingereicht und durch die **Fakultät für Maschinenwesen** am **27.06.2018** angenommen.

Danksagung

Nun prangt dies Werk im vollen Glanze,
welch Glück und Freud zugleich.
Der Schöpfer übt sich froh im Tanze,
den Dank beschert er reich.

Den Professoren wohl an erster Stelle,
für Chance, Kritik und Leit
und auch für mal gar manche Strenge,
sonst wär's noch nicht so weit.

Den Gefährten für die Korrekturen,
gleich wie die güldne Zeit.
Am Montag für die Prozeduren,
es hat mi ganz schön gfreit!

Das letzte Dankbekenntnis,
das ich hier noch schriebe:
für Kraft, Geduld, Verständnis
meiner großen Liebe!

Januar 2018

Kurzfassung

Die Vernetzung der Fahrzeuge ist maßgeblich für die zukünftige Entwicklung in der Automobilindustrie. Das Fahrzeug als Sensor sammelt eine Fülle an Daten. Als Herausforderung gilt es, einen Mehrwert aus diesen Daten zu erzeugen. So ist es das Ziel fahrzeugtechnische Innovationen zu generieren, bei denen der Kunde den Vorteil in der Verwendung seiner Daten erkennt. Als entscheidende Information für viele Funktionen gilt das Geschwindigkeitsprofil. Es ist die Haupteinflussgröße für den Verbrauch und wichtig für Sicherheitsfunktionen während der Fahrt. Es ist überwiegend abhängig von der Position, der gewählten Route und dem Verkehr. Um eine Vorhersage dieser Größe zu jedem Zeitpunkt zu ermöglichen ist darüber hinaus eine Prädiktion des nächsten Routenabschnitts notwendig. Daher werden in dieser Arbeit zwei Funktionen entwickelt, die Prädiktion des Trips sowie die Prädiktion des Fahrprofils entlang der erwarteten Route, auf deren Basis weitere prädiktive, standortbezogene Kundenfunktionen aufbauen können.

Für die Entwicklung werden 5 Mio. km reale Fahrprofile von über 200 verschiedenen Fahrern verwendet. Zur Prädiktion der nächsten Routenabschnitte wird ein neuronales Netz, im speziellen ein *Memory Network Model*, eingesetzt. Neben dem allgemeinen Ansatz, der hauptsächlich auf den Fahrten mehrerer Taxis basiert, wird die Methode um eine nutzerindividuelle Variante erweitert. Die Zahl der fehlerhaften Prädiktionen über die Testfahrten reduziert sich um ca. 5 % gegenüber einen vergleichbaren Ansatz aus dem Stand der Technik. Mit der nutzerspezifischen Erweiterung verbessert sich die Prädiktion um über 17 %.

Bei der Fahrprofilprädiktion werden Muster in dem Bewegungsverhalten der Fahrzeuge auf einem Streckensegment erkannt. Dabei wird die Ähnlichkeit in den Geschwindigkeitsprofilen berechnet und über ein spektrales Clusterverfahren werden Klassen gebildet. Als Prädiktionsmodell wird eine multinomiale logistische Regression verwendet, das neben Klassen der bereits zurückgelegten Straßensegmente vor allem den Verkehrsfluss berücksichtigt. Neben der Geschwindigkeit werden Konfidenzintervalle für die Gierrate und die Querbeschleunigung für die Straßensegmente und die zugehörigen Klassen ermittelt. Die zurückliegenden Streckensegmente zeigen einen relevanten Beitrag für die Prädiktion. Die Trefferquote beträgt im Durchschnitt ca. 84 % und erweist sich als abhängig von der Straßenklasse. In mehreren Experimenten wird das Prädiktionsmodell für Fahrten auf verschiedenen Routen eingesetzt und dadurch gezeigt, dass eine gute Übereinstimmung der prädizierten Größen der Geschwindigkeit und der Gierrate mit den aufgezeichneten Messgrößen besteht.

Um einen quantifizierbaren Nutzen der beiden Funktionen zu zeigen, wird eine Betriebsstrategie für ein vernetztes Plug-In-Hybridfahrzeug entwickelt. Die Vorgabe für den SOC-Verlauf wird abhängig von der prädizierten Route, dem Geschwindigkeitsprofil und der Ladeinfrastruktur gestaltet. Für über 200 Fahrten wurden Kraftstoffeinsparungen von ca. 8,0 % gegenüber einem nicht vorausschauenden Verfahren ermittelt.

Inhaltsverzeichnis

1	Einleitung.....	1
1.1	Motivation.....	2
1.2	Ziel der Arbeit	5
1.3	Aufbau der Arbeit	5
2	Vernetzung und Prädiktion im Fahrzeug	7
2.1	Grundlagen der Vernetzung von Fahrzeugen	7
2.1.1	Fahrzeuginterne Technologien	7
2.1.2	Kommunikation	7
2.1.3	Back-End Architektur	8
2.2	Location-based Services	9
2.2.1	Komponenten	9
2.2.2	Architekturen und Datenbeschreibung	10
2.2.3	Lokale dynamische Karte	11
2.2.4	Dienste im Automobil.....	11
2.3	Ableitung der Themenstellung.....	12
2.3.1	Forschungsfrage	12
2.3.2	Lösungsansatz	13
2.3.3	Vorgehen	15
3	Aufbereitung der Datenbasis	17
3.1	Datenerfassung und Selektion	17
3.1.1	Flottenversuche	17
3.1.2	Kartendaten	20
3.2	Vorverarbeitung und Aggregation	21
3.2.1	Mapmatching.....	21
3.2.2	Metadaten.....	23
3.2.3	Verkehrsinformationen	23
4	Trip-Prädiktion	25
4.1	Stand der Forschung	25
4.1.1	Verarbeitung und Bewertung von Mobilitätsdaten	25
4.1.2	Fahrtzielprädiktion.....	27
4.1.3	Routenprädiktion.....	28
4.1.4	Trip-Prädiktion	29
4.1.5	Zusammenfassung und Abgrenzung	33
4.2	Methodik	34
4.2.1	Systemansatz	34
4.2.2	Modellform	35
4.2.3	Eingangsdaten und Parametrierung	38
4.2.4	Nutzerindividueller Einfluss	40
4.2.5	Prädiktionshorizont.....	41
4.3	Ergebnisse	42
4.3.1	Prädiktion während der Fahrt.....	42
4.3.2	Prädiktionshorizont.....	43

4.3.3	Beispielfahrt	45
4.4	Bewertung	46
5	Fahrprofilprädiktion	48
5.1	Stand der Forschung	48
5.1.1	Mathematische Modelle	48
5.1.2	Datenbasierte Ansätze	51
5.1.3	Bewertung von Fahrprofilen	53
5.1.4	Zusammenfassung und Abgrenzung	55
5.2	Methodik	56
5.2.1	Systemansatz	56
5.2.2	Erzeugung charakteristischer Fahrprofile	57
5.2.3	Prädiktion des Fahrprofils	63
5.2.4	Prädiktion weiterer Bewegungsgrößen	68
5.2.5	Ableitung von Metadaten	71
5.3	Ergebnisse	73
5.3.1	Bewertung der repräsentativen Profile	73
5.3.2	Analyse eines Prädiktionsmodells	76
5.3.3	Bewertung der Prädiktionsmodelle	78
5.3.4	Prädiktionshorizont	82
5.3.5	Beispielfahrt	84
5.4	Bewertung	87
6	Anwendung in einer Betriebsstrategie	90
6.1	Systemansatz der Betriebsstrategie	90
6.2	Regelstrategie	91
6.2.1	Modell	91
6.2.2	Optimierung	92
6.3	Ergebnisse	93
6.3.1	Beispielfahrt	93
6.3.2	SOC-Verlauf	93
6.3.3	Erwartete Kraftstoffeinsparungen	95
7	Diskussion	96
7.1	Zielsetzung	96
7.2	Forschungsfragen	96
7.3	Einschränkungen und Übertragbarkeit	98
8	Zusammenfassung und Ausblick	99
	Abbildungsverzeichnis	101
	Tabellenverzeichnis	103
	Formelzeichen	104
	Abkürzungen	107
	Literatur	110
	Studentische Arbeiten	128
	Eigene Veröffentlichungen	129
	Anhang	130

1 Einleitung

Die technologischen Trends in der Automobilindustrie werden angeführt von dem Thema *Vernetzung und Digitalisierung*, dicht gefolgt von der *Elektromobilität* [1].

Die Vernetzung von Fahrzeugen ist ausschlaggebend für die Digitalisierung. Darunter wird zum einen die Kommunikation zwischen den Fahrzeugen verstanden. Dadurch können bspw. Objektlisten von Fahrerassistenzsystemen ausgetauscht und somit die Sicherheit erhöht werden [2]. Der weitere Bestandteil der Vernetzung ist die Verbindung des Fahrzeugs mit dem Back-End. Die digitale Karte im Fahrzeug kann dadurch mit temporären Verkehrsbehinderungen oder einer durch Baustellen geänderten Verkehrsführung aktualisiert werden. KPMG [1] bezeichnet die Vernetzung als maßgeblich für die Entwicklung in der Automobilindustrie. Bereits heute verändert sich die Wertentwicklung über den Lebenszyklus durch die Vernetzung [3]. Nach aktuellem Stand ist jedes zweite Neufahrzeug vernetzt [4]. Die derzeitigen Anteile am gesamten Fahrzeugbestand liegen bei unter 10 % und werden für 2020 auf rund 20 % prognostiziert [4]. Zudem besteht ein Markt für Nachrüstlösungen, die den Fahrern von gebrauchten Fahrzeugen über die On-Board-Diagnose (OBD)-Schnittstelle vernetzte Funktionen zu Verfügung stellen [5].

Die Vorteile der Vernetzung sind durch die hohen Marktanteilen der vernetzten Fahrzeuge bereits jetzt nutzbar. Die Kunden setzen die Erwartungen in erster Linie in eine Steigerung der Sicherheit, gefolgt von finanziellen Einsparungen und zuletzt die Optimierung der Fahrzeit und des Fahrkomforts [6]. Die Vernetzung ist zudem essentiell für autonome Fahrzeuge. Um aktuelle und hoch genaue Karteninformationen im Fahrzeug zu erhalten, ist eine Back-End-Verbindung erforderlich. Der große Vorteil, der durch die Einführung von autonomen Fahrzeugen erwartet wird, ist die mögliche Steigerung der Effizienz durch die Optimierung des Verkehrsflusses und einer Vernetzung der Teilnehmer [7]. Allerdings ist dieser Effekt stark abhängig von der Durchdringung der autonomen bzw. vernetzten Fahrzeuge und verbessert sich bei niedriger Penetration nicht signifikant [8].

McKinsey & Company [9] beschreibt das *autonome Fahren* zwar als Neudefinition der Mobilität, allerdings sind die Auswirkungen derzeit noch schwer abzuschätzen. Während ein hochautomatisierter Einsatz nach [10] (SAE Level 3) für Autobahnen bereits ab ca. 2018 [11, S. 107] in Serienfahrzeugen von mehreren Herstellern vorzufinden sein wird, ist mit einer Hochautomatisierung (SAE Level 4) im urbanen Bereich frühestens ab 2030 [4] zu rechnen. Aktuelle Studien schätzen die Neuzulassungsraten von Stufe 3 Fahrzeugen für 2030 auf 60 % bis 75 %, Stufe 4 Fahrzeuge nur auf 15 % bis 25 % [4, 12]. Damit liegt die voraussichtliche Marktdurchdringung von Stufe 3 Fahrzeugen in 2030 bei ca. 40 % und die der Stufe 4 Fahrzeuge bei unter 10 % [4].

Anders als bei autonomen Fahrzeugen hat der Wandel durch die Vernetzung bereits begonnen. Neuerungen sind im Bereich der datengetriebenen Dienste entstanden. Das Fahrzeug als Teil des Internet of Things (IOT) etabliert sich dabei als größter Sensor und ist in der Lage eine Fülle an Daten zu sammeln. Intel [13] beziffert die Zahl für Daten, die im Fahrzeug verarbeitet werden, auf ein Gigabyte pro Sekunde. Wird auch nur ein Teil dieser Daten weiter verarbeitet, ist es notwendig, dass ein vollständiges Ökosystem mitentsteht. Neben der Kommunikation entstehen sowohl weitere fahrzeuginterne, als auch beachtliche infrastrukturelle Anforderungen, um mit den Daten entsprechend umzugehen. Weiterhin sind geeignete Back-End-Lösungen erforderlich, um die Datenmenge nutzbar zu machen. Eine Übersicht nach McKinsey&Company [12] wird in Abb. 1.1 gegeben.

Grundsätzlich existiert eine Vielzahl an Möglichkeiten, einen Mehrwert aus den Daten zu erzeugen. Bratzel [14] unterscheidet dabei zwischen neuen Mobilitätsdienstleistungen und fahrzeugtechnischen Innovationen, wobei letztere in Bedien- und Anzeigeconzepte, Informations- und Kommunikationssysteme, Sicherheitssysteme, sowie sonstige Technologiefelder unterteilt werden. Unabhängig von der Anwendung sorgen sich die Kunden um die Sicherheit und den Datenschutz bei diesen Systemen [5, 6].

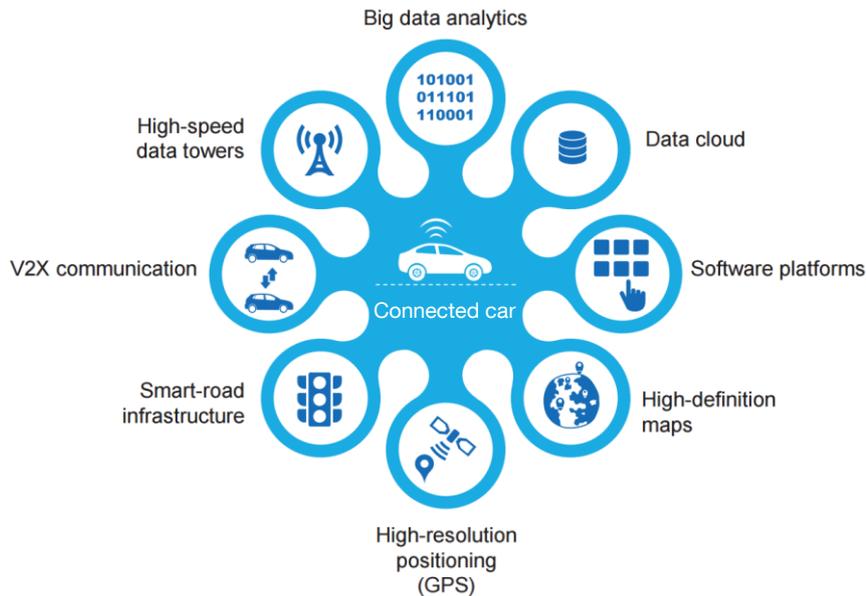


Abbildung 1.1 : Infrastrukturtechnologien für den Umgang mit Fahrzeugdaten [12]

Laut KPMG [1] wünschen die Kunden über ihre Daten zu verfügen, während viele Fahrzeughersteller den Besitz der Daten bei ihnen sehen. Je nachdem wie die Funktionen im Gegenzug für den Kunden angeboten werden können, anonymisiert oder individualisiert, entscheidet sich die Nutzung der Daten [15]. Ein verantwortungsvoller Umgang mit den Daten, sowie ein wirksamer Schutz vor Unberechtigten sind hierbei wichtige Leitlinien in Unternehmen [16, S. 82].

1.1 Motivation

Die Motivation, die sich für diese Arbeit ableitet, ist es Funktionen zu schaffen, die einen Mehrwert für vernetzte Fahrzeuge aus den gesammelten Daten generieren. Grundsätzlich können dabei mehrere Ziele verfolgt werden. Aus der Sicht des Kunden werden dabei *Energieeffizienz*, *Sicherheit* und *Fahrkomfort* unterschieden, dargestellt an drei bereits existierenden Beispielszenarien:

- **Energieeffizienz:** Über aktuelle und historische Verkehrsflussinformationen kann eine Prognose der Verkehrslage erzeugt werden. Online Navigationsdienste können aus den prä-dizierten Daten für die geplante Fahrt die schnellste, oder auch energieeffizienteste Route bestimmen. Dadurch sind Verbrauchseinsparungen von bis zu 14 % realisierbar [17].
- **Sicherheit:** Über das Elektronischen Stabilitätsprogramm (ESP) und dessen Sensorik wird im Fahrzeug kontinuierlich ein Reibwert für die befahrene Straße geschätzt. Wird dieser Wert in das Back-End übertragen, können Algorithmen die Schätzung durch mehrere Fahrzeuge, die auf dem Straßenabschnitt unterwegs waren, verbessern [18]. Bei Nässe oder Glätteis können Fahrzeuge auf einen niedrigen Reibwert eingestellt werden, bevor das Fahrzeug einen kritischen Zustand erreicht.
- **Komfort:** Mit Hilfe gesammelter Bewegungsprofile werden persönliche Point of Interest (POI), wie bspw. die Adresse des Zuhauses oder der Arbeitsstätte ermittelt. Navigationsdienstleistungen greifen auf das Bewegungsverhalten zurück und schließen über auftretende Bewegungsmuster oder den persönlichen Kalender auf das nächste Ziel und weisen den Fahrer bspw. darauf hin, wann er die Fahrt antreten muss, um rechtzeitig vor Ort zu sein [19, 20]. Damit kann auf die auftretende Verkehrslage frühzeitig reagiert werden.

Folgende Eigenschaften für derartige Funktionen lassen sich aus den genannten Beispielen extrahieren:

- **Datengetrieben:** Die Grundlage für die Untersuchung bilden Mobilitätsdaten. Diese werden entweder von einem Fahrzeug bzw. Fahrer oder von einer Vielzahl unterschiedlicher Verkehrsteilnehmer gesammelt.
- **Standortbezogen:** Die Daten sind zu meist bezogen auf eine Adresse oder einen Straßenabschnitt. Je nach Position werden unterschiedliche Informationen erhoben bzw. dem Fahrzeug mitgeteilt. Derartige standortbezogene Dienste oder Funktionen fallen unter die Kategorie Location-based Services (LBS).
- **Individuell:** Die Dienste erfolgen zum Teil personalisiert. Die aus den persönlichen Bewegungsprofilen gelernten Informationen können speziell für den Nutzer eingesetzt werden.
- **Prädiktiv:** Die Funktionen haben einen prädiktiven Charakter. Aus den historischen Daten werden Muster extrahiert und ein wahrscheinliches Verhalten für einen zukünftigen Zeitpunkt vorhergesagt.

Die Beispiele zeigen, dass durch die Prädiktion der Umgebungsbedingungen oder des Fahrzustands ein Mehrwert für das Fahrzeug und den Nutzer geschaffen wird. Die gewonnenen Informationen sind vielseitig einsetzbar. So wird der Verkehrsfluss in dynamischen Routingalgorithmen [21], in der Ermittlung des Energieverbrauchs [22] und zur Abschätzung von Emissionen durch den Verkehr [23] eingesetzt. Die Verkehrssituation und das Verkehrsaufkommen haben einen ausschlaggebenden Einfluss auf den Fahrzustand und damit auf die Energieeffizienz [24] und die Fahrsicherheit [25]. Die Ermittlung und die Prädiktion des Verkehrsflusses ist in der Forschung präsent und weit fortgeschritten [26, 27].

Um präzise Aussagen über den individuellen Fahrzustand, insbesondere in Hinblick auf die Bewegungsgrößen des Fahrzeugs, zu machen, ist eine genaue Betrachtung des Fahrverhaltens notwendig. Nach Ericsson [28] ist das Fahrverhalten mehrdimensional und wird durch mehrere Parameter beschrieben. Ericsson identifiziert dabei 16 Haupteinflussgrößen aus 62 Parametern, die speziell für die Emissionen und den Energieverbrauch ausschlaggebend sind. Die Haupteinflussgrößen sind in Tab. 1.1 zusammengefasst.

Aus Tab. 1.1 wird ersichtlich, dass elf der 16 Größen durch die Geschwindigkeit und die Längsbeschleunigung des Fahrzeugs beschrieben werden. Die restlichen fünf Größen sind von dem Antriebssystem abhängig. Eine Prädiktion von Geschwindigkeit und Beschleunigung eröffnet mehrere Möglichkeiten. Zum einen kann direkt auf diese Größen Einfluss genommen und auf die erwartete Situation hin optimiert werden [29]. Zum anderen können Systeme und Komponenten auf den auftretenden Betriebszustand hin geeignet konditioniert werden [30, 31].

Das Fahrverhalten mit mathematischen Modellen zu beschreiben erweist sich als Herausforderung auf Grund der vielfältigen Einflussgrößen [32]. Faktoren wie Straßenparameter oder die Umfeldsituation sind flächendeckend aufwändig zu erfassen. Ein daten- und ortsbasierter Ansatz wird für diese Fragestellung als zielführend erachtet.

Neben der Optimierung von Emissionen und dem Energieverbrauch sind weitere Ziele denkbar, die durch die Prädiktion von Bewegungsgrößen verfolgt werden können. Für eine Steigerung der Fahrsicherheit, insbesondere in Kurvensituationen, werden neben der Kurvengeschwindigkeit auch die auftretende Querschleunigung, sowie die Gierrate in Betracht gezogen. Mit Hilfe von Umweltinformationen kann bspw. eine Warnung in kritischen Kurven gegeben werden [33]. Zur Steigerung des Fahrkomforts werden Feder-Dämpfer-Systeme auf die auftretenden Vertikalbeschleunigungen, die durch unterschiedliche Straßeneigenschaften induziert werden, vorab eingestellt.

Am Beispiel eines Kurvenwarnsystems wird deutlich, dass der Prädiktionshorizont über die Reichweite der Fahrzeugsensorik hinaus geführt werden muss. Allerdings ist eine detaillierte Prädiktion der Bewegungsgrößen des Fahrzeugs für einen längeren Zeitraum nicht möglich. Der Einfluss von weiteren Verkehrsteilnehmern, besonders im Innenstadtbereich, führt zu spontanen Änderungen des Fahrzustands [34]. Nach dem Modell der Fahreraufgaben nach

Tabelle 1.1 : Haupteinflussgrößen aus Fahrprofilen für Emissionen und Verbrauch nach Ericsson [28]

	Name	Typischer Parameter
1	Verzögerung	Durchschnittliche Verzögerung
2	Beschleunigung mit hohem Energiebedarf	Relative positive Beschleunigung
3	Stillstand	Anteil der Zeit $v < 2 \frac{\text{km}}{\text{h}}$
4	Oszillation der Geschwindigkeit	Frequenz der Oszillation der Geschwindigkeit pro 100 s
5	Beschleunigung mit moderatem Energiebedarf	Zeitanteil va $3-6 \frac{\text{m}^2}{\text{s}^3}$
6	Extreme Beschleunigung	Zeitanteil $a > 2.5 \frac{\text{m}}{\text{s}^2}$
7	Geschwindigkeit von 15 bis 30 $\frac{\text{km}}{\text{h}}$	Zeitanteil v $15-30 \frac{\text{km}}{\text{h}}$
8	Geschwindigkeit von 90 bis 110 $\frac{\text{km}}{\text{h}}$	Zeitanteil v $90-110 \frac{\text{km}}{\text{h}}$
9	Geschwindigkeit von 70 bis 90 $\frac{\text{km}}{\text{h}}$	Zeitanteil v $70-90 \frac{\text{km}}{\text{h}}$
10	Geschwindigkeit von 50 bis 70 $\frac{\text{km}}{\text{h}}$	Zeitanteil v $50-70 \frac{\text{km}}{\text{h}}$
11	Spätes Hochschalten aus Gang 2 oder Gang 3	Zeitanteil Motordrehzahl $2\,500-3\,500 \frac{1}{\text{min}}$ in Gang 2, 3
12	Motordrehzahl $> 3\,500 \frac{1}{\text{min}}$	Zeitanteil Motordrehzahl $> 3\,500 \frac{1}{\text{min}}$
13	Geschwindigkeit $> 110 \frac{\text{km}}{\text{h}}$	Zeitanteil $v > 110 \frac{\text{km}}{\text{h}}$
14	Moderate Motordrehzahl in den Gängen 2 und 3	Zeitanteil Motordrehzahl $1\,500-2\,500 \frac{1}{\text{min}}$ in Gang 2, 3
15	Niedrige Motordrehzahl in Gang 4	Zeitanteil Motordrehzahl $< 1\,500 \frac{1}{\text{min}}$ in Gang 4
16	Niedrige Motordrehzahl in Gang 5	Zeitanteil Motordrehzahl $< 1\,500 \frac{1}{\text{min}}$ in Gang 5

Donges [35] orientiert sich ein möglicher Horizont zwischen den Bereichen von Fahrzeugführung und Navigation (Abb. 1.2).

Daraus wird ersichtlich, dass die Route des Fahrzeugs eine wichtige Information für die Prädiktion der Fahrzeugbewegung in dem Zeithorizont ist. Die Route, die der Fahrer wählt, ist dem Fahrzeug nicht zwingenderweise bekannt. Die Verbreitung von Navigationsgeräten ist mit über 75 % bereits hoch [36, 37]. Allerdings werden davon mobile Geräte und Smartphones bevorzugt [36, 38]. Nach Skoglund [39] und Karlsson [40] gibt die Mehrheit der Befragten an, in weniger als 25 % der Fälle ein Navigationsgerät zu verwenden. Forbes [41, S. 137] stellt fest, dass lediglich 10,2 % unter 872 Befragten in bekannten Gebieten die Geräte häufig bzw. immer aktiv nutzen. Ist der Zielort nicht bekannt wird das Navigationsgerät in 85,6 % der Fälle vergleichbar eingesetzt. Zu ähnlichen Erkenntnissen gelangt Göktürk [42]. Falls der Zielort unbekannt, bzw. die Route länger ist, wird eine Führung durch die Navigation häufiger verwendet [40, 43].

Einen weiteren Sonderfall stellen Taxifahrer dar. In der Ortskundeprüfung (Beispiel Berlin [44]) müssen die Fahrer über eine detaillierte Ortskenntnis des Straßennetzes verfügen. Um die Prädiktion daher jederzeit nutzbar zu machen, ist entweder die Routeninformation des Navigationsgeräts zu verwenden oder, falls diese nicht existiert, die Route über den Horizont zu prädizieren.

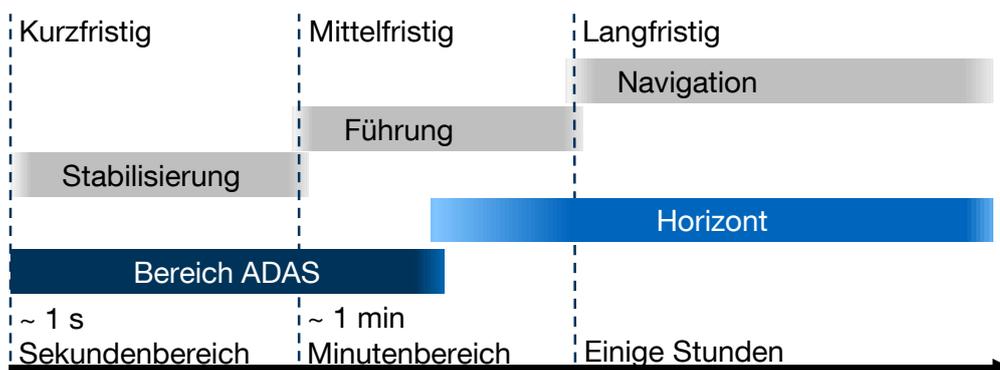


Abbildung 1.2 : Fahreraufgaben im Zeitbereich nach Donges [35].

1.2 Ziel der Arbeit

Diese Arbeit hat zum Ziel, eine Methode für die Prädiktion des Trips und des Fahrprofils über einen begrenzten Zeithorizont aufzuzeigen. Ein Trip bezeichnet dabei einen Teil der Route des Fahrzeugs. Unter dem Fahrprofil werden Zeitreiheninformationen der Bewegungsgrößen des Fahrzeugs verstanden. Dazu zählen insbesondere der Geschwindigkeitsverlauf, die Beschleunigungen und die Drehraten des Fahrzeugs. Es soll ein Mehrwert aus aufgezeichneten Fahrdaten unterschiedlicher Fahrer und Fahrzeuge geschaffen werden.

Dies soll über eine serverbasierte Funktion weiteren Fahrzeugen zu Verfügung gestellt werden. Dabei können nutzer- und fahrzeugindividuelle Eigenschaften berücksichtigt werden. Um jederzeit die Funktionalität gewährleisten zu können, wird ein datenbasierter Ansatz gewählt, der nutzerspezifische Daten im Back-End anonymisiert und fahrerindividuelle Optimierungen im Fahrzeug verarbeitet. Um eine hohe Detailtiefe bei der Prädiktion der Bewegungszustände zu erlangen, werden Zeitreiheninformationen generiert. Dabei werden aggregierte und gemittelte Werte, wie ein ortsspezifischer Verkehrsfluss, mit berücksichtigt.

Die Ergebnisse der Prädiktion sollen an einem ausgewählten Beispiel zur Optimierung des Fahrzeugs hinsichtlich Energieeffizienz dargestellt werden. Der dadurch erzielbare Mehrwert wird durch Simulationen quantifiziert.

1.3 Aufbau der Arbeit

Der Aufbau der Arbeit ist in Abb. 1.3 dargestellt. Kapitel 2 geht auf die Grundlagen der Vernetzung von Fahrzeugen ein. Anschließend werden standortbezogene Dienste ausführlicher erläutert. Daraus wird die Themenstellung für diese Arbeit abgeleitet, der gewählte Lösungsansatz dargestellt und auf das Vorgehensmodell eingegangen. Zur Erläuterung der Datenbasis wird in Kapitel 3 auf die Erfassung und Verarbeitung der verwendeten Daten eingegangen. Anschließend wird in Kapitel 4 die Prädiktion des Trips und in Kapitel 5 die Prädiktion des Fahrprofils erläutert. In beiden Kapiteln wird folgende Struktur verwendet: Zuerst wird auf den Stand der Forschung eingegangen. Anschließend wird die Methodik näher erläutert. Ergebnisse der Prädiktion werden darauf folgend vorgestellt und diese abschließend bewertet. In Kapitel 6 wird eine beispielhafte Anwendung skizziert. Die Diskussion der Ergebnisse dieser Arbeit erfolgt in Kapitel 7. Abschließend wird in Kapitel 8 eine Zusammenfassung und ein Ausblick für aufbauende Forschungen gegeben.

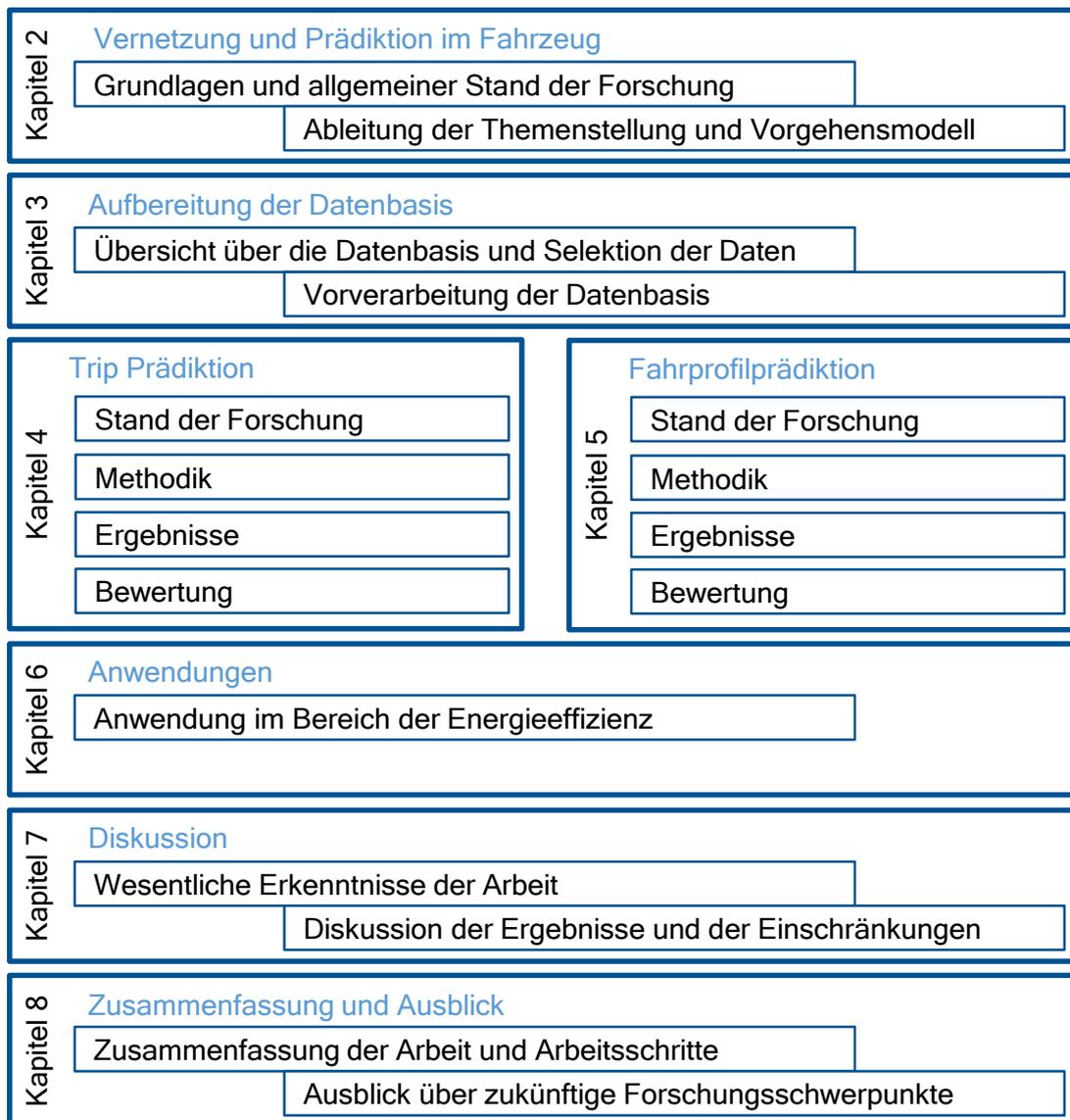


Abbildung 1.3 : Überblick über den Aufbau dieser Arbeit

2 Vernetzung und Prädiktion im Fahrzeug

In diesem Kapitel werden zuerst die Grundlagen für eine Vernetzung von Fahrzeugen mit einem Back-End-System vorgestellt. Dabei wird auf die erforderlichen Technologien und Komponenten sowohl im Fahrzeug als auch im Back-End eingegangen. Anschließend werden standortbezogene Dienste näher erläutert. Speziell im Hinblick auf datenbasierte Anwendungen werden dafür Beispiele aus Industrie und Forschung herangezogen. Aus den Kritikpunkten zum Stand der Forschung wird die Themenstellung für diese Arbeit abgeleitet und der gewählte Lösungsansatz aufgezeigt. Zuletzt wird auf das gewählte Vorgehen eingegangen.

2.1 Grundlagen der Vernetzung von Fahrzeugen

Wie bereits in Kapitel 1 eingeführt, müssen mehrere Bereiche bei der Vernetzung des Fahrzeugs mit dem Back-End betrachtet werden. In den folgenden Abschnitten wird hierbei nach den Bereichen fahrzeuginterne Technologien, Kommunikation und Back-End-Architektur unterschieden.

2.1.1 Fahrzeuginterne Technologien

Um fahrzeuginterne Daten im Back-End nutzbar zu machen, ist die Bordnetzarchitektur des Fahrzeugs dementsprechend zu entwerfen. Nach Buckl [45] spielt die Informations- und Kommunikationstechnologie (IKT) eine entscheidende Rolle und ist der Treiber für die Veränderung des Bordnetzes. Es existieren verschiedene Architekturkonzepte, die meist entweder nach Funktion oder nach Fahrzeugdomäne aufgebaut werden [46, 47]. Bei der Funktionsarchitektur werden in verschiedenen Lagen die Datenverarbeitung, Entscheidungsprozesse und Ausführung der Funktionen geregelt [48, 49]. Bei der am verbreitetsten eingesetzten Domänenarchitektur wird das Fahrzeug nach Baugruppen unterteilt. Für die Kommunikation über die Domänen hinweg ist ein zentrales Gateway erforderlich, das den Nachrichtenaustausch koordiniert und berechtigt [50, S. 45].

Um die Komplexität mit zunehmender Zahl an Funktionen und Komponenten zu reduzieren, geht der Trend in Richtung einer zentralisierten IKT-Architektur [45, 51]. Ein Beispiel für ein derartiges Konzept wird in [52] dargestellt. Die interne Kommunikation verläuft hierbei über die zentrale Einheit, die gleichzeitig als Gateway fungiert. Die externe Kommunikation erfolgt über einen mobilen Router, der eine duale Long Term Evolution (LTE)-Verbindung aufbauen kann. Entertainmentdienste und Fahrzeugfunktionen werden entkoppelt voneinander von verschiedenen Einheiten ausgeführt. Grundsätzlich ermöglicht ein derartiger Aufbau die sichere Übertragung aller im Fahrzeug generierten Daten.

An die Kommunikation mit externen Komponenten werden besondere Anforderungen hinsichtlich Übertragungssicherheit und Datenschutz gestellt [53]. Glas [54] zeigt einen Ansatz für eine echtzeitfähige Kommunikationsabsicherung im Fahrzeug. Über eine Hardware-Firewall werden die beiden Kommunikationsdomänen Intra-Fahrzeug und Inter-Fahrzeug miteinander verknüpft. Letztere übernimmt die Übertragung, Filterung, Sicherheitsüberprüfung und das Routing der Nachrichten. Damit können Angriffe von Außen frühzeitig abgewehrt werden. Sicherheitskritische Nachrichten werden priorisiert und die Auslastung des Systems gezielt gesteuert.

2.1.2 Kommunikation

Bei der Übertragung der Daten werden unterschiedliche Medien und Protokolle verwendet. Gemäß des Open Systems Interconnection Model (OSI) Schichtenmodells [55] der Internationale Organisation für Normung (ISO) dient bei der Kommunikation zwischen Fahrzeugen (Car-to-Car (C2C)) oder mit der straßennahen Infrastruktur (Car-to-Infrastructure (C2I)) der Standard IEEE 802.11p als Übertragungsmedium [56]. Für den Fernbereich mit dem Back-End werden

die örtlich verfügbaren Mobilfunkstandards (LTE, Universal Mobile Telecommunications System (UMTS)) verwendet. Je nach Netzabdeckung und Kapazität können große Datenmengen mit bis zu 300 Mbit/s übertragen werden [11]. Während im Bereich C2C die Transportprotokolle speziell für Automotive-Anwendungen standardisiert werden [56], wird bei der Kommunikation über Mobilfunk das Internettransportprotokoll Transmission Control Protocol (TCP) eingesetzt.

Ein Standard für das Nachrichtenformat im Bereich der Verkehrstechnik ist das Protokoll Global Automotive Telematics Standard (GATS) [57, S. 32–34],[58]. Hierbei werden drei Schichten definiert. In der Transportschicht wird die applikationsabhängige Adressierung und Organisation der Nutzdaten vorgenommen. Die zweite Schicht dient zur Verschlüsselung und der Authentifizierung. Die dritte Ebene beschreibt das Datenprotokoll der Applikation. Es können verschiedene Dienstypen wie bspw. Verkehrsinformationen, Sicherheitsdienste und die Übertragung von Floating Car Data (FCD) eingesetzt werden [59, S. 68]. Neben GATS findet auch das Application Communication Protocol (ACP) in der Fahrzeugtechnik Einsatz. Ein weiterer Standard zum Austausch von Mobilitätsdaten ist Datex II [60]. Das von der europäischen Kommission vorgeschlagene Protokoll definiert XML-Strukturen und plattformunabhängige Regeln zur Datenübertragung.

Die übertragenen Dateninhalte sind abhängig von der jeweiligen Applikation und in vielen Fällen nicht eindeutig spezifiziert. Das für die Verkehrslageerfassung durch Fahrzeuge verwendete FCD Datenpaket enthält dabei mindestens die Fahrzeugkennung, eine Zeitangabe und die Position des Fahrzeugs [59]. Weitaus umfangreichere Fahrzeug- und Umgebungsinformationen werden mit dem von BMW entwickelten extended FCD (xFCD) Format übertragen [61]. So können bspw. neben der Position und der Geschwindigkeit des Fahrzeugs, die Außentemperatur oder ABS-Eingriffe übertragen werden. Zudem ist eine Unterscheidung in anonymisierte (bspw. Warnblinker) und personalisierte Daten (bspw. Maut, Route) möglich [62].

2.1.3 Back-End Architektur

Der Informationsaustausch zwischen Systemen wird im Bereich der Telematiksysteme thematisiert. Dieser hat nach [63, S. 68–71] besonders durch die Verknüpfung von kundenorientierter Elektronik mit dem Automobil an Bedeutung gewonnen. In der Verkehrstechnik spielt dieser Bereich schon länger eine wichtige Rolle.

Unter anderem in den Projekten CVIS [64], MOBiNET [65, 66] und DITCM architecture [67, 68] werden kooperative Systemarchitekturen für den Austausch der Daten zwischen Fahrzeug, straßennaher Infrastruktur und dem Back-End beschrieben. Ein einheitlicher Standard existiert nicht. Im Bereich der Automobiltechnik wurde 2010 die zweite Version des Telematikschemas Next Generation Telematics Patterns (NGTP) von dem Hersteller BMW, dem Telematikdienstleister WirelessCar, dem Geodatenanbieter Navteq (jetzt HERE) und dem IT- und Telematikentwickler Ygomi als offener Architekturstandard vorgestellt [69]. Dieser ermöglicht den flexiblen und skalierbaren Aufbau von Back-End-basierten Diensten. In Abb. 2.1 werden die wesentlichen Elemente dargestellt.

Das Fahrzeug verfügt über eine Telematic Unit (TU), mit den notwendigen Schnittstellen (Kapitel 2.1.1), und kommuniziert über Mobilfunk mit dem Dispatcher (DSPT). Dieser regelt den Nachrichtenfluss zwischen Fahrzeug und dem Back-End und kann sich je nach Land oder Region unterscheiden. Die eingehende Nachricht wird nach Anfrage über den Provisioning Data Provider (PDP) mit Zusatzinformationen versehen und entsprechend weitergeleitet. Der Service Handler (SH) de- bzw. encodiert die Servicedaten der Nachricht und ergänzt über den Customer Data Provider (CDP) im Bedarfsfall fahrzeug- oder nutzerspezifische Informationen. Ein oder mehrere Service Integratoren (SI) führen die Telematikdienste zusammen. Die Architektur integriert damit unter anderem ein automatisches Notrufsystem (eCall), das den Notruf mit den relevanten Informationen an den Public Safety Answering Point (PSAP) weiterleitet. Über den SI ist es möglich Daten von weiteren Content Provider (CP) oder zusätzliche proprietäre Dienste über weitere Service Provider (SP) anzubieten. Darauf wird in Kapitel 2.2 detailliert eingegangen.

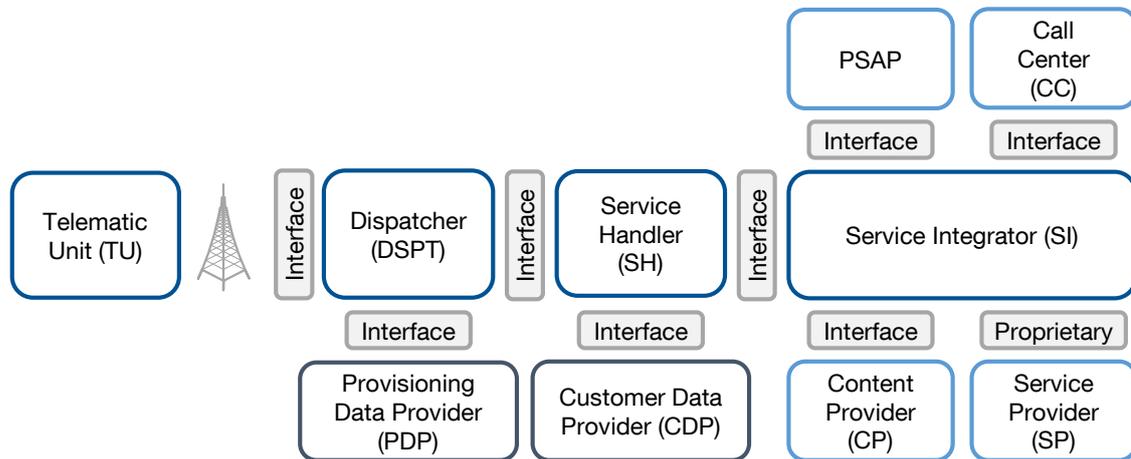


Abbildung 2.1 : Aufbau Next Generation Telematics Pattern nach [69]

Aufbauend auf der NGTP Architektur hat WirelessCar die Erweiterungen Telematics Service Delivery Platform [70] für eine sichere und effiziente Einbindung von Nutzerportalen, der Geschäftsprozessintegration des jeweiligen Herstellers und Kundendiensten vorgestellt. Neben NGTP existieren mehrere Architekturen, die ähnliche Ziele verfolgen. Haberle [71] stellt bspw. eine Plattform für die Prototypenentwicklung von Telematikdiensten vor.

2.2 Location-based Services

Die Schnittmenge aus Internet, mobilen Geräten der IKT und Geoinformationssystem (GIS) bezeichnet Brimicombe [72] als Konvergenz für die Erstellung von LBS. Vor allem durch die Verbreitung von Smartphones und die zunehmende Digitalisierung von Informationen hat dieser Bereich an Bedeutung gewonnen. Die Informationen und Dienste sind demnach abhängig von der aktuellen Position des Geräts. Die mit unter populärsten Dienste in dieser Kategorie sind die Navigation und der eCall. Folgend wird auf die benötigten Komponenten für LBS, mögliche Systemarchitekturen in Bezug auf Kapitel 2.1.3 und Dienste im Automobil eingegangen.

2.2.1 Komponenten

Steininger [73] beschreibt die fünf hauptsächlich benötigten Komponenten für LBS:

- Mobiles Gerät
- Kommunikationsnetzwerk
- Positionierung
- Service Provider
- Content Provider

In Abschnitt 2.1 wurden die Komponenten Mobiles Gerät als Telematikeinheit in Kapitel 2.1.1 und das Kommunikationsnetzwerk in Kapitel 2.1.2 bereits ausführlich beschrieben.

Eine möglichst akkurate Positionierung des Geräts bzw. Nutzers ist für eine zuverlässige Anwendung unabdingbar. Dabei können je nach Verfügbarkeit verschiedene Technologien zum Einsatz kommen. Neben satellitengestützten Verfahren wie Global Positioning System (GPS), Glonass oder Galileo finden auch Methoden Anwendung, die auf der Signalstärke von Mobilfunkzellen basieren. Eine Übersicht über gängige Positionierungsverfahren für LBS wird in [74, S. 21–42] gegeben. Eine zusätzliche Indoor-Positionsbestimmung für Fahrzeuge ist

vor allem in Parkhäusern eine Alternative, um die Präzision bei schwachen Satellitensignal zu erhöhen.

In Abb. 2.1 wurden für das Telematikprotokoll bereits der SP sowie der CP eingeführt. Diese spielen im Bereich der LBS eine entscheidende Rolle. Der SP bietet den Nutzer die entsprechenden Dienste an, übernimmt die Verarbeitung der Anfrage und bezieht die erforderlichen Daten über den CP. Es wird somit ein Mehrwert für den Fahrer oder das Fahrzeug erzeugt, der gegebenenfalls gebührenpflichtig erworben wird.

2.2.2 Architekturen und Datenbeschreibung

Das NGTP zeigt, wie sich SP und CP direkt in die Architektur eingliedern. Das Open Geospatial Consortium (OGC) hat bereits 2002 den LBS Standard OpenLS spezifiziert, der die offene Service Plattform GeoMobility Server beschreibt [75]. Dieser wird in Abb. 2.2 als Beispiel zur Veranschaulichung der Prozesskette in Verbindung mit dem SP und dem CP dargestellt.

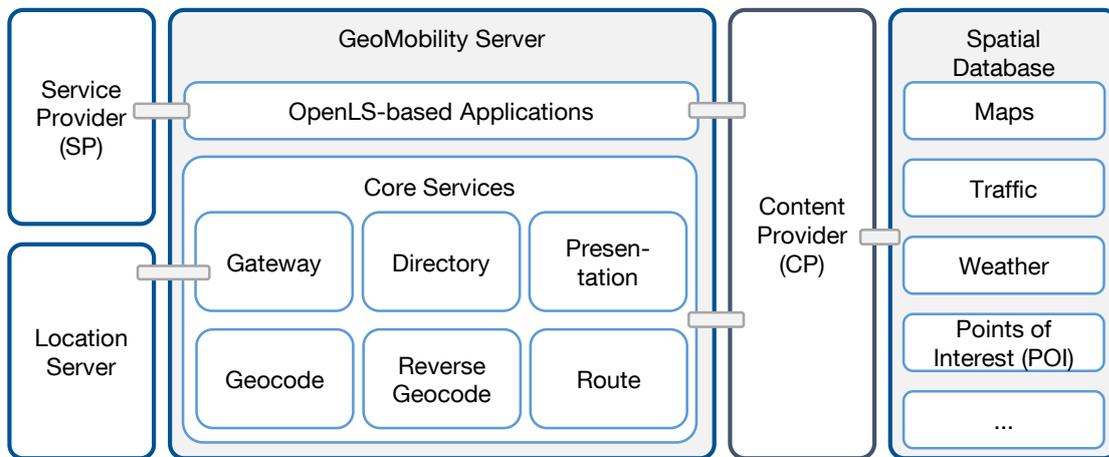


Abbildung 2.2 : Übersicht über Service und Content Provider

Der GeoMobility Server stellt sowohl grundlegende als auch nach OpenLS definierte erweiterte Dienste bereit. Zu den grundlegenden Diensten zählt das Gateway. Dieses ist die Schnittstelle zu einem Location Server, der die Position des Geräts bestimmt. Ein Verzeichnisdienst (Directory) liefert auf Anfrage ein nächstgelegenes oder spezifisches Objekt (Restaurant, Autohändler). Ein weiterer Dienst übernimmt die Darstellung der Ergebnisse. Die Georeferenzierung (Geocode) bezeichnet die Zuordnung von Geokoordinaten zu einem Objekt (Bsp. Straßename zu Koordinaten). Bei der inversen Variante werden die Geokoordinaten in ein Format durch Namensgebung überführt (Bsp. Koordinaten zu Straßename). Über einen Navigationsdienst wird die Route für einen Nutzer bestimmt. Erweiterte Dienste können auch auf grundlegende Dienste zugreifen.

Daneben greifen die Dienste über den CP auf die notwendigen Inhalte zu. Dem zugrunde liegen Karteninformationen und insbesondere im Bereich der Mobilität das Straßennetz. Es handelt sich vor allem um räumlich bezogene Daten, wie bspw. den Verkehrsfluss oder das Wetter, aber auch fahrer- oder fahrzeugspezifische Daten, die keinen räumlichen Bezug aufweisen. Diese Daten können unter anderem von Drittanbieterquellen bereitgestellt werden [75]. In Bezug auf räumliche Datenbanken ist der Bereich der GIS besonders von Bedeutung. GIS definierte Prozesse dienen zur Speicherung, Verarbeitung und Repräsentation der räumlich bezogenen Daten [76]. Einen detaillierten Überblick über GIS liefert [74].

Neben OpenLS gibt es weitere Standards und prototypische Alternativen. Unter anderem zeigt [77] einen sicherheitsbezogenen Ansatz für persönliche Daten mit einer Erweiterung nach dem OpenLS Standard auf. Ein gerätekonzentrischer Ansatz für LBS wird in [78] präsentiert. Eine Übersicht zu weiteren Standards im Bereich GIS bietet [79].

Neben der Architektur existieren vor allem Standards für die Beschreibung der Daten zum Austausch zwischen den Bereichen. Die Geography Markup Language (GML) ist eine weit verbreitete, XML-basierte Beschreibung zur Encodierung, unter anderem von Objekten und Geometrien (OGC) [80]. Anders als GML setzt das von TomTom initiierte Protokoll OpenLR auf die dynamische Referenzierung der Position [81]. Diese Beschreibung dient bspw. zum Datenaustausch zwischen Verkehrsmanagementbetrieben oder zur Kommunikation von variablen Geschwindigkeitsbegrenzungen. Das Europäische Komitee für Normung [82] versucht die unterschiedlichen Standards zur Beschreibung zu vereinheitlichen. Dabei werden sechs Ebenen definiert, die zur Beschreibung für eine multimodale, urbane Mobilität besonders von Bedeutung sind und die Objekte und Abläufe der Mobilität für statische und dynamische Vorgänge beschreiben:

- Nutzerinteraktion (u.a. Fahrzeugführung)
- Personenverkehrsdienste (u.a. Reisendeninformation, Busnummern)
- Zeitbezogene Ebene (u.a. Fahrpläne, bedarfsabhängige Routen)
- Räumliches Netzwerk (u.a. Haltestellen, zeitliche Beschränkungen)
- Infrastruktur (u.a. Straßennetzwerk, Oberflächenbeschaffenheit)
- Umwelt (u.a. Koordinatensysteme, Wetter)

2.2.3 Lokale dynamische Karte

Im Bereich der Verkehrstechnik werden die Informationen, welche die Grundlage für die Dienste darstellen, in Bezug zum Straßennetz gesetzt. Im Projekt Safespot [83] wurde eine Beschreibung für eine Local Dynamic Map (LDM) erarbeitet. Diese Art der Karte zielt auf die Aggregation von Daten für Intelligent Transportation Systems (ITS). Das Europäische Institut für Telekommunikationsnormen (ETSI) hat 2011 die LDM standardisiert [84]. Ebenso existiert der Standard ISO/TS 18750:2015 zur Festlegung eines globalen Konzeptes für lokale dynamische Karten. Die Karte wird in vier Ebenen beschrieben [84]:

- LDM 1: Statische Karte (Straßennetz)
- LDM 2: Erweiterte statische Karte (z.B. Landmarken, Verkehrsschilder)
- LDM 3: Temporäre und dynamische Informationen (z.B. Verkehrsfluss, Lichtsignalanlage (LSA))
- LDM 4: Hochdynamische Informationen (z.B. Fahrzeugposition, Geschwindigkeit)

Shimada [85] setzt die LDM für Sicherheitsfunktionen ein und evaluiert die Auswirkungen für Zugriffe auf die Karte durch eine Vielzahl von Fahrzeugen. Eiter [86] erweitert die Karte um semantische Informationen. Als Beispiel für Anwendungen auf der Karte wird ein Rotlichtverstoß und die Detektion der aktuellen Fahrspur genannt.

Um die Karteninformationen für Assistenzsysteme im Fahrzeug vergleichbar nutzen zu können, hat das Advanced Driver Assistance Systems Interface Specifications (ADASIS) Forum ein Protokoll entwickelt [87]. Im Fahrzeug dient ein elektronischer Horizont dazu, Teile des Datensatzes auf dem Fahrzeug-Bus zur Verfügung zu stellen [88]. In einer Erweiterung beschreibt Horita [89] eine fahrspurindividuelle Beschreibung speziell im Hinblick auf autonome Fahrzeuge. Burgstahler [90, 91] beschreibt wie ein offener SP für LDM unter Verwendung des ADASIS Protokolls aussehen könnte.

2.2.4 Dienste im Automobil

Steininger [73] beschreibt mögliche, allgemeine Kategorien für LBS. Die Bereiche Navigation, Informationsdienste, Notruf und Gebührenerhebung wurden bereits in Beispielen aufgegriffen. Weiter besteht die Möglichkeit der Nachverfolgung (Personen, Sendungen), der Anzeige von

Werbung oder der Umsetzung von Spielen. Die für diese Arbeit relevanten Bereiche ordnen sich für die Fahrzeugtechnik in Energieeffizienz, Sicherheit und Komfort ein, wie bereits in Kapitel 1.1 eingeführt. Neben diesen Kategorien erfolgt eine Einordnung der Dienste im Hinblick auf die Ebenen der LDM (Kapitel 2.2.3). Dabei greifen die Dienste direkt auf Attribute der digitalen Karte zu und integrieren diese in fahrzeugrelevante Funktionen. Analog zu dieser Einordnung kann der Zeitbezug wie in Abb. 1.2 berücksichtigt werden.

Eine Einordnung von typischen Diensten wird in Tab. 2.1 vorgestellt. Eine verkehrsflussabhängige Navigation [21], sowie der Bereich Ecorouting [17] lassen sich hauptsächlich in die Bereiche Energieeffizienz und Komfort einordnen. Neben den essentiellen Karteninformationen verwenden die Funktionen statische Zusatzattribute wie Verkehrsschilder. Die Funktionen generieren dynamische Kartenattribute und benötigen einen langfristigen Horizont. Zudem werden ausschließlich allgemeine Informationen verwendet. Ähnlich lassen sich Grüne-Welle-Assistenten einordnen. Zusätzlich zu den dynamischen Informationen der Ampelphasenzyklen generiert die Funktion eine Empfehlung für eine effiziente Fahrstrategie [92]. Als Beispiele für sicherheitsrelevante Funktionen können Unfallmeldungen [93], Reibwertschätzung [18] oder die Adaption von Komfortparametern durch die Berücksichtigung der Straßenqualität [94] gesehen werden. Der in Kapitel 1.1 angesprochene persönliche Assistent [19] wird hauptsächlich für einen Komfortgewinn eingesetzt. Anders als die bisher angegebenen Funktionen werden personalisierte Informationen benötigt. Auf die Funktion Trip- und Fahrprofilprädiktion wird in Kapitel 2.3.1 eingegangen.

Tabelle 2.1 : Beispiele für die Einordnung von Location Based Services

Funktion	Kategorie			Local Dynamic Map				Horizont			Nutzung	
	Energieeffizienz	Sicherheit	Komfort	LDM 1	LDM 2	LDM 3	LDM 4	Langfristig	Mittelfristig	Kurzfristig	Allgemein	Personalisiert
Verkehrsfluss, Ecorouting	x	-	o	n	n	g	-	x	o	-	x	-
LSA Grüne Welle Assistent	x	-	-	n	n	g	g	-	-	x	x	-
Unfallmeldungen	-	x	-	n	n	g	-	-	x	x	x	-
Reibwertschätzung	-	x	-	n	n	g	-	-	x	-	x	-
Straßenqualität	-	x	x	n	g	-	-	x	-	-	x	-
Persönlicher Assistent	-	-	x	n	g	n	-	x	-	-	-	x
Trip- und Fahrprofilprädiktion	o	o	o	n	n	g	g	o	x	o	x	o

x = explizit, o = optional, n = nutzt, g = generiert

2.3 Ableitung der Themenstellung

In diesem Abschnitt wird zuerst die für diese Arbeit relevante Forschungsfrage aufgestellt. Anschließend wird der Lösungsansatz kurz vorgestellt und dabei auf die Einbettung des Themas in die Grundlagen der Vernetzung von Fahrzeugen eingegangen. Abschließend wird das allgemeine Vorgehen aufgezeigt.

2.3.1 Forschungsfrage

In der Motivation für diese Arbeit wurde herausgearbeitet, dass es möglich ist, einen Mehrwert durch die Vernetzung von Fahrzeugen für den Fahrer und das Fahrzeug zu erzielen. Dabei hat sich herausgestellt, dass vor allem die Route und die Bewegungszustände des Fahrzeugs betrachtet werden müssen. Die Ziele für diese Arbeit wurden in Kapitel 1.2 definiert und erfordern eben diese für einen mittelfristigen Horizont zu präzisieren. In den Kapiteln 2.1 und 2.2 wurde herausgearbeitet, dass bereits geeignete Standards für Architekturen im Fahrzeug und im Back-

End, sowie definierte Übertragungsprotokolle vorhanden sind, um derartige standortbezogene Dienste umzusetzen. Die Forschungsfragen können daher ausschließlich auf die Funktionalität hin ausgerichtet werden:

- Wie kann die geplante Fahrtroute des Fahrzeugs ohne Angabe eines Fahrtziels prädiziert werden?
- Wie kann das Fahrprofil des Fahrzeugs möglichst detailliert für einen mittelfristigen Horizont prädiziert werden?
- Wie kann diese Funktion für unterschiedliche Ziele zur Optimierung von Fahrzeugfunktionen eingesetzt werden?
- Welcher quantifizierbare Nutzen ergibt sich aus einer derartigen Anwendung?

Diese Fragen werden zudem im Rahmen der Motivation unter dem Zielaspekt der Vernetzung von Fahrzeugen weiter eingegrenzt:

- Wie können diese Ziele unter der Verwendung einer Vielzahl gesammelter Fahrprofile erreicht werden?
- Inwiefern können diese Ziele anonymisiert bzw. personalisiert erreicht werden?

Alle Funktionen aus Tab. 2.1 zielen auf eine konkrete Einordnung nach den Gesichtspunkten Energieeffizienz, Sicherheit und Komfort ab und können nur zum Teil für übergreifende Bereiche verwendet werden. Die Kombination aus Trip- und Fahrprofilprädiktion ist nicht explizit einzelnen Kategorien zuzuordnen, sondern kann als Grundlage für darauf aufbauende Funktionen verschiedene Ziele verfolgen. Damit wird eine Zusatzfunktionalität geschaffen, die vielseitig eingesetzt werden kann und die Bewegungsgrößen des Fahrzeugs voraussagt. Die vorgestellten Dienste benötigen zudem für den Einsatz im Fahrzeug das Fahrtziel. In dieser Arbeit soll eine Steigerung der Energieeffizienz auch ohne diese Angabe erzielt werden.

Neben der Generierung von dynamischen Informationen werden auch Zeitreiheninformationen generiert. Die Funktion nutzt somit Daten der dynamischen Karte und generiert dabei sowohl Daten für Ebene LDM3 (Trip) als auch LDM4 (Fahrprofil). Anders als die vorgestellten Funktionen werden hierbei detaillierte Zeitreiheninformationen für die Fahrzeugbewegung erzeugt. Während bspw. für den Verkehrsfluss, im Bereich des Ecorouting oder für den Straßenreibwert aggregierte Kennwerte berechnet werden, kann über diese Funktion die Zeitreiheninformation für die voraussichtliche Geschwindigkeit, Beschleunigung und Gierrate vielseitig eingesetzt werden.

Die Basis für die Funktionalität bilden anonymisierte Daten von mehreren Fahrern. Aufbauend auf den allgemeinen Informationen können auch personalisierte Daten verwendet werden und die Funktion für Nutzer oder Nutzergruppen zielgerichtet erweitert werden.

2.3.2 Lösungsansatz

Im folgenden wird zusammenhängend der Lösungsansatz beschrieben. Dafür wird zuerst die Datenbasis erläutert. Anschließend wird auf die Trip-Prädiktion eingegangen. Der Ansatz für die Fahrprofilprädiktion wird erläutert und abschließend die Funktionsentwicklung zur Nutzung der gewonnenen Erkenntnisse vorgestellt.

Datenbasis

Diese besteht aus Fahrten, durchgeführt von unterschiedlichen Fahrern und Fahrzeugen. Die Fahrten werden beschrieben durch die GPS-Position, die Fahrgeschwindigkeit, sowie Beschleunigungen und Drehraten in alle Raumrichtungen. Zudem ist der Zeitpunkt der Aufzeichnung bekannt. Die Daten liegen als Zeitreiheninformationen in unterschiedlichen Abstraten zu Grunde. Um eine detaillierte Beschreibung des Fahrverhaltens für die Prädiktion zu ermöglichen, werden die Daten anders als in [28] und [95] nicht auf Kennwerte reduziert. Stattdessen spielt der Verlauf der Bewegungsgröße bei der Untersuchung eine entscheidende Rolle. Abb.

2.3 verdeutlicht die Datengrundlage für zwei exemplarische Fahrten. In diesem Beispiel wird ersichtlich, dass sich die Fahrten auf der Route überschneiden. Daneben ist ein ähnliches Geschwindigkeitsprofil der beiden Fahrten auf Teilen der Route erkennbar. Die Kartenattribute ergänzen die Zeitreiheninformationen.

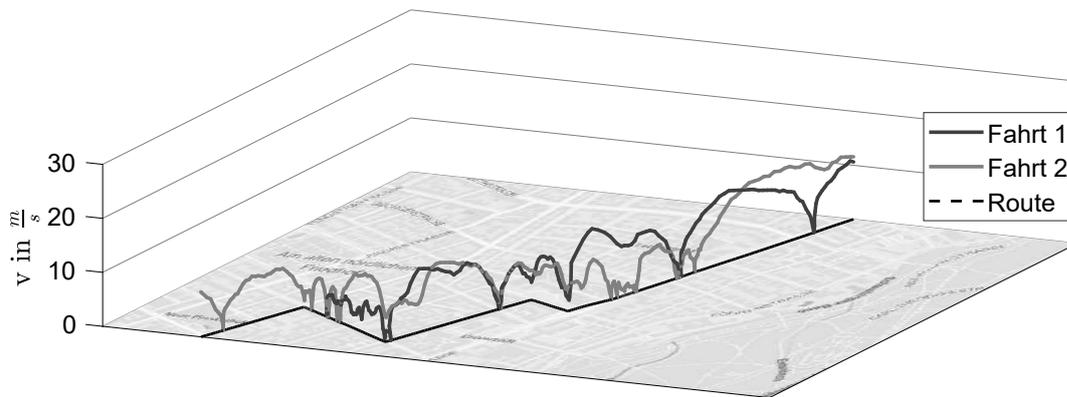


Abbildung 2.3 : Beispielhafter Verlauf von zwei Geschwindigkeitsprofilen auf einer Route (Karte: OSM )

Trip-Prädiktion

Das Ziel der Trip-Prädiktion ist es, Muster in der Bewegung, bzw. in der Position der Fahrzeuge zu erkennen und diese geeignet vorherzusagen. Der Freiheitsgrad der Bewegung ist im Grunde durch das Verkehrswegenetz (LDM1) beschränkt. Daher werden die GPS-Positionen auf Straßensegmente reduziert. Dies erfolgt durch ein Mapmatching-Verfahren, das in Kapitel 3.2 näher erläutert wird. Mit Hilfe von Data-Mining-Verfahren und Methoden des Maschinellen Lernens [96, S. 191–264] werden Muster in den reduzierten Datensätzen erkannt. Dabei werden zuerst alle Fahrer gleichzeitig anonymisiert betrachtet.

Die nächsten Streckensegmente werden anschließend prädiziert. Dabei dient zudem der bereits zurückgelegte Weg als Eingangsgröße. Weiterhin können zeitliche Informationen oder persönliche Präferenzen mit berücksichtigt werden (LDM2). Der Vorausschauhorizont ist variabel gestaltbar und wird je nach Zielanwendung angepasst. Neben dem wahrscheinlichsten Pfad ist es außerdem möglich, weitere Alternativen und deren Wahrscheinlichkeit anzugeben [97]. Das Ergebnis dieser Funktion ist ein wahrscheinlicher Pfad des Fahrzeugs auf den vorausliegenden Straßensegmenten, auf Basis real gefahrener Fahrten von mehreren Fahrern, für einen beliebig langen Horizont. Die Methodik wird in Kapitel 4 erläutert.

Fahrprofilprädiktion

Bei der Prädiktion des Fahrprofils ist das Ziel, Muster in dem Bewegungsverhalten der Fahrzeuge zu erkennen und diese vorherzusagen. Die Bewegung wird durch die Fahrgeschwindigkeit, Beschleunigungen und Drehraten beschrieben. Als Eingang dienen Zeitreiheninformationen, die dem Straßennetz zugeordnet werden können und zusätzlich der vorausliegende Pfad der Trip-Prädiktion. Falls alternativ das Ziel für die bevorstehende Fahrt bekannt ist, wird die vorab kalkulierte Route verwendet.

Die Problemstellung eignet sich analog zur Trip-Prädiktion zur Lösung mit Data-Mining Verfahren und Methoden des Maschinellen Lernens. Es werden Ähnlichkeitsmaße zwischen den aufgezeichneten, ortsbezogenen Fahrprofilen identifiziert. Anschließend werden diese in Cluster zusammengefasst und jeweils durch ein, für dieses Cluster repräsentatives Geschwindigkeitsprofil, abgebildet. Dadurch können die Zeitreiheninformationen als Zustände interpretiert werden. Die Betrachtung erfolgt individuell für jedes Streckensegment. Diese Unterscheidung ist erforderlich, weil bspw. das reale Fahrverhalten von dem erlaubten stark abweichen kann und diese Information in einer allgemeingültigen Betrachtung reduziert auf Kartenattribute (z.B. erlaubte Höchstgeschwindigkeit) verloren gehen würde [97]. Nach der Modellierung erfolgt die Prädiktion der Zustände (repräsentatives Fahrprofil) für die vorausliegende Strecke. Die Prädiktion ist dabei abhängig von dem zurückgelegten Weg und dem Fahrprofil. Zudem werden

Informationen wie der Verkehrsfluss (LDM3) mit berücksichtigt. Der Prädiktionshorizont ist, wie bei der Trip-Prädiktion, von variabler Länge und wird nur über den vorab identifizierten Pfad begrenzt. Als Ergebnis wird in dieser Funktion ein wahrscheinliches Geschwindigkeitsprofil für die vorausliegenden Streckenabschnitte erstellt.

Zusätzlich zur Fahrgeschwindigkeit können weitere Bewegungsgrößen mit betrachtet werden. Dabei dienen die ermittelten Cluster in den Geschwindigkeitsprofilen einer Einordnung von Beschleunigungen und Drehraten über den zeitlichen Verlauf. Es wird dabei ein Konfidenzintervall erstellt, das bei gleicher Fahrstrecke und ähnlichen Geschwindigkeitsverläufen aus anderen Fahrten eine Abschätzung der Gierrate für die vorausliegende Strecke ermöglicht. Gerade in Kurvenfahrten kann diese Information für Anwendungen zur Bewertung der Stabilität des Fahrzeugs herangezogen werden.

Anwendung im Fahrzeug

Für eine Anwendung im Fahrzeug können die Funktionen sowohl einzeln, als auch in Kombination verwendet werden. Ziel ist es, die Prädiktion zur Optimierung des Fahrzeugs einzusetzen. Als Beispiel wird eine Funktion mit einer datengetriebenen Betriebsstrategie für ein PHEV in Kapitel 6 beschrieben. Die Funktion ist dabei nicht notwendigerweise im Fahrzeug verankert, sondern kann auch im Back-End ausgeführt werden. Ebenso ist eine hybride Form möglich. Der Nutzen, der sich aus der Vorausschau ergibt, wird in einer Simulation bestimmt.

2.3.3 Vorgehen

In dieser Arbeit wird ein datengetriebener Ansatz verfolgt. Die Thematik ordnet sich daher in den Bereich der Wissensgenerierung in Datenbanken ein. Fayyad [98] hat ein weitreichendes, allgemeines Vorgehensmodell für diese Aufgabe entworfen. Dieses verdeutlicht, dass mehrere Schritte vor dem eigentlichen Data-Mining-Verfahren notwendig sind, um das Ziel eines interpretierbaren Wissens aus den Daten zu erreichen. Das in dieser Arbeit angewandte Vorgehen in Anlehnung an [98] wird in Abb. 2.4 dargestellt.

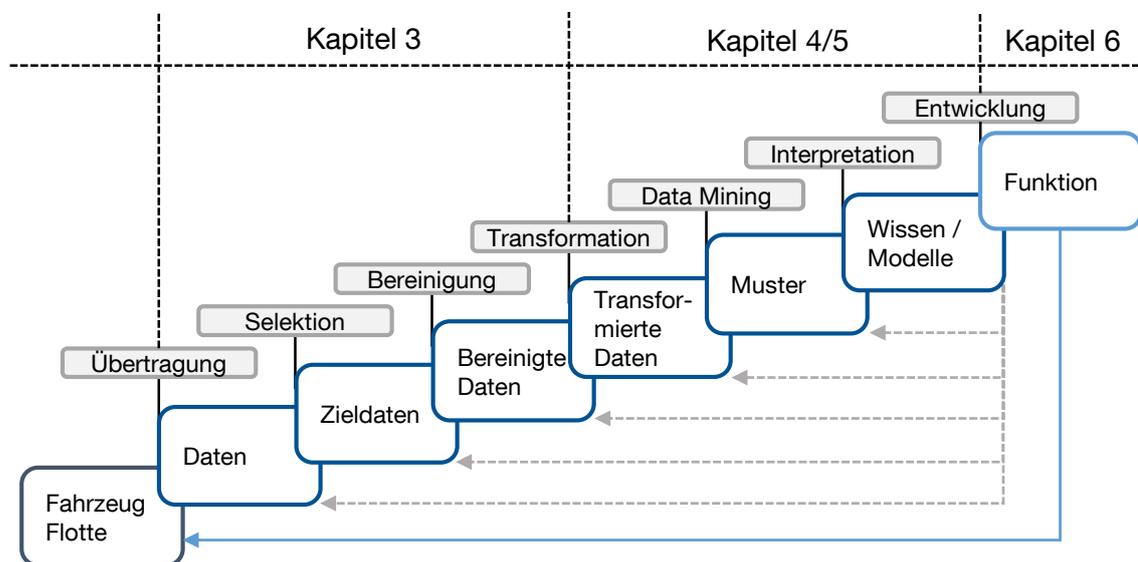


Abbildung 2.4 : Vorgehensmodell mit Wissensgenerierung in Datenbanken aufbauend auf Fayyad [98]

Der erste Schritt beschreibt den Aufbau der Datenbasis. In Flottenversuchen wurde eine breite Datenbasis durch die übertragenen Fahrprofile erstellt. Die Datensammlung war nicht Bestandteil dieser Arbeit. Eine Übersicht über die Flottenversuche und die aufgezeichneten Daten wird in Kapitel 3.1 gegeben. Anschließend erfolgt die Selektion der Daten in einen Zieldatensatz. Wie in Kapitel 2.3.2 beschrieben, werden die Untersuchungen abhängig von dem Kartenmaterial durchgeführt. Die Selektion erfolgt hauptsächlich anhand der definierten Straßensegmente aus Kapitel 3.2. Bei der Datenerfassung und Vorverarbeitung können zum einen Fehler auftreten, zum anderen kann es bspw. durch Ausnahmesituationen im Straßenverkehr zu Ausreißern in

der Datenbasis kommen. Deshalb wird im darauffolgenden Schritt die Bereinigung der Daten durchgeführt. Die Daten werden auf eine aussagekräftige Basis reduziert. Dabei muss auch eine Entscheidung für den Umgang mit fehlenden Daten getroffen werden.

Anschließend wird eine Transformation der Daten durchgeführt. Die Daten werden dadurch auf Kernelemente reduziert. Dieser Vorgang unterscheidet sich für die Trip- (Kapitel 4) und Fahrprofilprädiktion (Kapitel 5) wesentlich, weil unterschiedliche Ziele bei der Betrachtung der Daten verfolgt werden. Die Daten können um Zusatzinformationen ergänzt werden. Auf den transformierten Daten werden je nach Zielsetzung (Mustererkennung, Klassifikation, Regression, etc.) geeignete Data-Mining-Verfahren angewendet. Darauf aufbauend werden die Ergebnisse interpretiert und eine Entscheidung getroffen, ob die angewandten Verfahren dafür geeignet sind, die Aussagen zu erreichen. Abhängig davon wird ab diesem Punkt der Vorgang iterativ verbessert. Die Schritte im Prozess können angepasst werden und bspw. die Bereinigung der Daten angepasst oder anderweitige Datenreduktionsverfahren durchgeführt werden.

Abschließend werden die interpretierbaren Ergebnisse bei der Entwicklung einer Anwendung in Kapitel 6 in Dienste überführt. Die Funktionen werden geeignet erweitert, um im Fahrzeug anwendbar zu sein. Ziel dieser Arbeit ist es, die Funktionen durch Simulationen zu überprüfen.

Neben dem Vorgehensmodell nach Fayyad existieren weitere Ansätze gerade im Hinblick auf die Verwendung von Streaming-Daten [99]. Die kontinuierliche Verarbeitung von Daten kann sich besonders für große Datenmengen und einen gleichzeitigen Echtzeiteinsatz eignen. Da die verwendete Datenbasis dafür allerdings nicht geeignet ist und ein Echtzeiteinsatz nicht Fokus dieser Arbeit ist, wird dieses Modell nicht betrachtet.

Folgend wird auf die Entwicklungsumgebung eingegangen. Diese ist schematisch in Abb. 2.5 dargestellt. In dieser Arbeit werden reale Messdaten verwendet. Dabei handelt es sich ausschließlich um historische Daten. Diese werden in einer relationalen Datenbank hinterlegt. Die Erstellung der datengetriebenen Modelle erfolgt serverseitig mit den historischen Daten. Dies wird mit einem Hochleistungsrechner (HPC) bewerkstelligt. Die Bewertung der Genauigkeit der Funktion wird mit unabhängigen Testdaten durchgeführt, die nicht bei der Erstellung der Modelle verwendet werden. Die Funktionen sind Back-End-Algorithmen, die nicht im Fahrzeug berechnet werden. Des Weiteren wird in Kapitel 6 eine Anwendung vorgestellt, die auf der Trip- und Fahrprofilprädiktion basiert. Für eine Echtzeitanwendung im Fahrzeug ist eine Kommunikation mit dem Back-End erforderlich. In dieser Arbeit erfolgt keine Integration der Anwendung ins Fahrzeug. Die Bewertung der Anwendung wird in einer Simulationsumgebung mit realen Messdaten auf einer Workstation durchgeführt. Da die Kommunikation mit dem Back-End kein zentraler Bestandteil dieser Arbeit ist, wird der Applikations-Server mit der Workstation für die Simulation zusammengefasst.

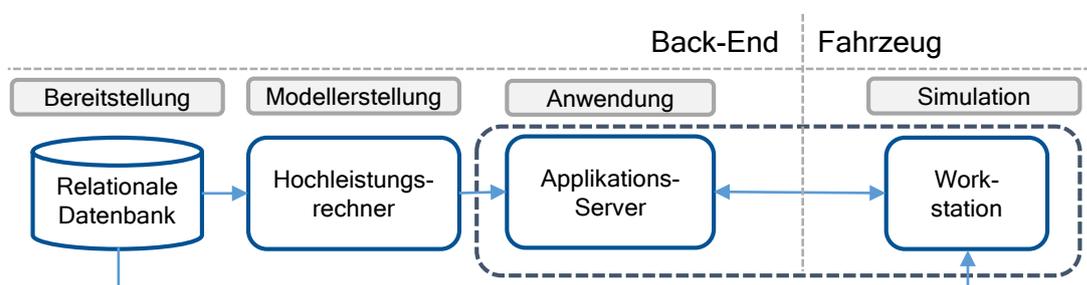


Abbildung 2.5 : Entwicklungsumgebung

3 Aufbereitung der Datenbasis

In den folgendem Kapitel wird die Aufbereitung der Datenbasis näher erläutert. Es werden die Datenquellen, sowie die Datenverarbeitung veranschaulicht. Anschließend wird das verwendete Kartenmaterial genauer dargestellt. In Kapitel 3.2 wird auf die notwendigen Verarbeitungsschritte eingegangen, um auf den aggregierten Zieldatensatz für die Untersuchung zu gelangen.

3.1 Datenerfassung und Selektion

Zuerst werden mit den aufgezeichneten Flottenversuchen die wichtigsten Eingangsdaten erläutert. Anschließend werden die zu Grunde liegenden Karteninformationen beleuchtet. Das Ziel dieser Arbeit ist es, einen Mehrwert aus Mobilitätsdaten zu generieren. Deswegen werden reale Fahrdaten für die Untersuchung herangezogen. Da die Aufzeichnung der Daten mit einer Smartphone-Applikation erfolgte, ist eine ausführliche Bewertung der Genauigkeit und der Fehler erforderlich, um festzustellen, ob diese für die Entwicklung der Funktionen verwendet werden können.

3.1.1 Flottenversuche

Im Rahmen von mehreren geförderten Forschungsvorhaben wurden am Lehrstuhl für Fahrzeugtechnik Flottenversuche durchgeführt. Eine Übersicht wird in [100] gegeben. Für diese Arbeit werden die Daten mit Privatfahrern aus den Versuchen eFlott [101], eMUC [102] und sun2car [103] verwendet. Des Weiteren wurden mehrere Fahrzeuge der Flotte eines ausgewählten Handwerksunternehmens betrachtet. Einzelne Fahrzeuge von Handwerksbetrieben wurden auch im Projekt Virtuelle Elektromobilität für den Taxi- und Handwerksverkehr München (VEM) [104] untersucht. In VEM wurde zwischen 2012 und 2015 das Bewegungsverhalten von 90 Taxis in München aufgezeichnet. In Abb. 3.1 werden die für diese Arbeit relevanten Flottenversuche zusammengefasst. Zudem ist das Aufzeichnungsgebiet aller Fahrzeuge dargestellt. Insgesamt wird auf rund 5 Mio. km an Fahrdaten zurückgegriffen. Die gesamten Straßenkilometer von Bayern werden in der Datenbank zu 18,7 % abgedeckt, die häufiger befahrenen Straßen von Oberbayern zu 40,3 %. Die Abdeckung des Münchner Netzes beträgt 92,4 %. Besonders der Flottenversuch VEM hat nahezu das gesamte Straßennetz Münchens während der Laufzeit befahren.

Datenbasis

Flottentest	Zeitraum	Fahrzeuge
eFlott	10.2010-08.2012	20 BEV, 1 ICE
eMUC	07.2012-02.2014	25 BEV
VEM	10.2012-12.2015	115 ICE, 3 BEV
sun2car	09.2013-09.2014	10 BEV, 10 ICE
Handwerk	04.2016-06.2016	44 ICE

# Fahrten	Distanz	Fahrzeuge insg.
374 782	4 841 201 km	228

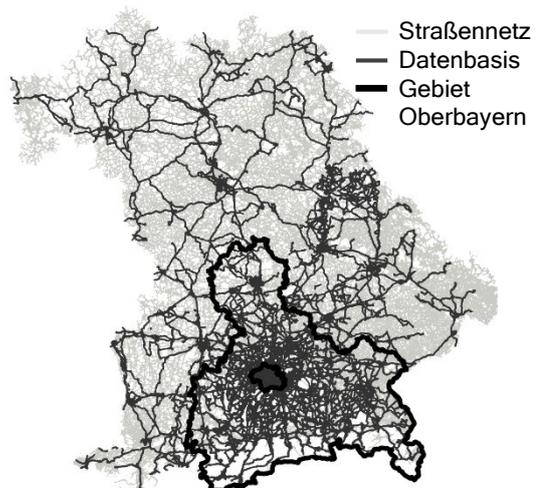


Abbildung 3.1 : Datenbasis der relevanten Flottenversuche nach [100] (Karte: OSM © ⓘ ⓘ)

Die Aufzeichnung erfolgt mit Hilfe einer Smartphone-Applikation. Dabei wurden GPS-Position und Bewegung der Geräte im Fahrzeug erfasst und über Mobilfunk zu einem Server im Back-End

übermittelt [100]. Die Aufzeichnungsrate der Positionsdaten beträgt 1 Hz. Die Geschwindigkeit wird über die GPS-Position ermittelt und liegt in 1 Hz vor. Die Beschleunigungen und Drehraten werden mit bis zu 25 Hz (in alle Raumrichtungen) abgetastet. Ein Web-Server empfängt die Daten und hinterlegt diese in einer relationalen PostgreSQL Geodatenbank [105]. Diese verfügt über eine PostGIS Erweiterung [106], die unter anderem georeferenzierte Abfragen für GIS ermöglicht. Die Daten werden dabei zu einzelnen Fahrten kombiniert, die über eine Nutzeridentifikationsnummer und den Zeitbereich nachvollziehbar zugeordnet werden können.

Im Rahmen der Flottenversuche wurden mehrere Geräte von verschiedenen Herstellern zur Datenerfassung verwendet. Die verbauten Sensoren unterscheiden sich daher. Eine einheitliche Aussage über die Güte und Auflösung der Sensoren ist nicht möglich. Zudem ist die Messdatenerfassung mit Smartphones im Vergleich zu internen Fahrzeugdaten bspw. über OBD fehlerbehaftet. Ritzer [107] betrachtet in seiner Arbeit die auftretenden Fehler und quantifiziert diese. Der quadratische Mittelwert des horizontalen Positionsfehlers liegt für verschiedene Smartphones zwischen 4,06 m und 9,12 m. Der Geschwindigkeitsfehler ist im Bereich zwischen $0,31 \frac{m}{s}$ und $0,49 \frac{m}{s}$. Die Lage der Smartphones im Fahrzeug kann durch die Bedienung während der Fahrt (unerlaubt oder durch den Beifahrer) die aufgezeichneten Daten verfälschen. Insbesondere sind die Drehraten in das Fahrzeugkoordinatensystem zu überführen [107]. Die Kalibrierung ist ebenfalls zu beachten, weil andernfalls systematische Messfehlern auftreten können. Lee [108] ermittelt für die mittlere Gierrate und Varianz der Längs- und Querschleunigung keine signifikante Unterschiede zwischen dem Fahrzeug und mobilen Geräten.

Um einen Eindruck über die Güte der Messungen mit Hilfe der Smartphones zu bekommen, wird im folgenden Abschnitt ein Vergleich von Geschwindigkeit, Gierrate und Querschleunigung mit der Messtechnikplattform RT2500 [109] des Unternehmens Oxford Technical Solutions dargestellt. Ein Testdatensatz von rund 250 km wurde im Rahmen zur Erstellung querdynamischer Fahrzyklen von Schultze [110] erhoben.

Geschwindigkeit

Ein Ausschnitt des aufgezeichneten Geschwindigkeitsverlaufs wird in Abb. 3.2 dargestellt. Die Auflösung beträgt $0,25 \frac{m}{s}$. Vereinzelt treten größere Ausreißer auf. Negative Geschwindigkeiten in Folge von Rückwärtsfahrten können nicht dargestellt werden. Zudem wird die Häufigkeitsverteilung der Messabweichung bezogen auf die Referenz dargestellt. Der quadratische Mittelwert des Geschwindigkeitsfehlers über alle Messdaten beträgt $0,53 \frac{m}{s}$. In Folge von Signalabbrüchen oder bspw. Interferenzen [74, S. 26–27] können größere Abweichungen zur Referenz auftreten. Falls nach der Definition für Ausreißer bei Boxplots, über das 1,5-fache des Interquartilsabstandes [111, S. 59–64], diese nicht berücksichtigt werden, beträgt der Fehler $0,2 \frac{m}{s}$ und liegt damit im Bereich der Auflösung.

Gierrate

Der Verlauf der Gierrate wird, wie die Geschwindigkeit, qualitativ gut wiedergegeben. Die Auflösung ist gerätespezifisch. Im Gegensatz zur Berechnung der Geschwindigkeit ist die Einbaulage des Smartphones im Fahrzeug von Bedeutung. Daher können Offsets im Vergleich zur Messtechnik auftreten. Ritzer [107] beschreibt, neben dem Bias, Skalenfaktorfehler, Achsenfehlstellungen und Rauschen als weitere mögliche Fehler, die zu einer Abweichung der Messung führen können. Der quadratische Mittelwert des Fehlers der Gierrate beträgt $0,63 \frac{^\circ}{s}$, unter Ausschluss der Ausreißer liegt dieser bei rund $0,31 \frac{^\circ}{s}$. In Abb. 3.2 wird ein Ausschnitt der Messung dargestellt.

Querschleunigung

Ähnlich wie die Gierrate ist die mit dem Smartphone ermittelte Querschleunigung abhängig von der Einbaulage und fehlerbehaftet. Die Auflösung ist gerätespezifisch. Der quadratische Mittelwert des Fehlers der Querschleunigung beträgt $0,54 \frac{m}{s^2}$, unter Ausschluss der Ausreißer liegt dieser bei rund $0,35 \frac{m}{s^2}$. In Abb. 3.2 wird ein Ausschnitt der Messung dargestellt.

Die weiteren notwendigen Vorverarbeitungsschritte werden in Kapitel 3.2 vorgestellt.

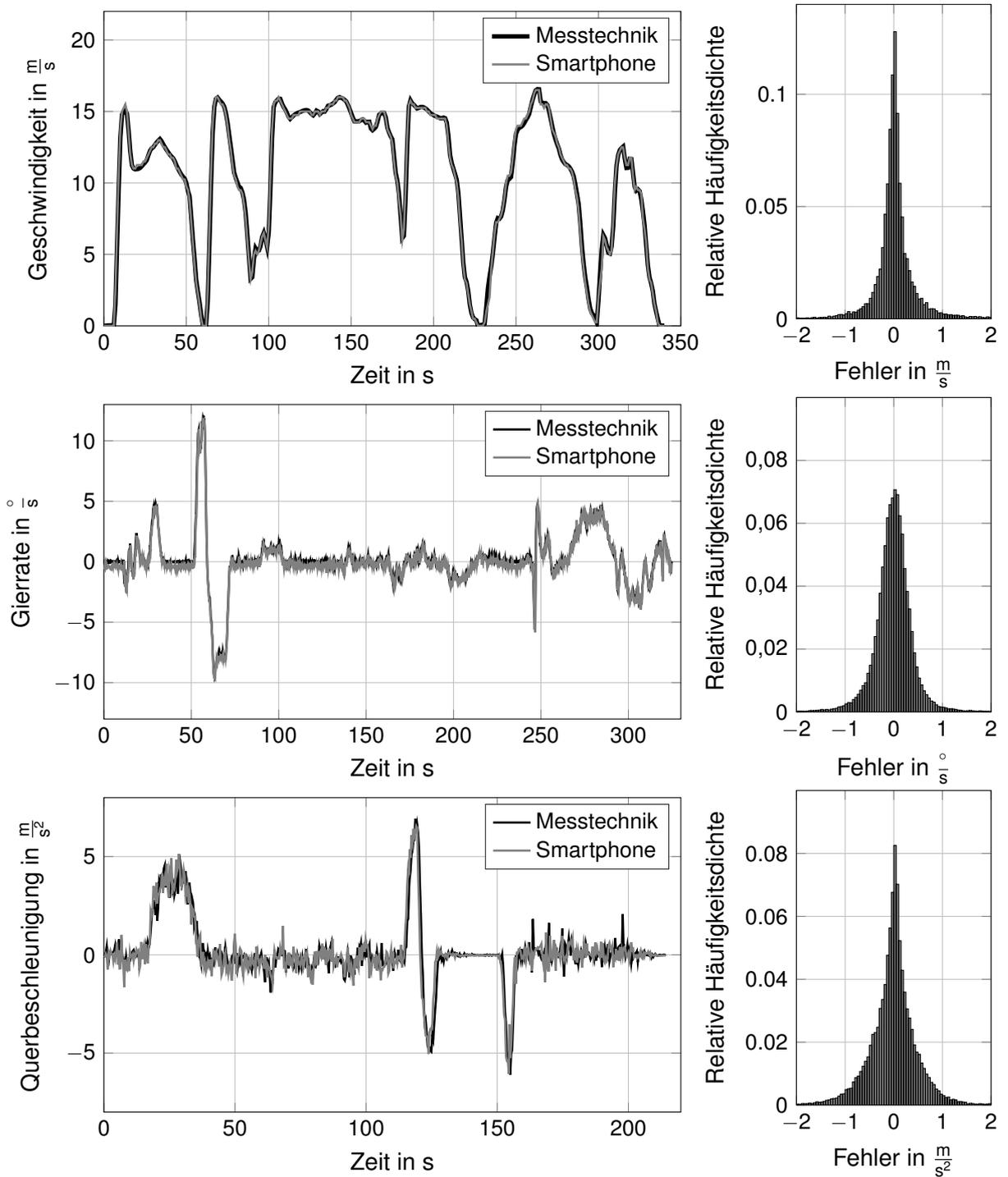


Abbildung 3.2 : Vergleich der Zeitreihen und Fehlerhistogramme

3.1.2 Kartendaten

Das verwendete Kartenmaterial entstammt OpenStreetMap (OSM) [112], das unter der Open Data Commons Open Database Lizenz frei verfügbar ist. Es ist freigestellt, die Daten zu kopieren, weiterzugeben, zu übermitteln sowie anzupassen. Es handelt sich hierbei um eine Open Source Plattform, die eine kostenlose Nutzung des Kartenmaterials bietet, das im Zuge eines Crowdsourcing Ansatzes von Nutzern gesammelt und bereitgestellt wird.

Die Karte wird durch einen Graphen aus Knoten und Kanten (Straßensegmente) beschrieben. Das OSM-Kartenmaterial ist ohne vorherige Überarbeitung nicht routing-fähig. Der Konverter osm2po [113] bringt die Daten in das gewünschte Format, bei dem sich die Knoten nur an den Kantenenden befinden. Allerdings werden mit dem Verfahren auch Straßensegmente eingefügt, die durch auffallende Knoten begrenzt werden. Dies ist bspw. an Brücken, Orts- oder Straßenschildern und sich ändernden Geschwindigkeitsbegrenzungen der Fall. Folglich wird eine Vielzahl von kurzen Straßensegmenten erzeugt. Daraus entstehen mehrere Nachteile: Zum einen wird die Position des Fahrzeugs nur jede Sekunde erfasst, dadurch können bei kurzen Straßensegmenten nur wenige bis gar keine GPS-Punkte dem Abschnitt zugeordnet werden. Das Fahrprofil wird im Folgenden auf Straßenabschnitte bezogen und wird daher über wenige Datenpunkte auf dem Abschnitt nur schlecht beschrieben. Zum anderen können typische Attribute wie die nach Ericsson [28] auf längeren Abschnitten besser bestimmt werden [95]. Für die Fragestellung ist eine Definition für einen Straßenabschnitt als Segment zwischen zwei Kreuzungspunkten ausreichend. Das Kartenmaterial wird dahingehend überarbeitet und kürzere Straßenabschnitte, die nicht an Kreuzungen ausgerichtet sind, werden vereint.

Nach der Überarbeitung stehen die in Tab. 3.1 beschriebenen Attribute für jedes Straßensegment zu Verfügung. Neben der Identifikationsnummer sind die Straßenknoten von besonderer Bedeutung, weil über diese die Definition erfolgt und angrenzende Abschnitte bestimmt werden können. Zu den Knoten werden zusätzlich die GPS-Koordinaten mit angegeben. Die Länge des Segments ist ebenso definiert wie die erlaubte Höchstgeschwindigkeit. Letztere sind in OSM nicht immer korrekt angegeben [95]. Die Routing-Verfahren ermitteln anhand von Kostenattributen den optimalen Pfad. Die angegebenen Kosten bestimmen sich als die benötigte Fahrzeit über die Länge des Segments und die erlaubte Höchstgeschwindigkeit. Die Kosten für die entgegen gerichtete Fahrtrichtung ist daher identisch, außer es handelt sich um ein Segment, dass nur in eine Richtung befahren werden kann. An Einbahnstraßen werden diese Kosten auf einen konstanten Extremwert gesetzt. Dies führt zu extremen Kosten im Routing-Verfahren und infolgedessen zu einer Vermeidung der Wahl des Segments bei der Routenfindung. Zudem sind weitere Attribute wie die ursprüngliche OSM-ID, der Straßename und die erlaubten Verkehrsmittel gegeben.

Tabelle 3.1 : Übersicht zu den Kartenattributen

	Name	Beschreibung
1	id	Identifikationsnummer des Straßensegments
2	osm_id	Ursprüngliche OSM-Identifikationsnummer
3	osm_name	Straßename
4	osm_source_id	Startstraßenknoten
5	osm_target_id	Zielstraßenknoten
6	clazz	Straßenklasse nach Tab. 3.2
7	flags	Freigegebenes Verkehrsmittel (Fahrzeug, Fahrrad, Fußgänger)
8	km	Länge des Segments in km
9	kmh	Erlaubte Höchstgeschwindigkeit in $\frac{\text{km}}{\text{h}}$
10	cost	Routingkosten, berechnet aus 8 und 9
11	reverse_cost	Routingkosten für die entgegen gerichtete Fahrtrichtung
12	x1, y1, x2, y2	GPS-Koordinaten der Straßenknoten

Die Kartendaten werden um Höheninformationen in Form eines Digitalen Geländemodells (DGM) ergänzt. Dies ist vor allem für die energetischen Betrachtungen in Kapitel 6 erforderlich. In Abb.

3.3 werden verschiedene Datenquellen miteinander verglichen. Die Daten des DGM25 stammen von der Bayerischen Vermessungsverwaltung und sind für Hochschulen in einem 25 m Raster kostenfrei verfügbar [114]. Das DGM1 mit einem 1 m Raster dient als Referenz. Die Daten werden in Befliegungen mit einem Laserscanning Verfahren erhoben. Die Genauigkeit des DGM1 beträgt $\pm 0,2$ m, die des DGM25 $\pm 2-3$ m [115]. Als zusätzlicher Vergleich werden die häufig verwendeten Shuttle Radar Topography Mission (SRTM) Daten dargestellt [116]. Diese liegen im 30 m Raster mit einer Genauigkeit von ± 6 m vor. Die DGM25 Daten zeigen im Vergleich zu den SRTM Daten eine bessere Übereinstimmung mit der Referenz. Allerdings können vereinzelt Ausreißer auftreten. Daher werden diese vorab mit einer lokal gewichteten Regressionsglättung [117] überarbeitet.

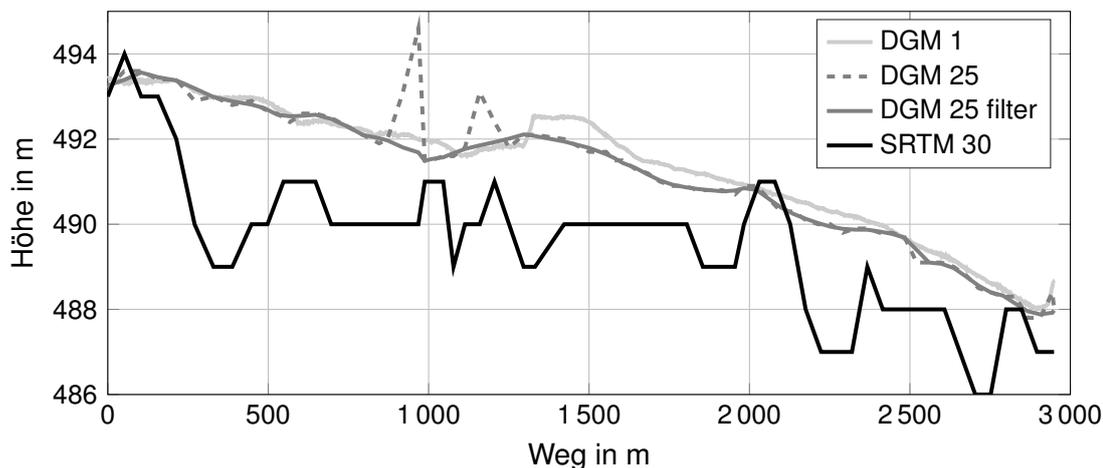


Abbildung 3.3 : Vergleich von Datenquellen für Höheninformationen

3.2 Vorverarbeitung und Aggregation

Im nächsten Abschnitt wird die Zuordnung der Flottendaten zu den Kartendaten im Kapitel 3.2.1 betrachtet. Anschließend werden weitere Metadaten, die direkt aus den Daten bestimmt werden können, betrachtet. Abschließend werden die Daten mit Verkehrsflussinformationen angereichert.

3.2.1 Mapmatching

Um die aufgezeichneten Daten der Smartphones der entsprechenden Straße eindeutig zuzuweisen zu können, muss die Information des befahrenen Streckenabschnitts zu den Sensordaten hinzugefügt werden. Die Rohdaten müssen um die zugehörige Identifikation der Straßensegmente ergänzt werden. Dieser Vorgang wird als Mapmatching bezeichnet und ermöglicht folglich eine straßensegment-, straßentyp- oder routenspezifische Auswertung der Daten [118, S. 59–63],[119].

Der in [22] beschriebene Algorithmus auf Basis des Verfahrens von [120] eignet sich vor allem für GPS-Signale mit niedriger Abtastrate. In dem Offline-Verfahren wird die gesamte zusammenhängende Fahrt betrachtet. Die aufgezeichneten Datenpunkte können zu 97 % zugeordnet werden [22]. Das Verfahren wird zu einem zweistufigen Prozess erweitert [100]. In einem ersten Schritt wird die Zahl der Datenpunkte reduziert, um die in Frage kommenden Straßensegmente zu identifizieren. In einem zweiten Schritt werden alle weiteren Datenpunkte der Route zugeordnet. Dies reduziert die Rechenzeit. Zudem werden zusätzliche Punkte im Datensatz an befahrenen Straßensegmenten erzeugt, die keine Messpunkte aufweisen. Dies tritt bspw. bei besonders kurzen Segmenten ein oder in Tunneln, falls kein GPS-Empfang vorliegt. Die zusätzlichen Datenpunkte ermöglichen im Nachgang die vollständige Route nachzuvollziehen. Zusätzlich wird die Fahrtrichtung hinterlegt, um auch Kehrtwenden zu detektieren.

Anschließend ist eine straßensegmentenspezifische Betrachtung der Daten möglich. Die Anzahl der befahrenen Abschnitte ist unterschiedlich. Extrem hoch frequentierte Straßenabschnitte und Kreuzungen in München und Umgebung wurden über 30 000-mal befahren. Zur Veranschaulichung ist die Abdeckung der Fahrdaten mit ihrer Häufigkeit qualitativ in Abb. 3.4 dargestellt. Daraus wird ersichtlich, dass der Großteil der Hauptverkehrswege mit über 1 000 Fahrten befahren wurde.

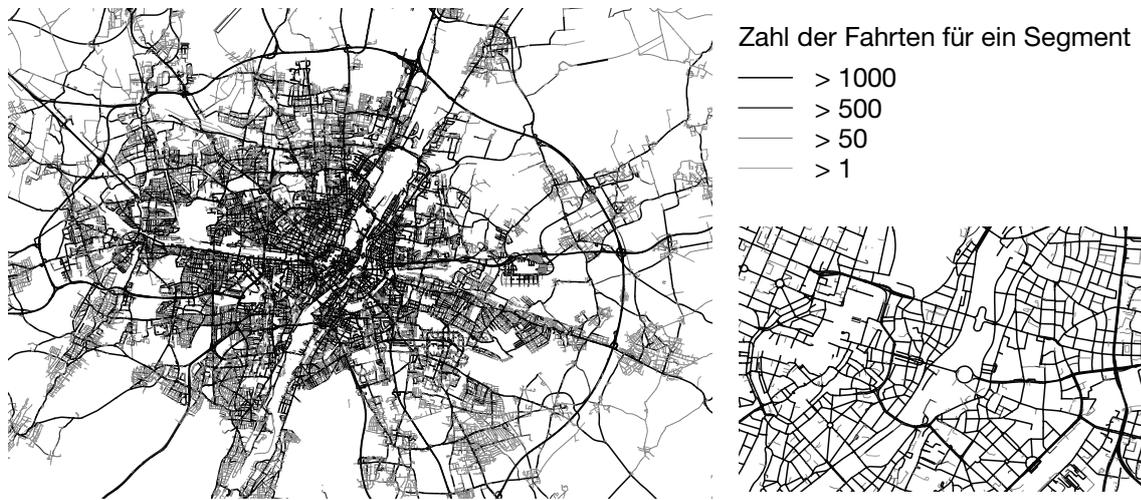


Abbildung 3.4 : Befahrene Straßenabschnitte im Stadtgebiet München (Ausschnitt zwischen Altstadt und Ostbahnhof)(Karte: OSM ©📍📍)

Eine Übersicht ist dazu in Tab. 3.2 zu finden. Straßen in Wohngebieten werden im Verhältnis deutlich seltener befahren und sind demnach im Zieldatensatz weniger vertreten. Es zeigt sich auch eine Abhängigkeit des Anteils der befahrenen Straßenklassen von der Nutzergruppe. In Abb. 3.5 sind für Taxi, Handwerksbetriebe und Privatfahrer die durchschnittlichen Streckenanteile nach [100] über alle Fahrten dargestellt. Der auch pro Fahrt relativ geringe Anteil von Wohngebiets- und tertiären Straßen begründet den geringen Anteil dieser im Zieldatensatz.

Tabelle 3.2 : Übersicht zu den Straßensegmenten im Zieldatensatz

Klasse	Beschreibung	Anzahl			Kilometer		
		insgesamt	vorhanden	%	insgesamt	vorhanden	%
11/12	motorway/link	2 186	1 899	86,9	1 482	1 408	95,0
13/14	trunk/link	580	544	93,8	187	178	95,2
15/16	primary/link	7 449	6 360	85,4	1 972	1 619	82,1
21/22	secondary/link	16 434	13 554	82,5	3 902	2 920	74,8
31/32	tertiary	16 949	9 989	58,9	4 834	2 420	50,1
41	residential	122 459	29 987	24,5	16 248	3 715	22,9
43	unclassified	22 845	4 845	21,2	8 983	1 320	14,7
51	service	427	249	58,3	36	18	50,0
63	living-street	4 810	1 013	21,1	399	88	22,1

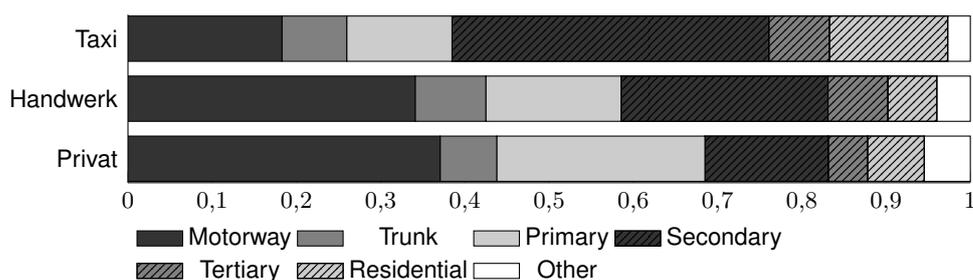


Abbildung 3.5 : Anteile an den Straßenklassen für unterschiedliche Nutzertypen nach [100]

3.2.2 Metadaten

Für die aufgezeichneten Fahrten, sowie die einzelnen Straßensegmente, werden zusätzliche Metadaten erzeugt. Neben den in Tab. 1.1 vorgestellten Parametern, die sich anhand der Geschwindigkeit ableiten lassen, wurde für jeden Straßenabschnitt pro Fahrt die durchschnittliche zeitliche \bar{v}_t und räumliche Geschwindigkeit \bar{v}_s ermittelt. Die Anzahl der Stopps n_{stop} sowie die gesamte Haltdauer t_{stop} auf dem Segment wird bestimmt. Weiterhin werden anhand der Geschwindigkeit Werte für die mittlere Beschleunigung \bar{a} , deren positive \bar{a}_p und negative Anteile \bar{a}_n , sowie die entsprechenden Zeitanteile a_{tp} , a_{tn} und die Zahl der Beschleunigungs- n_{ap} bzw. Verzögerungsvorgänge n_{an} berechnet.

Die Straßensegmente werden außerdem Gebieten zugeordnet. Diese werden anhand der Straßenklassen untergliedert [97]. Bereiche, die von Sekundärstraßen umschlossen werden, bilden die Sekundärgebiete. Übergeordnet sind diesen Primärgebiete, die von Primärstraßen umschlossen werden. Die Unterteilung ermöglicht einen skalierbaren Aufbau von Modellen, an die weitere Gebiete angekoppelt werden können. Für den Bereich Oberbayern wurden rund 100 Primärgebiete und 900 Sekundärgebiete bestimmt.

3.2.3 Verkehrsinformationen

In Kapitel 2.2.3 wurde der Verkehrsfluss als temporäre dynamische Information der Karte eingeführt. Mehrere Anbieter, darunter auch HERE [121] stellen Verkehrsflussinformationen oder Stauhinweise zur Verfügung [122]. HERE stellt eine Programmierschnittstelle bereit, die den Zugang zu straßensegmentspezifischen Daten ermöglicht. Darunter kann für jedes Segment die durchschnittliche Geschwindigkeit und die Reisezeit unter Berücksichtigung des Verkehrs bezogen werden. Zudem wird ein Maß für den Stau berechnet, der den freifließenden Verkehr widerspiegelt [123]. Eine grundsätzliche Vorgehensweise zur Bestimmung der Verkehrsgeschwindigkeit und der Reisezeit aus GPS-Daten und deren Abhängigkeiten wird in [124] gegeben.

Zur Berechnung der mittleren Verkehrsgeschwindigkeit werden in der Regel mehrere Datenquellen herangezogen. Sowohl mobile Geräte, FCD oder stationäre Messstationen dienen zur Ermittlung [27, 125, 126]. Darüber hinaus existieren Verfahren, die versuchen, den Verkehrsfluss zu ermitteln, falls keine aktuellen Messwerte vorliegen [26, 127, 128].

Die Prädiktion des Fahrprofils nutzt aktuelle Verkehrsinformationen. Diese sind im Datensatz allerdings nicht für alle Fahrten vorhanden, weil diese nicht gleichzeitig mit den Flottenversuchen erhoben wurden. Somit muss dies im Nachgang evaluiert werden. Dafür wird über einen Zeitraum von über drei Monaten die mittlere Verkehrsgeschwindigkeit von HERE im Raum München untersucht. Um die Daten dem OSM-Kartenmaterial zuordnen zu können, werden identische Straßensegmente mit Hilfe eines Routingverfahrens bestimmt. Insgesamt werden somit knapp 22 000 Straßenzüge im Hauptverkehrswegenetz identifiziert. HERE übermittelt den Verkehrsfluss mit einer Auflösung von $0,01 \frac{\text{m}}{\text{s}}$ [123]. Die Daten werden im 15-Minutentakt aufgezeichnet.

Der Verkehrsfluss weist deutlich zeitliche und räumliche Muster auf [125, 129]. Diese zeigen sich über den Tagesverlauf. Für Wochentage, Wochenenden bzw. Feiertage werden für die Abschnitte der Median, sowie 25-Perzentil und 75-Perzentil bestimmt. Um einen ähnlichen Verlauf anzunähern, wird die mittlere räumliche Geschwindigkeit \bar{v}_s , die pro Straßenabschnitt für jede Fahrt bestimmt wird, herangezogen [130, S. 14]. Gemittelt über alle Fahrten in dem jeweiligen Abschnitt können ebenso Tagesganglinien erstellt werden. Dies wird in den Abb. 3.6 und 3.7 für zwei Abschnitte veranschaulicht. Die Medianwerte weisen eine hohe Ähnlichkeit auf. Für die Fragestellung ist die qualitative Änderung über den Tagesverlauf entscheidend. Der Absolutwert ist abhängig von der Datenbasis und kann auch dimensionslos angegeben werden. Für 236 Straßenabschnitte, bei denen die Verteilung der Fahrten über den Tag ausreichend ist, um repräsentative Tagesverläufe zu erstellen, werden die mittleren Tagesverläufe miteinander verglichen. Dafür werden aufgrund von vereinzelt Ausreißern die Ganglinien mit einem gleitenden Mittelwert geglättet. Unter der Annahme eines linearen Zusammenhangs wird das Bestimmtheitsmaß R^2 für die Qualität der Approximation ermittelt. Es zeigt sich, dass mit einem

Median von $R^2 = 0,82$ für alle betrachteten Abschnitte die mittlere räumliche Geschwindigkeit die Verläufe plausibel wiedergibt.

In vielen Fällen wird der Wert für den Verkehrsfluss in Klassen ausgegeben, die farblich auf einer Karte dargestellt werden. Die Auflösung der Daten von HERE zeigt, dass bereits eine detailliertere Auswertung möglich ist. Für den Verkehrsfluss in dieser Arbeit wird eine Abstufung in $1 \frac{m}{s}$ Schritten vorgenommen.

Damit bestimmt sich der segmentspezifische Verkehrsfluss zu:

$$TI_S = \begin{cases} \bar{v}_{s25}^S & \bar{v}_s < \bar{v}_{s25}^S \\ \bar{v}_s & \\ \bar{v}_{s75}^S & \bar{v}_s > \bar{v}_{s75}^S \end{cases}, \quad (3.1)$$

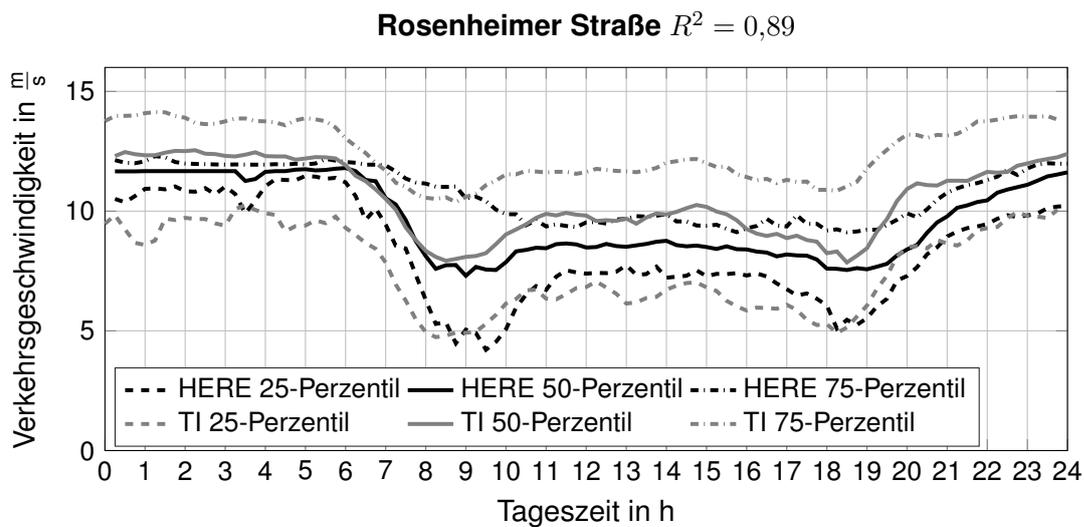


Abbildung 3.6 : Tagesverlauf des Verkehrsflusses auf Rosenheimer Straße

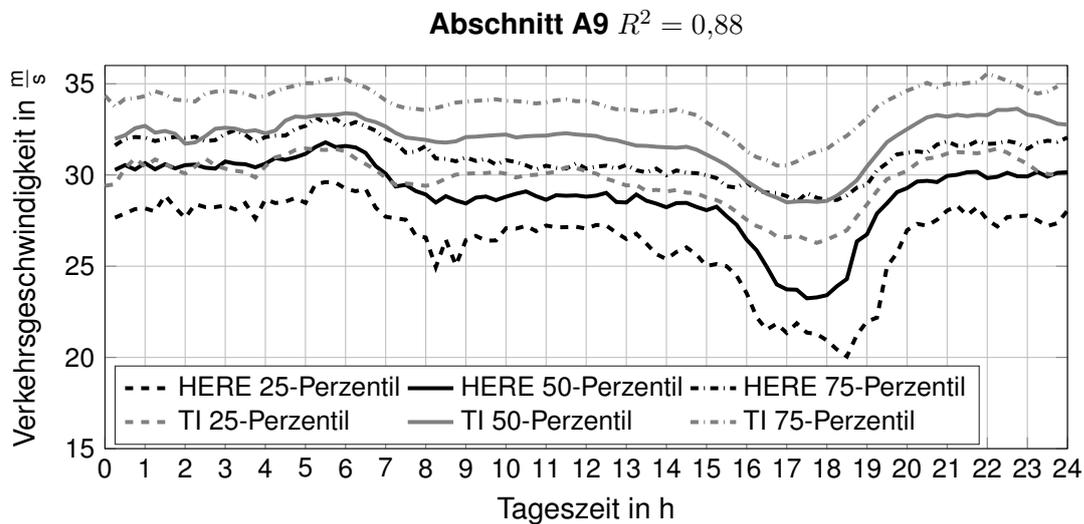


Abbildung 3.7 : Tagesverlauf des Verkehrsflusses auf einem Abschnitt der A9

4 Trip-Prädiktion

Nachfolgend wird erst auf den Stand der Forschung eingegangen. Anschließend wird die angewandte Methodik erläutert. Daraufaufgehend werden die erzielten Ergebnisse vorgestellt und abschließend bewertet.

4.1 Stand der Forschung

Der Stand der Forschung lässt sich in die jeweiligen Motive der Prädiktion des Mobilitätsverhaltens einordnen. Je nach Fragestellung wird versucht, entweder das Fahrtziel (Kapitel 4.1.2), das Fahrtziel und die vollständige Route dazu (Kapitel 4.1.3), bzw. einen Teil der Route (Trip) (Kapitel 4.1.4) zu prädizieren. Zuerst wird ein kurzer Überblick zu den übergeordneten Unterscheidungsmerkmalen für die Prädiktion des Mobilitätsverhaltens und die Verarbeitung der Daten gegeben.

4.1.1 Verarbeitung und Bewertung von Mobilitätsdaten

Mobilität wird durch einen räumlich-zeitlichen Zusammenhang beschrieben. Dies kann in unterschiedlichen Datenformaten und Abstraktionsgraden erfolgen. Die Daten können identifizierbare Muster aufweisen. Andrienko [131] stellt diese bspw. zur visuellen Interpretation in einem Raum-Zeit-Würfel dar, bei dem die 2D-Positionsdaten in einer dritten Dimension zeitlich abgehoben werden. Somit werden wiederkehrende Fahrten ersichtlich und es können Muster im Mobilitätsverhalten identifiziert werden.

Lin und Hsu [132] geben eine Übersicht über Ansätze der Auswertung von GPS-Daten zur Identifikation von darin enthaltenen Mustern. Sie beschreiben diese Aufgabe als wichtige Basis, um das Bewegungsverhalten von Nutzern verstehen zu können. Die Bewegung von mehreren Nutzern unterscheidet sich in einer geographischen Region. Das individuellen Verhalten weist jedoch einen stark, wiederkehrenden Charakter auf. Offene Fragestellungen liegen im Bereich der Datenkompression, der Erstellung von Mobilitätsmodellen [132], sowie der Prädiktion der Mobilität und darin liegende Unsicherheiten [133].

Pelekis und Theodoridis [118] behandeln das Management und die Modellierung von Mobilitätsdaten. Das extrahierbare Wissen beschränkt sich nicht auf eine individuelle Person, sondern ist auch in der Lage, das Kollektiv zu bewerten. Die Herangehensweise ist dabei auf andere Fragestellungen anwendbar. Es dient bei der Modellierung von Fahrzeugflotten [134] oder im Bereich der Schifffahrt [135]. Bei der Modellierung der Mobilität stellt Yan [136] den Einsatz von semantischen Trajektorien vor. Das Vorgehen extrahiert aus den Rohdaten ein konzeptionelles Modell bestehend aus Sequenzen, die Daten logisch partitionieren. Als nächste Ebene folgt ein semantisches Modell, das die Sequenzen mit Zusatzinformationen von geographischem Bezug ergänzt. Letzteres benötigt häufig zur Erstellung Daten von Drittquellen, weil diese rein aus den Karteninformationen meist nicht abgeleitet werden können. Ein Beispiel wird in Abb. 4.1 dargestellt.

Yang [137] beschreibt das Labeln von Trajektorien aus einem großen Datensatz mit probabilistischen Mitteln. Als Vorstufe setzt er dabei auf ein Mapmatching-Verfahren, das ebenso als Labeling verstanden werden kann und die Datenqualität verbessert. Als Anreicherung der Daten wird häufig die Wahl des Verkehrsmittels bestimmt. Die zusätzlichen Kontextinformationen können dafür aus dem räumlich-zeitlichen Bezug generiert werden [138]. Es existieren hierfür mehrere Ansätze, unter anderem wird der Einsatz von künstlichen neuronalen Netzen untersucht [139, 140].

Die wichtigsten Unterscheidungsmerkmale der Forschungsthemen in diesem Bereich gliedern sich wie folgt:

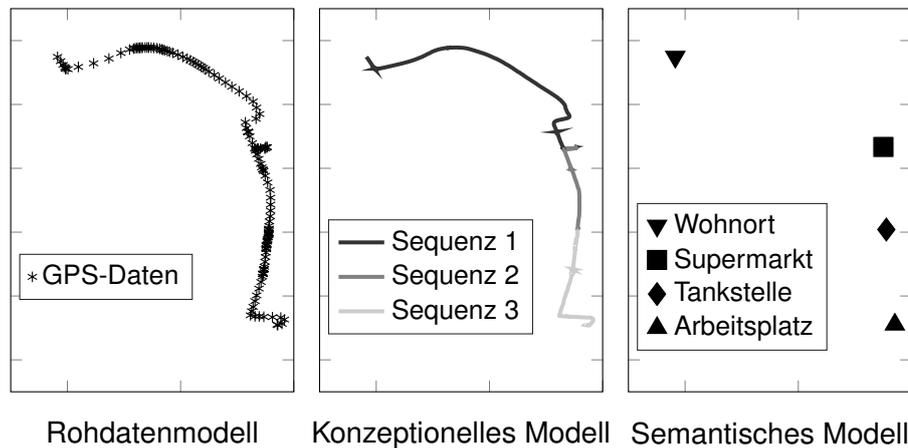


Abbildung 4.1 : Semantische Beschreibung von Trajektorien

- Wahl der Datenbasis
- Gewählter Abstraktionsgrad der Daten
- Nutzerbasis für die Betrachtung
- Erweiterungen und Modellergänzungen

Diese werden kurz näher erläutert.

Datenbasis

Eine Prädiktion der Position des Fahrzeugs kann grundsätzlich in einer rein geometrischen Betrachtung erfolgen. Dabei werden anhand der aktuellen Position Kartenattribute für die vorausliegenden Streckenabschnitte, ohne Hintergrund über historische Bewegungsmuster, bewertet.

Als Datenbasis können auch Bildsequenzen dienen. Mit einer videobasierten Untersuchung werden Trajektorien auf Kreuzungen geclustert und Muster extrahiert. Häufig wird dies mit Hilfe von einer Self-Organizing-Map gelöst [141, 142]. Kim [143] verwendet eine Gaußprozess-Regression zur inkrementellen Prädiktion der Abbiegevorgänge auf Basis von Videos. Die Datenbasis begrenzt die Bewertung räumlich auf vergleichsweise kurze Abschnitte.

Die für diese Arbeit relevante Datenbasis bezieht sich auf GPS-Daten. Für gesamte Trajektorien ermittelt Vlachos [144] bspw. Muster mit Hilfe des Ähnlichkeitsmaßes Longest Common Subsequence. Die längsten gemeinsamen Unterteilungen können räumlich oder zeitlich indiziert werden [118, S. 82–87]. Eine Betrachtung kann bezogen auf das Straßennetzwerk für ganze Trajektorien oder Teilabschnitte erfolgen [145].

Abstraktionsgrad

Eine weitere Unterscheidungsmöglichkeit ist der Abstraktionsgrad. Zum einen kann eine Prädiktion direkt auf Basis der GPS-Daten durchgeführt werden. Zum anderen können die Daten abstrahiert werden. Zhang [146] führt eine Generalisierung der Daten vor dem Cluster-Verfahren durch. Damit wird die Datenmenge deutlich reduziert und das Verfahren beschleunigt. Weiterhin kann die Datenmenge auch mit Hilfe von Mapmatching reduziert werden. Dies schränkt die Bewegungsmöglichkeiten auf das Straßennetz ein. Die Daten können weiter auf ein semantisches Modell wie in Abb. 4.1 abstrahiert werden. Im weitesten Fall wird die Fahrt durch eine Quelle-Ziel-Beziehung (Origin-Destination (OD)) ausgedrückt [147]. Damit ist keine Beschreibung der Route mehr möglich, aber ein zeitabhängiger Bezug kann dargestellt werden [148].

Nutzerbasis

Weiterhin kann unterschieden werden, ob als Datenquelle die Fahrten von einem Fahrer oder die von mehreren Fahrern verwendet werden. Ebenso kann bei der Prädiktion versucht werden, das Mobilitätsverhalten von einem einzelnen Fahrzeug oder einer ganzen Flotte vorherzusagen.

Erweiterungen

Die Datenbasis kann darüber hinaus mit weiteren Quellen ergänzt werden. So können bspw. POIs bei der Betrachtung hinzugezogen werden. Weiterhin kann eine Einschränkung oder Priorisierung von Regionen oder Zeiträumen erfolgen.

4.1.2 Fahrtzielprädiktion

Bei dieser Kategorie wird rein das Ziel einer Fahrt prädiziert. Dies erfolgt entweder vor Fahrtbeginn oder während einer Fahrt. Im letzteren Fall werden die verwendeten Daten kontinuierlich aktualisiert.

Casott [19] stellt ein Navigationssystem vor, das vergangene Ziele im Fahrzeug speichert. Diese werden um zeitliche und örtliche Informationen zu Beginn der Fahrt ergänzt. Zu Beginn einer neuen Fahrt schlägt das Navigationsgerät abhängig von den aufgezeichneten Kriterien die wahrscheinlichsten Ziele vor. Das Gerät kann zudem mit einem persönlichen Terminkalender verknüpft werden, um weitere personalisierte Daten zu beziehen.

Vorreiter auf dem Gebiet der Prädiktion der Mobilität ist Microsoft Research. Krumm [149] versucht das Ziel während der Fahrt zu prädizieren. Er verwendet eine probabilistische Methode, bei der Fahrzeiten anderer Fahrer mit berücksichtigt werden. Über den Betrachtungsraum wird ein Raster mit einer Weite von einem Kilometer gelegt. Die aktuelle Fahrt repräsentiert dabei eine Sequenz von Zellen über dieses Raster. Über den Satz von Bayes werden bedingte Wahrscheinlichkeiten wiedergegeben, ob eine Zelle ein Ziel für die betrachtete Fahrt ist. Der Median des Fehlers für die Prädiktion des Ziels beträgt nach 25 % der Fahrt 21 km, nach 75 % der Fahrt reduziert sich der Wert auf 3 km.

Alvarez-Garcia [150] ermittelt das Ziel einer Fahrt ohne die Zuhilfenahme einer digitalen Karte. Es entsteht ein Kennfeld aufbauend auf GPS-Daten, das häufig angefahrne Ziele und markante Punkte aus den Daten extrahiert. Mit dem begrenzten Set an Datenpunkten wird ein Hidden Markov Model (HMM) angelernt. Nur Teilinformationen der aktuellen Fahrt werden benötigt, um ein mögliches Ziel zu bestimmen. Der personenbezogene Ansatz kann das Ziel bei 50 % der Fahrt zu 46 % korrekt bestimmen. Bei 75 % der Fahrt steigt der Wert auf bis zu 81 %. Einen darauf aufbauenden Ansatz stellt Gams [151] vor. Dabei werden die Markov Ketten um eine zeitliche Dimension erweitert.

Einen weiteren Ansatz stellen Manasseh und Sengupta [152] vor. Sie statten zehn Personen mit Smartphones für einen Zeitraum von zwei bis neun Wochen aus. Im Laufe der Datenaufzeichnung legt jeder Teilnehmer zwischen 13 und 49 Fahrten zurück. Zudem fährt jeder Teilnehmer nur eine begrenzte Anzahl an wiederkehrenden Zielen an. Die Anzahl der Ziele liegt zwischen vier und 19. Auf den Datensatz aufbauend wird eine OD-Matrix basierend auf dem United States National Grid (USNG) Netz mit einer Weite von einem Kilometer erstellt. Verschiedene Verfahren des maschinellen Lernens werden untersucht. Am besten schneidet ein vereinfachter Entscheidungsbaum (Decision Tree with Pruning) ab. Mit bis zu 96 % Genauigkeit kann für alle Teilnehmer das nächste Ziel bestimmt werden.

Chen [153] stellt heraus, dass es Vorteile bringt, die Prädiktion der individuellen Fortbewegung mit der des Kollektivs zu vereinen. Damit kann bei wenig persönlich aufgezeichneten Daten eine Nutzung erfolgen und bei vorhandenen Daten die individuellen Muster die Prädiktion verbessern. Bei den verwendeten Daten handelt es sich um die Aufzeichnungen von 300 Überwachungskameras im Straßennetz. Mit einem kollektiven und einem individuellen Markov Model werden die Orte und Zeitpunkte, bei denen Fahrzeuge über die Nummernschilder erkannt werden, angelernt. Das hybride Modell zeigt deutliche Verbesserungen gegenüber den einzelnen Modellen.

Mit Hilfe eines Variable Order Markov Model (VOMM) prädiziert Yang [154] das nächste Ziel. Dafür werden die GPS-Daten von den Aufenthaltsorten um semantische Umgebungsinformationen erweitert. Um diese zu erlangen, wird eine Dichtebasierte räumliche Clusteranalyse mit Rauschen (DBSCAN) auf den Datensätzen angewendet, um signifikante Orte zu ermitteln.

Das VOMM ermittelt anschließend den wahrscheinlichsten nächsten Aufenthaltsort. Mit 18 000 Trajektorien werden die Modelle erstellt und evaluiert. Markov Modelle können das nächste Ziel mit 48 % korrekt bestimmen, das VOMM erreicht einen Wert von 50 %.

Liu [155] verwendet räumlich-zeitliche rekurrente neuronale Netze für die Prädiktion des nächsten Ziels. Es zeigt eine deutliche Verbesserung gegenüber dem Stand der Technik bei der Prädiktion der Bewegung von Nutzern des standortbasierten sozialen Netzwerks Gowalla und einem Datensatz zu terroristischen Anschlägen.

Mit Hilfe von künstlichen neuronalen Netzen prädiziert [156] das Ziel einer Taxifahrt. Ein derartiges System kann dabei dienen, die Verteilung der Taxis frühzeitig zu steuern. Im Rahmen der Konferenz ECML/PKDD 2015 wurde ein Wettbewerb veranstaltet, das Ziel der Fahrt in Form von GPS-Koordinaten mit Hilfe der initialen Trajektorien der Fahrzeuge zu Beginn der Fahrt vorausszusagen. Zusätzlich wurden Metadaten wie Zeit, Fahreridentifikationsnummer und Kundenidentifikationsnummer (falls vorhanden) erhoben, die verwendet werden konnten. Insgesamt wurden Daten von 442 Taxis über ein Jahr in der Stadt Porto erfasst. Die GPS-Punkte wurden alle 15 Sekunden erhoben. Das Gewinnerteam [156] hat verschiedene Arten an neuronalen Netzen untersucht. Darunter ein mehrlagiges Perzeptron (multilayer perceptron), speicherbasierte Netzwerke (memory network) und ein bidirektionales, rekurrentes neuronales Netz (bidirectional recurrent neural network). Letzteres hat aufbauend auf den fünf ersten Datenpunkten und den Metadaten das beste Ergebnis erzielt. Der gewählte Algorithmus ist nicht auf die Taxifragestellung beschränkt, sondern ist generell in der Lage, mit einer Sequenz von variabler Länge einen Ausgang mit festgesetzter Länge zu prädizieren.

Lv [157] untersucht den gleichen Datensatz zur Prädiktion von Taxizielen. Der Algorithmus T-CONV erzielt mit einem mehrlagigen, faltenden neuronalen Netz (multilayer convolutional neural networks) ähnliche Ergebnisse. Die Trajektorien werden dabei als Bilder interpretiert. Diese Form der Netze haben sich im Bereich der Verarbeitung von Bilddaten etabliert.

Tabelle 4.1 : Literaturübersicht Fahrtzielprädiktion

Quelle	Jahr	Methode	Datenbasis	Abstaktion	Nutzer
[149]	2006	Satz von Bayes	4 300 Fahrten	Raster	K
[19]P	2008	Kriterien	-	Adresse	I
[150]	2010	HMM	645 Fahrten	Punkte	I
[151]	2012	HMM + Zeit	-	POI	I
[152]	2013	Decision Tree wP	269 Fahrten	OD-Matrix	I
[153]	2014	Markov Modell	> 1,7 Mio. Fahrten	Punkte	I+K
[154]	2014	DBSCAN + VOMM	> 17 000 Fahrten	POI	K
[156]	2015	BRNN	> 1,7 Mio. Fahrten	Punkte	K
[157]	2016	CNN	> 1,7 Mio. Fahrten	Raster	K
[155]	2016	RNN	Gowalla	Punkte	I

P = Patent, K = Kollektiv, I = Individuell

4.1.3 Routenprädiktion

Als Erweiterung der Fahrtzielprädiktion wird bei der Routenprädiktion zusätzlich die komplette Route ermittelt. Dies kann auch während der Fahrt erfolgen.

Simmons [158] erstellt für 46 Fahrten in Michigan ein Prädiktionsmodell basierend auf Straßensegmenten. Dabei prädiziert der Algorithmus basierend auf dem letzten Segment das darauffolgende wahrscheinlichste Segment. Zudem wird überwacht, ob es sich dabei um ein präferiertes Fahrtziel handelt. Hinter der Methode befindet sich ein HMM. Es wird ermittelt, wie häufig ein Abschnitt überfahren wurde, was das vorausgegangene Segment war und welches Ziel die Fahrt hatte. Zusätzlich werden weitere Faktoren wie die Tageszeit berücksichtigt und das HMM entsprechend ergänzt. Als Ergebnis wird festgehalten, dass der nächste Abschnitt in den meisten Fällen mit 98 % genau bestimmt werden kann. Allerdings handelt es sich bei

95 % der Übergänge zwischen den Straßensegmenten um direkte Verbindungen, ohne die Möglichkeit eines Abbiegevorgangs. Bei Straßenabschnitten ohne erzwungenen Übergang liegt die Genauigkeit bei 70-80 %. Der Kern dieser Untersuchung liegt nicht auf dem Fahrtziel, daher kann es gleichzeitig in dem Bereich der Trip-Prädiktion eingeordnet werden.

Das Verfahren von Carlsson [159–161] setzt auf eine Datenbank im Fahrzeug. Darin werden häufige Routen, sowie Fahrtziele und Ereignisse festgehalten. Als Ereignis kann bspw. eine Kurve, bestimmte Steigungen oder die Position einer erkannten LSA dienen. Während der Fahrt wird ein Abgleich von Situationen und der Route vorgenommen. Tritt eine Situation ein, wird in der Region nach vergleichbaren, bereits aufgezeichneten Situationen gesucht. Falls ein Eintrag gefunden wird, kann damit das Fahrtziel der aufgezeichneten Fahrt und die dazugehörige Route identifiziert werden. Diese wird folgend als wahrscheinlichste angenommen.

Froehlich und Krumm [162] haben die Mobilität von 252 Fahrer für jeweils ca. 15 Tage mit GPS-Sensoren erfasst. Für die Privatpersonen wurde festgestellt, dass bei einem Betrachtungszeitraum von mehr als 40 Tagen knapp 60 % der Fahrten einen repetitiven Charakter haben. Um eine vollständige Route innerhalb der ersten Kilometer zu identifizieren, werden Muster in den Trajektorien gesucht. Die gefilterten GPS-Daten der aufgezeichneten Fahrten werden über die Hausdorff-Metrik miteinander verglichen. Ähnliche Wege werden dabei herausgearbeitet. Darauf aufbauend wird ein hierarchisches Clusterverfahren angesetzt, das die einzelnen Wege zu Routen zusammenfasst. Bei Beginn einer neuen Fahrt wird über die Metrik die aktuelle Fahrt mit den identifizierten Routen verglichen. Die ähnlichste Route wird dann als die wahrscheinlichste angenommen. Der Algorithmus kann keine belastbare Aussage über eine zukünftige Route treffen, falls die aktuelle Route vorher noch nicht befahren wurde. Sind 50 % der Fahrt erfolgt, liegt die tatsächliche Route zu 25-30 % unter den zwei wahrscheinlichsten. Werden bevorzugt wiederkehrende Routen betrachtet, erhöht sich das Ergebnis auf 60 %.

Tiwari [163] baut auf den Algorithmus von Froehlich [162] auf. Allerdings werden die GPS-Daten vorab einem Mapmatching unterzogen. Dadurch reduziert sich der Speicherbedarf für die Daten um 99,51 %. Die Segmentfolgen werden wie in [162] über die Übereinstimmung mit dem kleinsten Abstand gewählt. Diese Variante liefert eine ähnliche Genauigkeit wie [162], ist jedoch deutlich schneller.

Necula [164] setzt bei der Prädiktion ebenso auf Straßensegmente und verwendet ein HMM zur Bestimmung des wahrscheinlichsten Fahrtziels. Des Weiteren können zusätzliche Attribute wie der Wochentag oder die Tageszeit mitberücksichtigt werden. Die Methode wird ebenso mit [162] verglichen. Es werden ähnlich gute Ergebnisse erzielt. Es zeigt sich außerdem eine Tendenz für bessere Ergebnisse, falls die Fahrer mehrere unterschiedliche Ziele ansteuern, bzw. die Zahl der Fahrten steigt. Die Methode zeichnet sich wie [163] durch einen reduzierten Datensatz aus.

Tabelle 4.2 : Literaturübersicht Routenprädiktion

Quelle	Jahr	Methode	Datenbasis	Abstaktion	Nutzer
[158]	2006	HMM + Zeit	46 Fahrten	Segmente	I
[160]	2008	Kriterien	-	Segmente	I
[162]	2008	Hausdorff + Hier. Cluster	27 000 Fahrten	-	I
[163]	2013	Hausdorff + Hier. Cluster	27 000 Fahrten	Segmente	I
[164]	2014	HMM	269 Fahrten	Segmente	I

K = Kollektiv, I = Individuell

4.1.4 Trip-Prädiktion

Zuerst wird auf das System des elektronischen Horizonts eingegangen. Anschließend werden Patentschriften vorgestellt, die auf die Ermittlung des Trips schließen. Zuletzt werden wissenschaftliche Methoden aufgezeigt, die versuchen, den wahrscheinlichen Pfad zu bestimmen.

Elektronischer Horizont

Wie in Kapitel 2.2.3 bereits angesprochen, besteht die Möglichkeit über das ADASIS v2 Protokoll

[87] dem Fahrzeug einen elektronischen Horizont zur Verfügung zu stellen. Im Protokoll werden grundlegende Straßengeometrien und Attribute wie Geschwindigkeitsbegrenzungen oder die Zahl der Fahrspuren festgeschrieben. Des Weiteren liegt der Horizont um einen möglichen Pfad (Most Probable Path (MPP)) und optionalen alternativen Routen. Das Protokoll ist dabei herstellerübergreifend und unabhängig von einem bestimmten Bussystem.

Ress [165] beschreibt den MPP als obsolet, wenn der Fahrer das Ziel in das Navigationsgerät eingibt. Allerdings weist er auch auf die in Kapitel 1 beschriebene Problematik hin, dass dies nicht immer der Fall ist. Ein wahrscheinlicher Pfad muss demnach ermittelt werden. Dabei können Fahrereingaben wie Blinker, Pedalstellung oder der Lenkradwinkel mitberücksichtigt werden. Dies wird für die nächste Kreuzung berücksichtigt. Ein intelligenter Algorithmus kann eine vollständige Route aus den Entscheidungen an den einzelnen Kreuzungen zusammensetzen. Die tatsächlich befahrenen Routen können dabei als Entscheidungshilfe dienen und werden in einer Datenbank im Fahrzeug hinterlegt und verbessert so das Ergebnis. Ähnlich beschreibt Carlsson [160] eine Variante, die während der Fahrt eine Datenbank im Fahrzeug nach Einträgen durchsucht, die zu ähnlichen Zeitpunkten und Ereignissen bereits befahrene Routen als MPP vorschlägt.

Einen offenen Ansatz für einen elektronischen Horizont beschreibt Burgstahler [90, 91, 166] mit dem System Horizon.KOM. Er bemängelt, dass die Systeme keinen Einblick in die Funktionsweisen und die Ermittlung des MPP erlauben. Horizon.KOM basiert auf quelloffenen Daten von OSM. Bei der Ermittlung des MPP werden wie in [165] höhere Straßenklassen, bzw. kleinere Abbiegewinkel bevorzugt. Es können häufig gefahrene Routen aufgezeichnet und berücksichtigt werden.

Patentschriften

In einer Patentschrift beschreibt Bechtolsheim [167] mögliche Kostenfunktionen, von denen der wahrscheinliche Pfad abhängt. Je nach Assistenzfunktion kann sich die Ausprägung des Horizonts auch unterscheiden. Die Länge des Horizonts bestimmt sich demnach aus der Fahrgeschwindigkeit und einer Reisezeit bzw. der Distanz ab der aktuellen Position. Ebenso können spezielle Straßensegmente oder Situationen wie Kehrtwenden, kreisläufige Straßenzweige oder auch unzugängliche Segmente mitberücksichtigt werden. Die Berechnung des MPP erfolgt nach den funktionalen Klassen der Straßensegmente, bzw. der Kategorie. Falls diese bei einer Kreuzung gleichwertig sind, wird die Straße mit dem kleineren Abbiegewinkel bevorzugt. Falls auch hier keine Unterscheidung getroffen werden kann, wird sich für einen Rechtsabbiegevorgang entschieden.

Darüber hinaus existieren mehrere Erweiterungen für den MPP. Gee [168] erstellt eine Voraussichtstabelle zur Rechenzeitoptimierung im Fahrzeug. Haug [169] berücksichtigt zur Bestimmung auch dynamische Informationen wie Sperrungen oder Baustellen. Zudem ist es möglich, den MPP abhängig von dem Verkehrsfluss zu gestalten oder vorausfahrende Fahrzeuge mitzuberechnen. Ghisio [170] fügt zur Erstellung des MPP Umgebungsinformationen mit zeitlicher Abhängigkeit hinzu. Dabei kann es sich um die Öffnungszeiten von Einkaufszentren, Industrieanlagen, oder Sportzentren handeln. Eine dynamische Anpassung der Daten ist möglich.

Engel [171] stellt vor, den MPP abhängig von Fahrwahrscheinlichkeitsdaten zu erstellen. Dabei sollen Daten von mindestens einem Fahrzeug verwendet werden. Die Daten können auch unpersonalisiert von mehreren Fahrern über eine externe Quelle übertragen und zur Verbesserung zusätzlich Daten im Fahrzeug berücksichtigt werden. Neue Fahrten können an die Quelle zur Verbesserung zurückgemeldet werden. Die verwendete Methodik wird nicht beschrieben. Mattmüller [172] verbessert dies durch eine Trennung der Daten für fahrzeugbezogene und allgemeine Berechnungen des MPP auf zwei Datenbanken. Diese können unabhängig voneinander durch Selbstlernen aktualisiert werden. Die Methodik wird nicht näher erläutert. Pu [88] stellt eine Möglichkeit vor, die Daten von LSA dynamisch über das ADASIS Protokoll zu übertragen.

Wissenschaftliche Methoden

Tao [173] beschreibt die Prädiktion der Bewegung von Objekten im Allgemeinen. Dabei geht er von einer nichtlinearen Bewegung aus. Die Methode ermöglicht frei wählbare Bewegungs-

muster, die nicht notwendigerweise vorab bekannt sein müssen. Die Objekte beschreiben das Verhalten über rekursive Bewegungsfunktionen und hinterlegen diese in einer Datenbank. Mit den Funktionen können die nächsten 20 Zeitschritte beschrieben werden. Die Methode ist bewegungsbasiert und unabhängig von Fahrzeugen oder dem Straßennetz.

Karimi [174] prädiziert die nächsten Streckenabschnitte mit Hilfe von Übergangswahrscheinlichkeiten an den Kreuzungen. Zu Beginn sind diese an den Kreuzungen gleich gewichtet. Abhängig von der Anzahl der Überfahrten ändert sich die Wahrscheinlichkeit. Mit einem dynamischen Fenster können für weitere anschließende Segmente vollständige Pfade bestimmt werden. Das Fenster ist dabei abhängig von der aktuellen Geschwindigkeit und dem Prädiktionshorizont. Der somit wahrscheinlichste Pfad basiert auf den Segmenten mit den höchsten Übergangswahrscheinlichkeiten und ist jeweils nur von dem letzten Segment abhängig. Der Algorithmus wird stabiler, je häufiger eine gleiche Route befahren wird.

Diese Methode führt Krumm [175] fort. Er erstellt für die Segmente ein Markov Modell und prädiziert die nächsten Segmente. Mit Daten von 100 Fahrern lernt der Algorithmus. Zudem wird die Genauigkeit abhängig von der Ordnung des Markov Modells untersucht. Bei einer höheren Ordnung reduziert sich die Zahl an gleichen Sequenzen. Mit einem Modell zehnter Ordnung kann das nächste Segment zu 90 % korrekt vorhergesagt werden. Nach zehn Segmenten reduziert sich der Wert auf 50 %. Krumm stellt keinen Zusammenhang von Genauigkeit und Tageszeit fest. Wenn nur die Daten von einem Fahrer ausgewertet werden, ist eine Steigerung von einem Prozent festzustellen.

Aufbauend auf [175] entwickelt Xue [176] ein VOMM. Aus einer Vielzahl an Taxidaten werden mit dem VOMM häufig befahrene Routen extrahiert. Über einen probabilistischen Suffixbaum werden die nächsten Segmente prädiziert. Die variable Ordnung ermöglicht es, möglichst viele Muster aus den Realdaten zu erkennen. Zudem werden verschiedene Modelle zu Zeiten mit unterschiedlichem Verkehrsfluss erstellt. Xue stellt heraus, dass dies einen positiven Einfluss auf die Prädiktionsgüte hat. Für ein spezifisches Fahrzeug werden 40 % der nächsten Abschnitte mit einer Zuverlässigkeit von 60 % korrekt bestimmt.

Für die Trajektorien von mehreren Fahrern erstellt Monreale [177] Entscheidungsbäume. Aus den Trajektorien werden Muster über ein Gitternetz extrahiert, die zudem eine ähnliche Reisezeit aufweisen. Für die Region wird ein Entscheidungsbaum erstellt und die nächsten Ziele anhand dessen identifiziert.

Jeung [178] entwickelt in Anlehnung an [174] eine Methodik, um für einen längeren Zeitraum die wahrscheinlichste Route zu identifizieren. Dabei werden die Übergangswahrscheinlichkeiten mit der Maximum Likelihood Methode für alle Segmente in einem Suchraum abhängig von der Startposition und der Reisezeit bestimmt. Dabei werden über mehrere Segmente hinweg die Abbiegewahrscheinlichkeiten multipliziert. Da hierbei der komplette Raum durchsucht wird, ist die Methode rechenintensiv. Daher wird ein Greedy-Algorithmus angewandt, bei dem schrittweise die Folgezustände bestimmt werden. Über zusätzliche Metriken, unter anderem mit der Berücksichtigung von Abbiegewinkeln, werden ähnlich gute Ergebnisse wie bei der Maximum Likelihood Variante erreicht. Bei einer Reisezeit von 10 min liegt die Abweichung mit der Methode nach [174] bei ca. 18 km, mit dieser Variante liegt der Fehler bei ca. 5 km.

Gidofalvi [179] sucht im Vergleich zu [178] Subroutinen, also durchgängige Straßensegmente, die häufig befahren werden. Dies erfolgt über ein Mindestmaß an in sich geschlossenen und wiederkehrenden Routen. Anhand der Routen werden Übergangswahrscheinlichkeiten bestimmt. Die Ergebnisse sind im Vergleich zu [178] um ca. 10 % besser.

Aus über 250 000 Fahrten von 357 Taxis erstellt Mao [180] eine Karte mit POIs. Dafür wird ein gewichteter Page-Rank-Algorithmus verwendet. Die Punkte werden als eine Zelle im Gitternetz beschrieben. Zur Prädiktion der nächsten Zellen werden die Übergangswahrscheinlichkeiten mit einem Markov Modell erstellt, das zudem mit der POI-Karte gewichtet wird. Durch die Gewichtung verbessert sich die Methode um 5 %.

Eldaw [181] wählt einen kollaborativen Ansatz für mehrere Fahrer. Insgesamt 8 154 Tage aufgeteilt auf 38 Nutzer werden GPS-Daten gesammelt. Auf Basis der Daten werden Cluster über den DBSCAN-Algorithmus gebildet. Der Schwerpunkt des Clusters wird bestimmt und in ein Gitternetz überführt. Zudem wird die Tageszeit mitberücksichtigt. Die Zeiteinheiten für die Aufenthalte in den Zellen werden dafür berechnet. Die Cluster werden mit der Zeitabhängigkeit in einen Suffixbaum eingebettet. Eldaw führt einen Vergleich von Kollektiv und Einzelnutzer durch und stellt fest, dass das Modell des Kollektivs sehr gute, nahezu vergleichbare Ergebnisse liefert.

Da die meisten Algorithmen rein auf Daten basieren, besteht die Schwierigkeit, den richtigen Weg zu präzisieren, wenn keine oder nur unzureichend viele Daten vorhanden sind. Zhou [182] wählt daher ein hybrides Modell, das dynamisch Referenztrajektorien wählen kann, wenn diese zu Verfügung stehen. Über den Aufzeichnungszeitraum kann sich das System kontinuierlich verbessern.

Huo [183] bestimmt wie [178] die wahrscheinlichste Route über Abbiegevorgänge. Zudem berücksichtigt er Aufenthaltsorte während der Fahrt. Der Algorithmus kann mit auftretenden kreisläufigen oder sich überlappenden Straßensegmenten umgehen. Falls keine Daten vorhanden sind, werden Straßenkrümmung und Abbiegewinkel herangezogen. Die Variante basiert auf einem Backtracking-Algorithmus und ist genauer als die Ergebnisse von [178].

Um das Verhalten von vollständigen Strömen von Fahrzeugen in einer Stadt zu beschreiben wählt Zhang [184] einen Deep-Learning-Ansatz. Hierbei wird der Übergang von Fahrzeugströmen zwischen Zellen im Gitternetz betrachtet. Mit Deep Residual Learning können extrem mehrlagige, faltende neuronale Netze mit über 1 000 Ebenen erstellt werden, die gute Resultate im Bereich der Bilderkennung und Segmentierung liefern. Zusätzlich werden Attribute wie Wetter und Events mitberücksichtigt. Die Methode übertrifft die Ergebnisse von vergleichbaren Zeitreihen-Prognose-Modellen und neuronalen Netze zur Prädiktion von Menschenströmen.

Einen hybriden Ansatz aus individueller und kollektiver Prädiktion wählt auch Trasarti [185]. Über mobile Geräte erfolgt eine kontinuierliche Datenerfassung von GPS-Trajektorien. Aus den Daten werden ähnliche Profile über ein Distanzmaß extrahiert. Von 5 000 Nutzern werden 326 000 Trajektorien gesammelt. Ein Algorithmus sucht für eine neue Bewegung das ähnlichste bereits aufgezeichnete Profil aus dem persönlichen oder allgemeinen Datensatz und prädiziert über den Zeithorizont die Position. Der hybride Ansatz schneidet besser ab als der kollektive Ansatz und als die Referenz [177].

Tabelle 4.3 : Literaturübersicht Trip-Prädiktion

Quelle	Jahr	Methode	Datenbasis	Abstaktion	Nutzer
[174]	2003	Fahrwahrscheinlichkeit	-	Segmente	I
[173]	2004	Bewegungsfunktion	-	-	I
[165]	2006	-	-	Segmente	I
[175]	2008	Markov Modell	27 000 Fahrten	Segmente	I/K
[167]P	2009	Straßenklassen	-	Segmente	I
[176]	2009	VOMM + Zeit	ShanghaiGrid	Segmente	K
[177]	2009	Decision Tree	4 500 Fahrten	Raster	K
[178]	2010	Maximum Likelihood / Greedy	> 100 000 Fahrten	Segmente	K
[168]P	2011	Straßenklassen + Tabelle	-	Segmente	I
[179]	2011	Präfixbaum	17 000 Fahrten	Segmente	K
[169]P	2011	Dyn. Kartenattribute	-	Segmente	I
[170]P	2012	Dyn. Umgebungsdaten	-	Segmente	I
[180]	2012	Gewichtetes Markov Modell	> 250 000 Fahrten	Raster + POI	K
[181]	2013	DBSCAN + Suffixbaum	Nokia	Raster	I/K
[171]P	2013	Fahrwahrscheinlichkeit	-	Segmente	I+K
[182]	2013	Dyn. Referenztrajektorien	SGP Taxi	Raster	K
[91]	2015	Straßenklassen	-	Segmente	I
[183]	2016	Backtracking	> 18 000 Fahrten	Segmente	K
[172]P	2016	Wahrscheinlichkeit + Lernen	-	Segmente	I+K
[184]	2016	Deep Residual Learning	TaxiBJ & BikeNYC	Raster	K
[185]	2017	Distanzmaß	> 326 000 Fahrten	-	I+K

P = Patent, K = Kollektiv, I = Individuell

4.1.5 Zusammenfassung und Abgrenzung

Eine zusammenfassende Übersicht zu den einzelnen Quellen und eine Einordnung in die relevanten Bereiche ist in Tab. A.1 dargestellt. In den letzten Jahren wurden viele Möglichkeiten aufgezeigt, wie Daten für die Prädiktion der Mobilität verwendet werden können. Diese haben zum Teil unterschiedliche Zielsetzungen; die meisten Methoden und Ansätze sind jedoch übertragbar. So können Methoden zur Prädiktion des Fahrtziels oder der Route auch auf den Bereich der Trip-Prädiktion angewendet werden.

Für diese Arbeit ist, wie in Kapitel 2.3.1 vorgestellt, eine Prädiktion für die nächsten Straßen-segmente gemäß einem elektronischen Horizont von Bedeutung. In den dargestellten Patenten wird die Zielsetzung hinter dieser Anwendung gut beschrieben. Zudem werden die Zielfunktionen sowie mögliche Architekturen erläutert. Allerdings wird kaum auf die zu Grunde liegende Methodik eingegangen. Die Berücksichtigung von Fahrwahrscheinlichkeiten kann, wie in den vorherigen Abschnitten dargestellt, auf verschiedene Weise realisiert werden.

Die vorgestellten Methoden sind meist nicht direkt miteinander vergleichbar, weil diese auf unterschiedlichen Datenbasen trainiert werden. In manchen Publikationen werden vergleichsweise kleine Datensätze herangezogen. In den meisten Fällen werden entweder die Daten des Kollektivs oder speziell die eines individuellen Fahrers betrachtet, bzw. diese voneinander getrennt untersucht. Potential zeigt jedoch eine kombinierte Betrachtung. Falls noch keine persönliche Aufzeichnung eines Fahrers vorhanden ist, kann das Verhalten von mehreren Fahrern als Referenz dienen.

Die eingesetzten Methoden sind mit wenigen Ausnahmen Methoden des maschinellen Lernens. Es zeigt sich eine Vielfalt an Möglichkeiten, mit denen diese Fragestellung angegangen werden kann. Markov Modelle in den unterschiedlichen Ausprägungsstufen zeigen dabei gute Ergebnisse. In den letzten Jahren wurden vor allem im Bereich der Zielprädiktion verstärkt Formen von neuronalen Netzen eingesetzt, die hohes Potential zeigen.

Diese Arbeit grenzt sich wie folgt vom Stand der Forschung ab: Grundlage der Untersuchung stellen die in Kapitel 3.1.1 beschriebenen Flottenversuche. Im Vergleich zu den vorgestellten Projekten handelt es sich um einen der größten Datensätze. Ebenso spannt der Untersuchungsraum Oberbayern sowohl städtische Zonen wie auch ländliche Gebiete auf. Die Größe der Datenbasis ermöglicht außerdem eine zeitliche Betrachtung, ohne den Datensatz zu sehr einzugrenzen.

Der Abstraktionsgrad wird wie in vergleichbaren Arbeiten auf Straßensegmente gesetzt. Die Vorverarbeitung der GPS-Daten erlaubt so eine deutliche Reduzierung der Datenmenge ohne für die Aufgabenstellung negative Einschränkungen. Dabei ist es das Ziel, die vorausliegenden Streckensegmente vorherzusagen. Die Prädiktion des Fahrtziels, bzw. der vollständigen Route ist nicht primär entscheidend.

Ziel dieser Arbeit ist eine kombinierte Betrachtung von Fahrwahrscheinlichkeiten des Kollektivs und den individuellen Präferenzen für die Wahl der Streckensegmente. Dabei sollen möglichst wenige nutzerspezifische Informationen für die Anpassung der Modelle erforderlich sein. Neben Einzelpersonen werden auch größere Nutzergruppen mit einer ähnlichen Routenwahl untersucht. Der Unterschied zwischen den beiden Varianten und der Vorteil werden genau herausgearbeitet.

Für die Aufgabenstellung wird als Methode für die Prädiktion des Trips, mit einer Abstraktion auf Straßensegmente, ein neuronales Netz untersucht. Da diese Form des maschinellen Lernens im Bereich der Zielprädiktion gute Ergebnisse aufweist, wird das Prinzip auf diese Fragestellung angewendet. Die gewählte Methode wird mit klassischen Verfahren verglichen.

4.2 Methodik

Im folgenden Abschnitt wird auf die gewählte Methodik eingegangen. Es wird der Systemansatz vorgestellt, anschließend werden die einzelnen Schritte daraus hervorgehoben.

4.2.1 Systemansatz

In Kapitel 2.3.2 wurde die Forschungsfrage und der Lösungsansatz zur Trip-Prädiktion bereits kurz vorgestellt. Das Ziel hierbei ist die Vorhersage des wahrscheinlichsten Pfades, den der Fahrer wählen wird. Mit einem datengetriebenen Ansatz auf Basis einer Vielzahl von Fahrten mehrerer Fahrer, die anonymisiert verarbeitet werden, erfolgt eine Prädiktion. Zusätzlich können weitere Eingangsgrößen die Vorhersagegenauigkeit erhöhen. Über nutzer- oder nutzergruppenspezifische Informationen wird das allgemeine Modell auf den Anwender zugeschnitten.

In Abb. 4.2 wird das Verfahren der Trip-Prädiktion vorgestellt. Der Ansatz wird in vier Schritte unterteilt. Diese werden in den folgenden Kapiteln detailliert betrachtet.

In Kapitel 4.2.2 wird das verwendete Modell vorgestellt. Ein End-to-End Memory Network (MEM) als eine Form von rekurrenten neuronalen Netzen dient in dieser Arbeit dazu, die Wahrscheinlichkeit des nächsten Streckenabschnitts auf Basis zuvor gefahrener Routen zu bestimmen. Das Modell wird in einem End-to-End Verfahren mit Ein- und Ausgangsdaten des selben Typs trainiert. Das Netz berücksichtigt die zeitlichen Zusammenhänge der bisher gefahrenen Streckenabschnitte. Zudem wird in diesem Kapitel kurz auf das Vergleichsmodell eingegangen. Hierfür wird ein VOMM verwendet, das die Tageszeit als weiteren Eingangsparameter berücksichtigt.

Alle gewählten Eingangsparameter werden in Kapitel 4.2.3 erläutert. Die Untersuchung erfolgt auf der Grundlage von gematchten GPS-Informationen. Die aufgezeichneten Informationen der Fahrt werden auf die befahrenen Straßensegmente reduziert. Für die Modularität wird eine gebietsweise Betrachtung gewählt. Ebenso ist die Tageszeit ein wichtiger Eingangsparameter für das Modell. Des Weiteren wird die Parametrierung des Modells vorgestellt.

Neben dem allgemeinen Modellansatz wird ein nutzerindividueller Ansatz verfolgt. Dieser wird in Kapitel 4.2.4 aufgegriffen. Hierbei wird das allgemeine Modell in einem kontinuierlichen Lernprozess für den Nutzer verbessert. Die Modelle werden bei einer fehlerhaften Prädiktion mit

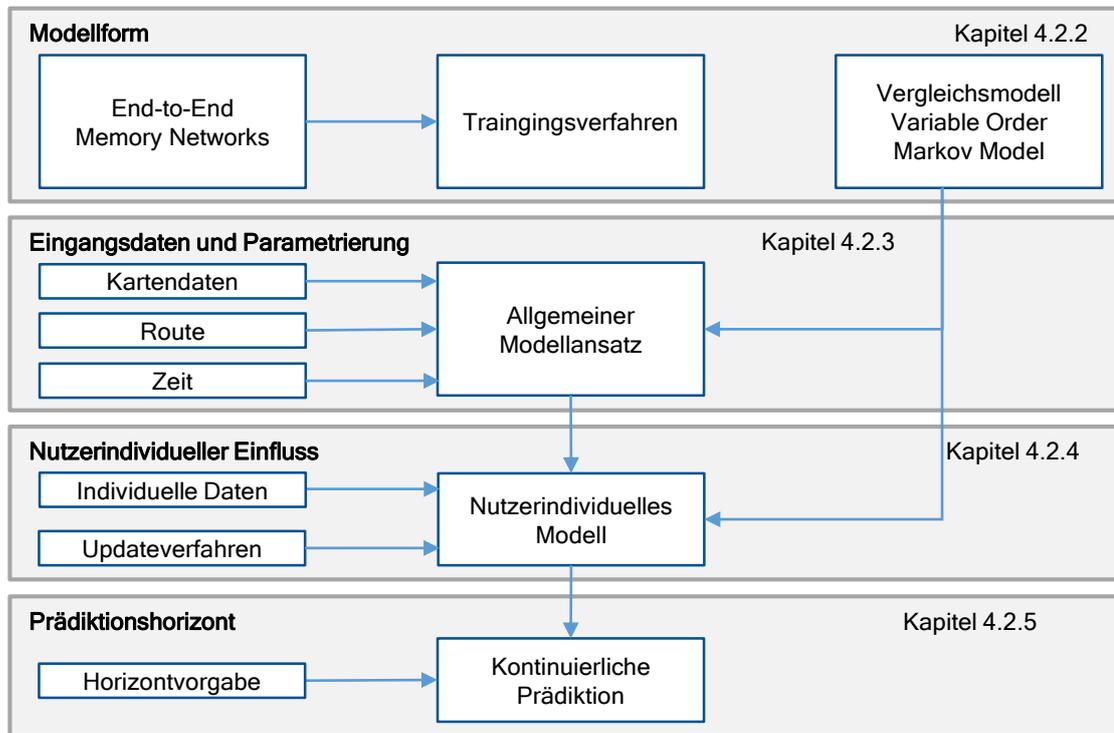


Abbildung 4.2 : Ansatz der Trip-Prädiktion

den korrekten Daten trainiert. Somit findet eine Anpassung an das spezifische Verhalten statt und die Güte der Prädiktion steigt.

Abschließend wird in Kapitel 4.2.5 das Verfahren des Prädiktionshorizonts vorgestellt. Der Horizont kann dabei nach der Zahl der Streckensegmente oder der Fahrstrecke aufgesetzt werden. Der Horizont der kontinuierlichen Prädiktion ist dabei an die spezifische Fahrzeuganwendung anzupassen.

4.2.2 Modellform

In diesem Abschnitt wird auf das verwendete Modell eingegangen. Anschließend wird der Algorithmus zum Trainieren des Modells vorgestellt. Zum Abschluss wird das verwendete Referenzmodell kurz dargestellt.

End-to-End Memory Networks

Als *Memory Network* [186] wird ein rekurrentes neuronales Netz mit einem möglichen externen Datenspeicher verstanden. Dieser Speicher ermöglicht, dass auch längerfristige Informationen von dem neuronalen Netz berücksichtigt werden können. Der Einsatzbereich liegt hierbei im Bereich von *Questions and Answers* (Q&A) [187]. Aus einer Reihe von aufeinander aufbauenden Zusammenhängen wird versucht, die richtige Antwort auf eine Frage zu finden. Außerdem können diese Modelle als sogenannte Chat-Bots eingesetzt werden. Dabei wird aus längeren Unterhaltungen auf den Kontext geschlossen und dann entsprechende Aussagen generiert.

Ebenso können andere neuronale Netze diese oder ähnliche Aufgaben erfüllen. Bspw. sind hier Long-short-term-memory-Netzwerke (LSTM) oder rekurrente neuronale Netze zu nennen. Weston [186] beschreibt, dass deren Gedächtnis allerdings mit den Informationen, die durch die Gewichte und verborgenen Zustände im Netz abgebildet werden, typischerweise zu klein bzw. zu stark miteinander verknüpft sind. Daraus entsteht ein Nachteil bei der Betrachtung von längerfristigen oder aufeinander aufbauenden Informationen, wie bspw. vollständigen Unterhaltungen. Das MEM und auch die Neural Turing Machine (NTM) [188] unterliegen diesen Beschränkungen nicht. Ein externer Speicher findet hierbei Einsatz. Dieser ermöglicht es, nahezu unbegrenzte Informationen zu hinterlegen und über Lese- und Schreibprozesse darauf zuzugreifen. Die NTM wird für abstrakte Operationen verwendet. Der Einsatz von MEM wird für die Erstellung von

Schlussfolgerungen auf aufeinander aufbauende Informationen angestrebt. In der von Sukhbaatar [187] vorgestellten Form ist es möglich, das Modell mit vorgegebenen Ein- und Ausgängen derselben Art vollständig zu trainieren.

Nachfolgend wird das Modell nach Sukhbaatar [187] beschrieben. In Abb. 4.3 ist die Funktionsweise grafisch dargestellt.

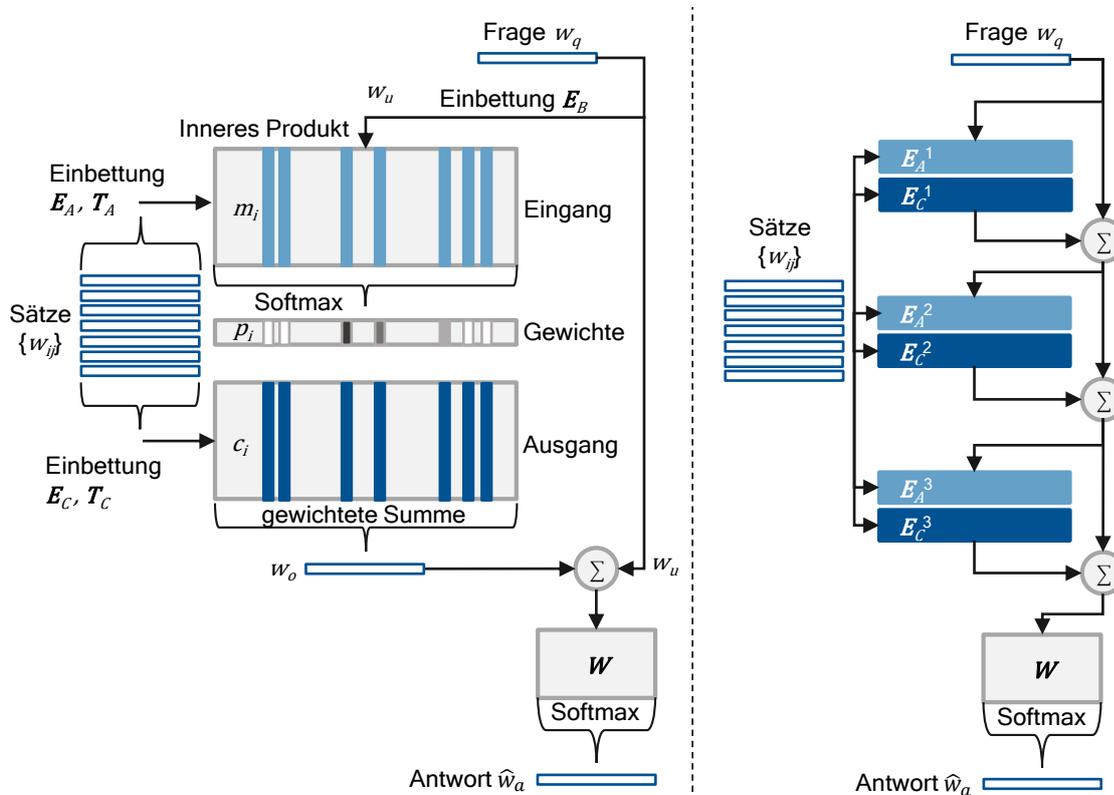


Abbildung 4.3 : Links: Einlagiges Netzwerk, Rechts: Dreilagiges Netzwerk nach Sukhbaatar [187]

Die Eingangsdaten stellen eine Abfolge von Sätzen w_{ij} dar, die in einem externen Speicher hinterlegt werden. Auf eine Frage w_q findet das Modell die wahrscheinlichste Antwort \hat{w}_a . Alle möglichen Wörter, welche die Sätze, Fragen und Antworten bilden, sind in einem gemeinsamen Verzeichnis V abgebildet.

Bei der Prädiktion von \hat{w}_a werden die Sätze in die Speichervektoren m_i mit der Einbettung E_A überführt. In diesem Fall bezeichnet eine Einbettung die Abbildung der diskreten Wörter des Verzeichnisses in einen kontinuierlichen Vektorraum, was einer Gewichtungsmatrix gleicht. Dies geschieht ebenso für die Frage w_q mit der Einbettung E_B um den inneren Zustand w_u zu erhalten. Um die zeitliche, aufeinander aufbauende Reihenfolge der Sätze zu berücksichtigen, erfolgt eine Encodierung über die Matrizen T_A und T_C , welche die zeitlichen Informationen hinterlegen.

$$m_i = \sum_j E_A w_{ij} + T_A(i) \quad (4.1)$$

$$c_i = \sum_j E_C w_{ij} + T_C(i) \quad (4.2)$$

Über das innere Produkt von w_u und m_i und der Softmax-Funktion [189, S. 81] wird der Wahrscheinlichkeitsvektor p_i der Eingänge bestimmt.

$$p_i = \text{Softmax}(w_u^T m_i) \quad (4.3)$$

Die Softmax-Funktion stellt dabei eine Generalisierung der Sigmoid-Funktion dar und wird als Klassifizierung verwendet. Damit kann eine Wahrscheinlichkeitsverteilung über kategoriale Variablen erstellt werden [189, S. 184].

$$\text{Softmax}(x_i) = \frac{\exp x_i}{\sum_j \exp x_j} \quad (4.4)$$

Mit der gewichteten Summe aus p_i und einer weiteren Einbettung E_C der Sätze in c_i wird der Speicherausgang w_o ermittelt.

$$w_o = \sum_i p_i c_i \quad (4.5)$$

Die letztendliche Prädiktion \hat{w}_a erfolgt über eine weitere Softmax-Funktion nach einer Gewichtung der Summe von w_o und w_u mit der Gewichtungsmatrix W .

$$\hat{w}_a = \text{Softmax}(W(w_o + w_u)) \quad (4.6)$$

Das Modell kann zudem verbessert werden, indem mehrere dieser Einbettungen E vorgenommen werden. Dies ist in Abb. 4.3 rechts dargestellt. Die kontinuierliche Darstellung der Sätze wird hintereinander in mehreren Stufen durchgeführt. Dabei hat jede Stufe eine eigene Einbettungsmatrix für E_A und E_C .

Trainingsvorgang

Der Vorgang für die Prädiktion der Antwort wird durch die Gl. 4.1-4.6 vollständig beschrieben. Der Trainingsvorgang des Modells besteht darin, die Einbettungs- und Gewichtungsmatrizen zu bestimmen. In einer stapelweisen Verarbeitung der Trainingsdaten werden E_A , E_B , E_C und W schrittweise verbessert [187].

Im Gegensatz zur Prädiktion, bei der die Sätze und die Frage als Eingang vorgegeben und vorwärts durch das Modell propagiert werden, findet beim Trainieren eine sog. Back-Propagation [190] statt. Hierbei wird zuerst vorwärts eine Antwort \hat{w}_a bestimmt und über eine Kreuzentropie mit Hilfe der korrekten Antwort w_a der Fehler ermittelt. Ziel ist es, diesen Fehler zu minimieren. Die Back-Propagation bezeichnet nun den Vorgang ausgehend von dem Modellausgang rückwärts durch das Modell zu gehen und die Gradienten zu bestimmen [189]. Dabei wird hauptsächlich auf die Kettenregel zurückgegriffen.

Den eigentliche Lernvorgang übernimmt ein stochastisches Gradientenverfahren [191], um eine große Menge an Trainingsdaten zu ermöglichen. In dem Batch-Processing wird zufällig ein Set an Trainingsdaten gezogen. Der damit bestimmte Gradient ist somit eine Näherung. Die Gewichte werden über den Gradienten angepasst. Die Lernrate η wird mit den Gradienten multipliziert und reguliert damit den Vorgang [189]. Über die Trainingsintervalle wird die Lernrate schrittweise reduziert.

Um bei mehrlagigen Netzwerken den Lernprozess zu vereinfachen werden die benachbarten Einbettungen von E_B zu E_A , E_A zu E_C , sowie von E_C zu W gleich gesetzt [187].

Referenzmodell

Als Referenzmodell dient ein VOMM. Ein vergleichbares Modell basierend auf dem ShanghaiGrid-Datensatz wird in [176] vorgestellt. Hierbei werden segmentweise alle Fahrten über diesen Abschnitt und die zuvor befahrenen Abschnitte betrachtet. Diese werden von diesem Segment an rückwirkend bis zu n_z Segmente aufgelistet. Über die Häufigkeit der auftretenden Routenabschnitte mit der entsprechenden Segmentreihenfolge kann die Wahrscheinlichkeit ermittelt werden, welcher Abschnitt als nächstes befahren wird. Über einen Grenzwert kann der notwendige Support, also die minimale Zahl an Fahrten, bei der diese Route gewählt wurde, die Auswahl einschränken.

Bei der Prädiktion werden die bereits zurückliegenden Segmente dem Modell zugeführt. Aus dem Datensatz wird die Wahrscheinlichkeit für das nächste Segment auf Basis dieser Route bestimmt. Falls keine aufgezeichneten Fahrten mit allen n_z Segmenten übereinstimmen, wird

die Ordnung auf weniger zurückliegende Segmente reduziert. In [176] und [158] wurde der Einfluss der Tageszeit der Fahrten auf das Modell mitberücksichtigt und als positiv bewertet. Daher wird die Möglichkeit der zeitlichen Begrenzung der Daten eingeführt. Falls keine Daten zu der untersuchten Tageszeit aufgezeichnet wurden, wird die zeitliche Begrenzung verworfen.

Zur Prädiktion eines längeren Horizonts, bis zu n_p Segmente, wird ein Suffixbaum aufgebaut. Dabei werden alle möglichen Abbiegevorgänge ausgehend von der aktuellen Position betrachtet. Bei dem iterativen Vorgang werden die Wahrscheinlichkeiten der nacheinander folgenden Abbiegevorgänge multipliziert, bis n_p Segmente erreicht sind. Um den Vorgang zu beschleunigen, werden Äste mit einer Wahrscheinlichkeit unter einer vorgegebenen Grenze nicht länger betrachtet.

4.2.3 Eingangsdaten und Parametrierung

Folgend werden die verwendeten Eingangsdaten kurz näher erläutert. Anschließend wird auf die Unterteilung der Datenbasis in Gebiete eingegangen. Abschließend wird die getroffene Modellparametrierung vorgestellt. Für die Untersuchung werden alle Fahrten der Datenbasis aus Kapitel 3.1.1 herangezogen. Für die Ergebnisbewertung wird ein Teil davon als Testdatensatz von dem Trainingsvorgang ausgeschlossen.

Eingangsdaten

Wie im Abschnitt 4.2.2 vorgestellt, benötigt das MEM eine Abfolge von Sätzen, die aus einzelnen Wörtern w bestehen. Ebenso bildet sich die Frage und die Antwort aus Wörtern, die alle in einem gemeinsamen Verzeichnis V angelegt werden.

Die Fragestellung zur Prädiktion des nächsten Straßensegments weist Ähnlichkeiten zu dem geplanten Einsatzgebiet der MEM auf, bei dem aus einer Reihe von aufeinander aufbauenden Informationen und einer Frage versucht werden soll, die korrekte Antwort zu bestimmen. In diesem Fall entsprechen die Informationen der bisher vom Fahrer zurückgelegten Strecke mit den entsprechenden Segmenten. Die Frage entspricht, in Abhängigkeit der aktuellen Position, der Richtung, in welche sich der Fahrer voraussichtlich bewegen wird. Als Antwort wird das nächste, wahrscheinlichste Segment erwartet.

Das Straßennetzwerk ist daher das Verzeichnis V mit den Segmenten als Wörter w . Im Vergleich zum MEM in [187] bestehen die Sätze allerdings nur aus jeweils einem Wort. Dennoch ist die Reihenfolge der Wörter von Bedeutung. Besonders dann, wenn Straßensegmente in beide Richtungen befahren werden können.

Weiterhin wird, wie im Referenzmodell, die Tageszeit mit berücksichtigt. Dies wird durch die Einbindung von 24 Stunden als weitere mögliche Wörter in das Verzeichnis realisiert. In den Speicher wird dann neben den Segmenten in der entsprechenden Reihenfolge die Uhrzeit angefügt.

Bei dem Trainingsvorgang gilt es zu berücksichtigen, wie viele Straßensegmente in den Speicher geladen werden. Um eine Vergleichbarkeit mit dem Referenzmodell zu gewährleisten, wird die Prädiktionsgüte des VOMM für das nächste Segment in Abhängigkeit der Zahl der zurückliegenden Segmente untersucht. Das Resultat ist in Abb. 4.4 dargestellt. Demnach steigt die Güte zuerst mit zunehmender Anzahl an Segmenten. Nach zehn Segmenten findet keine Verbesserung mehr statt. Folglich wird für den Lernvorgang des MEM und die Auswertung des VOMM die Zahl der zurückliegenden Segmente auf zehn festgesetzt.

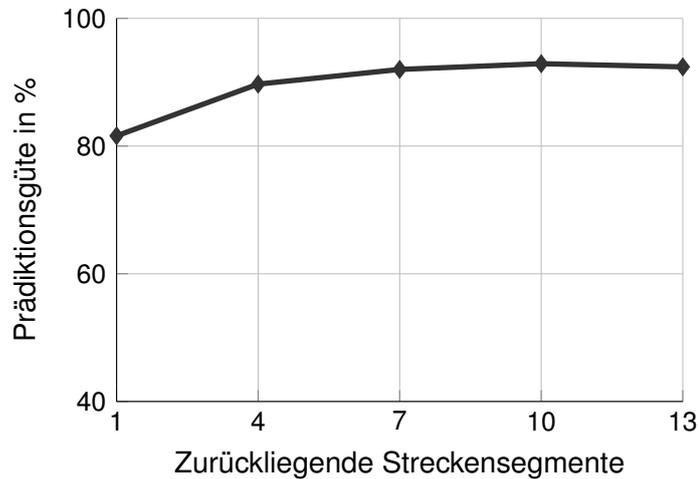


Abbildung 4.4 : Prädiktion des nächsten Streckensegments mit VOMM

Gebietsweise Betrachtung

Bei dem Einsatz des MEM im Bereich Q&A wird die Beziehung zwischen den einzelnen Wörtern und deren Auftreten im Satz berücksichtigt. Demzufolge werden Wörter, die in dem Kontext relevant sind, oder während des Lernvorgangs in einem vergleichbaren Bezug standen, für die Beantwortung der Frage von Bedeutung. Ein ähnlicher Zusammenhang besteht im Straßennetz. Die Beziehungen zwischen angrenzenden oder benachbarten Straßenabschnitten haben einen höheren Einfluss auf die Wahl des nächsten Segments, als bspw. Straßenabschnitte, die in weiter Entfernung liegen. Ebenso wird wie im vorherigen Abschnitt begründet, die Zahl der zurückliegenden Streckensegmente eingegrenzt.

Folglich können die Eingangsgrößen der Prädiktion auf lokale Gebiete reduziert werden. Im einfachsten Fall könnte für jedes Straßensegment ein Prädiktionsmodell zusammen mit den in den Datensätzen befahrenen, zurückliegenden Segmenten aufgebaut werden. Allerdings können auch lokal benachbarte Elemente, die explizit nicht in Verbindung mit dem betrachteten Segment befahren werden, bei der Prädiktion berücksichtigt werden. Daher wird im folgenden der Lernvorgang auf Gebiete mit mehreren Straßensegmenten begrenzt.

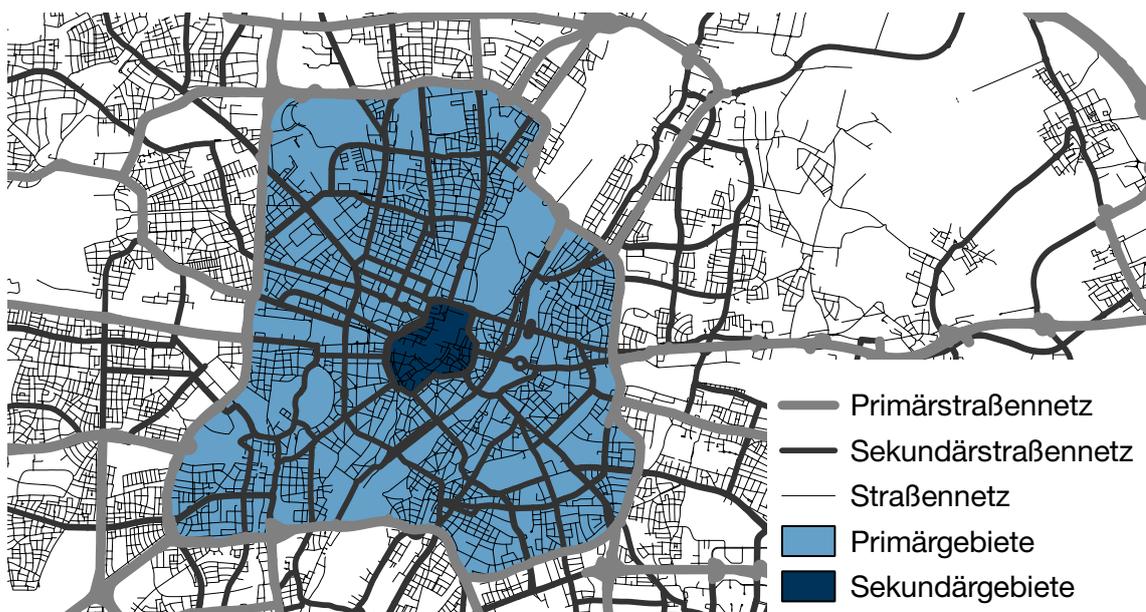


Abbildung 4.5 : Unterteilung in Gebiete nach [97] (Karte: OSM © ⓘ ⓘ)

Die Einteilung erfolgt anhand der OSM-Straßenklassen. Abb. 4.5 veranschaulicht die Gruppierung. Es erfolgt eine Aufteilung in Primär- und Sekundärgebiete [97]. Die Straßenklassen 11 bis 16 nach Tab. 3.2 definieren das Primärstraßennetzwerk. Die Gebiete, die durch das

Primärstraßennetz abgeschlossen werden, werden als Primärgebiete bezeichnet. Ebenso bilden die Straßenklassen 21 und 22 das Sekundärstraßennetz mit den umschlossenen Sekundärgebieten. Letztere bilden die Grundlage für die Modelle. Das Sekundärstraßennetz wird nicht zu einzelnen Sekundärgebieten zugeordnet. Daher werden die Sekundärstraßen des Primärnetzes gruppiert. Das Primärstraßennetz wird für Oberbayern geographisch in 50 weitere Gruppen unterteilt.

Dadurch entstehen ca. 1200 lokal gültige MEM, die sich durch die Zahl der zurückliegenden Elemente und dem zu präzisierenden Segment an den Gebietsgrenzen überschneiden. Durch die gebietsweise Betrachtung folgt auch ein Zeitersparnis beim Lernvorgang, weil die einzelnen Gebiete unabhängig voneinander parallelisiert berechnet werden können.

Modellparametrierung

Wie in [187] und in Abb. 4.3 illustriert, wird ein dreilagiges Netzwerk modelliert. Der Trainingsvorgang erfolgt in 100 Epochen mit einer initialen Lernrate $\eta = 0,01$. Zudem erfolgt eine Halbierung der Lernrate alle 25 Epochen. Die Einbettungen werden mit einem Zufallswert initialisiert. Die Dimension der Einbettungen wird individuell für jedes Gebiet in Abhängigkeit des Verzeichnisses V auf die Hälfte der Anzahl der Straßensegmente gesetzt. Die zeitliche Encodierung über T_A und T_C wird auf bis zu 10 Segmente begrenzt. Falls bei der Prädiktion Straßensegmente in den Speicher gerufen werden, die nicht in dem Verzeichnis hinterlegt sind, ersetzt diese ein Leersymbol.

4.2.4 Nutzerindividueller Einfluss

In Kapitel 3.1.1 wird die Zusammensetzung der Datenbasis beschrieben. Der Großteil der Fahrten, die auch für den Trainingsvorgang verwendet werden, stammt aus dem Taxiflottenversuch (Abb. 3.1). Durch die große Anzahl an Fahrten mit unterschiedlichen Start- und Endpunkten wird der Großraum München größtenteils abgedeckt. Die Taxis bilden daher eine Grundlage für eine Prädiktion ohne personalisierte Informationen.

Im Bezug auf personalisierte Informationen können Privatpersonen oder Nutzergruppen, wie bspw. Unternehmen, betrachtet werden. Abb. 4.6 zeigt die Start- und Endpunkte von Fahrten eines Taxis, bzw. von 44 Fahrzeugen eines Handwerksbetriebs, um den Unterschied zu demonstrieren. Bei den jeweils ca. 9000 Fahrten zeigt sich bei einem Taxi eine Konzentration der Punkte im Zentrum, innerhalb des mittleren Rings. In den Randgebieten herrscht eine nahezu gleichmäßige Verteilung.

Im Gegensatz zum Taxi zeigt sich bei 44 Fahrzeugen eines Unternehmens eine Konzentration an wenigen Standorten. Hierbei handelt es sich um den Unternehmenssitz und längerfristige Einsatzorte. Es ist ersichtlich, dass die angefahrenen Orte häufig in gleichen Gebieten liegen. Dadurch unterscheiden sich häufig befahrene Routenabschnitte von denen der Taxis.

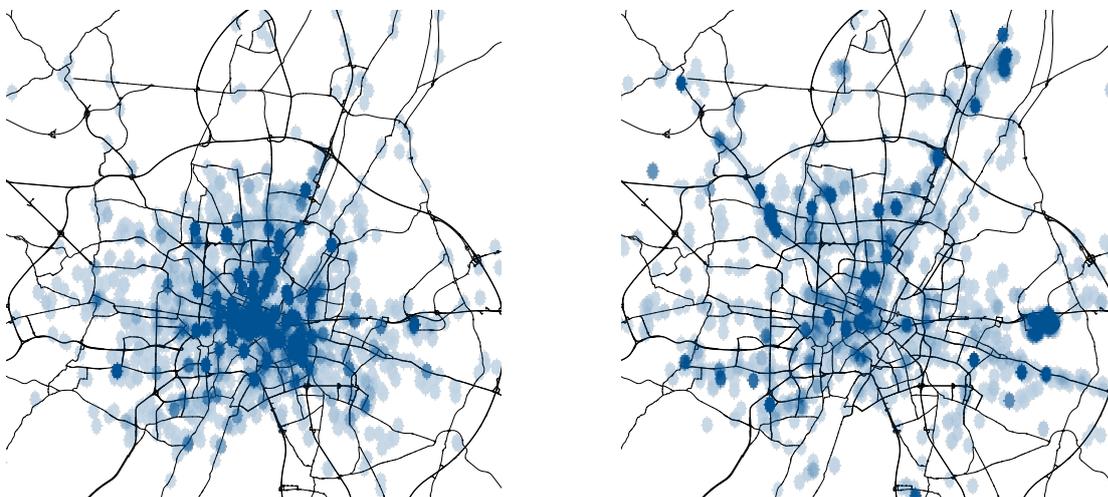


Abbildung 4.6 : Start- und Endpunkte: Links ein Taxi, Rechts ein Handwerksbetrieb (Karte: OSM © (i) (cc))

Um den nutzerindividuellen Einfluss bei der Prädiktion mit zu berücksichtigen, wird eine Anpassung des Modells durchgeführt. Wird während der Fahrt eine fehlerhafte Prädiktion detektiert, wird das lokale Modell mit den korrekten Daten erneut trainiert. Dazu wird die Lernrate erneut zu $\eta = 0,01$ gesetzt. Der somit anders befahrene Routenabschnitt wird dadurch verstärkt gewichtet. Das lokale Modell steht danach aktualisiert zur Verfügung und wird ausschließlich für den Nutzer oder die Nutzergruppe verwendet. Wird derselbe Abschnitt nach einer ähnlichen Route erneut befahren, führt das Update zu einer Priorisierung des zuvor befahrenen Segments.

4.2.5 Prädiktionshorizont

Die Prädiktion des Trips erfolgt sequentiell. Mit dem vorgestellten Verfahren wird jeweils das nächste Straßensegment prädiziert. Um einen längeren Prädiktionshorizont aufzuspannen, wird das Ergebnis der Prädiktion als Eingang für das darauffolgende Modell verwendet. Einen vergleichbaren Ansatz wählt [192] zur Prädiktion von Evakuierungen in Folge von nahenden Wirbelstürmen, bzw. Letham [193] als Empfehlungssystem auf Basis zuvor getroffener Entscheidungen.

Das Vorgehen ist in Abb. 4.7 exemplarisch dargestellt. Die initialen Informationen zu Beginn einer Fahrt setzen sich aus der aktuellen Position und der Fahrtrichtung zusammen (1). Auf dieser Basis erfolgt die Prädiktion des ersten Segments (2). Die nachfolgenden Segmente werden sequentiell prädiziert. Die Länge des Horizonts wird nach der Anzahl an Segmenten oder der resultierenden Distanz festgesetzt. Ist die Horizontlänge erreicht (3), beginnt die Fahrt.

Während der Fahrt wird die aktuelle Position mit der Prädiktion abgeglichen. Ist diese korrekt, wird entweder ein weiteres Segment oder der weitere Weg bis zur vorgegebenen Distanz prädiziert. Befindet sich das Fahrzeug nicht auf der vorhergesagten Strecke (4), ist die Prädiktion fehlerhaft. Anschließend erfolgt erneut eine initiale Prädiktion (5) auf Basis der zurückliegenden Streckensegmente.

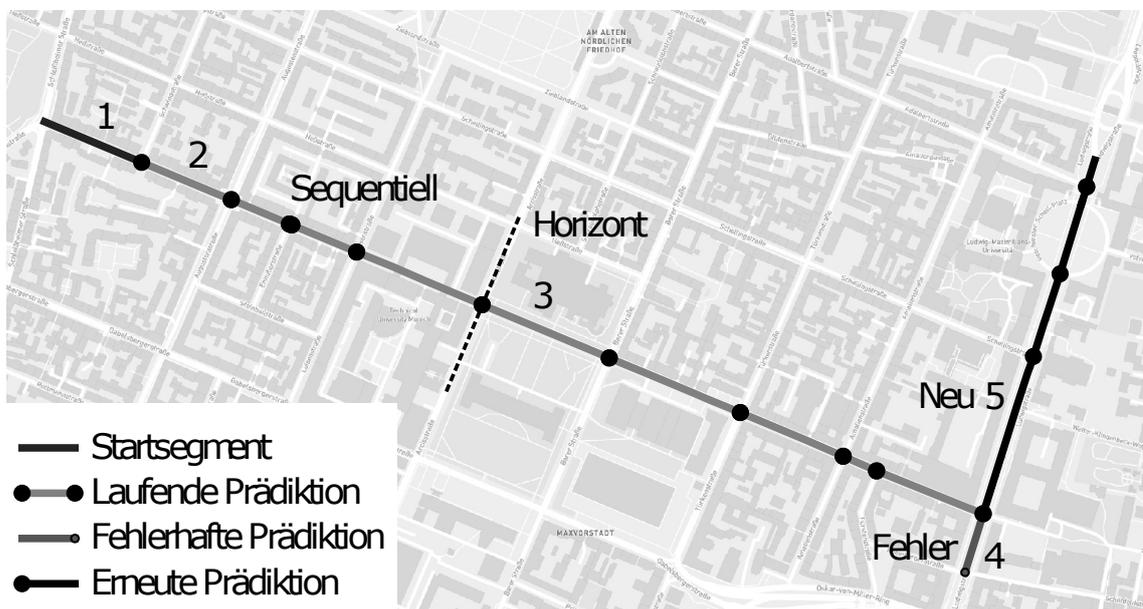


Abbildung 4.7 : Vorgehen bei der Prädiktion (Karte: OSM © ⓘ ⓘ)

Ebenso ist es möglich, wie auch im VOMM, mehrere mögliche Pfade aufzuspannen. Dies ist exemplarisch in Abb. 4.8 in einer Isodistanz-Darstellung veranschaulicht, bei der als Abbruchkriterium beispielhaft eine vorgegebene Distanz festgelegt wird. In einem identischen Vorgehen, wie in Kapitel 4.2.2 beschrieben, wird ein Suffixbaum auf den Ergebnissen der Prädiktion mit Hilfe des MEM erstellt. Aufgrund der Softmax-Funktion nach der Gewichtung in Gl. 4.6 können die Resultate der Berechnung für die einzelnen Straßensegmente als Wahrscheinlichkeiten aufgefasst werden.

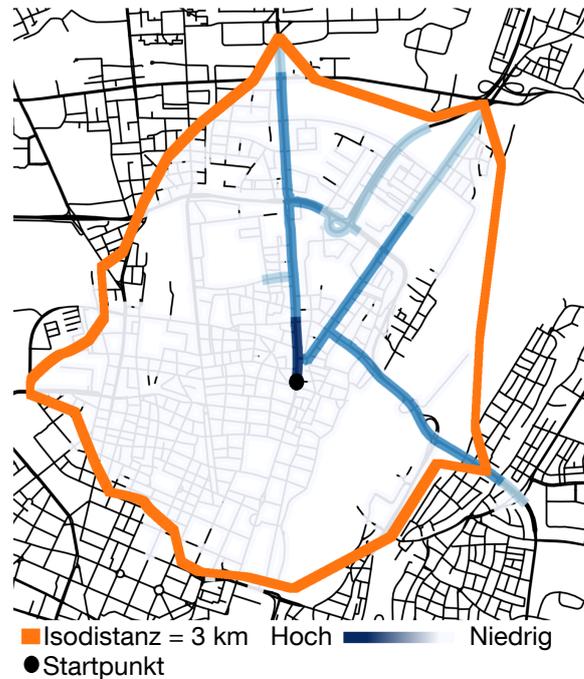


Abbildung 4.8 : Isodistanz mit mehreren wahrscheinlichen Pfaden nach [97] (Karte: OSM © ⓘ ⓘ)

4.3 Ergebnisse

Die Ergebnisse zur Trip-Prädiktion werden wie folgt untergliedert. Es wird der Testdatensatz und die Ergebnisse der Prädiktion für diese Fahrten vorgestellt. Anschließend wird die Prädiktionsgüte über den Horizont untersucht. Abschließend wird eine Einzelfahrt näher betrachtet.

4.3.1 Prädiktion während der Fahrt

Zur Bewertung der Prädiktion wird ein Testdatensatz von 1 224 Fahrten gewählt. Dieser wird nicht bei der Modellerstellung verwendet. Insgesamt gilt es 85 717 Segmente zu prädizieren. Dabei liegt die Zahl der Segmente pro Fahrt zwischen 26 und 448, mit durchschnittlich 84 Segmenten. Die Fahrten stammen, wie in 4.2.4 beschrieben, von einer Nutzergruppe mit 44 Fahrzeugen. Damit ist es auch möglich, den nutzerindividuellen Einfluss zu betrachten. Im Folgenden wird auf die Ergebnisse der Prädiktion mit dem Referenzmodell (VOMM), dem allgemeinen Modell (MEM) und dem nutzerspezifischen Modell (MEM Update) eingegangen.

In Abb. 4.9 wird die relative Zahl der Fehlprädiktionen in Bezug auf das Referenzmodell dargestellt. Bei 11 834 Segmenten sagt das VOMM einen anderen Weg voraus. Das allgemeine MEM erreicht einen um 5 % besseren Wert. Eine deutliche Verbesserung zeigt das um nutzerspezifische Informationen ergänzte Modell. Die Zahl der Fehlprädiktionen wird um über 17 % gesenkt.

Im Durchschnitt wird demnach beim Referenzmodell alle 7,24 Segmente eine erneute initiale Prädiktion benötigt. Das MEM liegt alle 7,62 Segmente falsch und das personalisierte Modelle nach 8,77 Segmenten.

Der deutliche Vorteil der nutzerspezifischen Anpassung des Modells im Vergleich zum Referenzmodell wird in Abb. 4.10 ersichtlich. Es wird über die zeitlich aufeinanderfolgenden Fahrten die fehlerhaften Prädiktionen im Vergleich zur Referenz dargestellt. Das nutzerunabhängige Modell zeigt einen linearen Verlauf. Folglich ist es über die Fahrten hinweg konstant besser als das Referenzmodell. Das nutzerspezifische Modell weist einen quadratischen Verlauf auf. Die Updates der Modelle erfolgen im Anschluss an die Fehlprädiktion und können direkt von der Anpassung profitieren. Der Verlauf zeigt, dass sich die Modelle kontinuierlich verbessern und das Verhalten der Nutzergruppe über die Zeit hinweg bestimmter vorhergesagt werden kann.

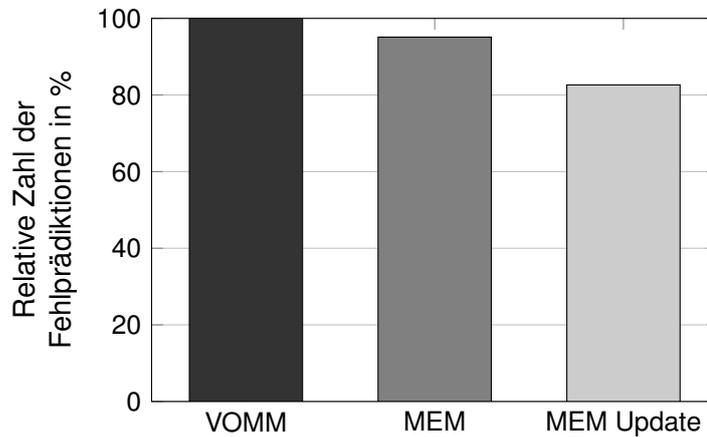


Abbildung 4.9 : Relative Zahl der Fehlprädiktionen im Vergleich der Modelle

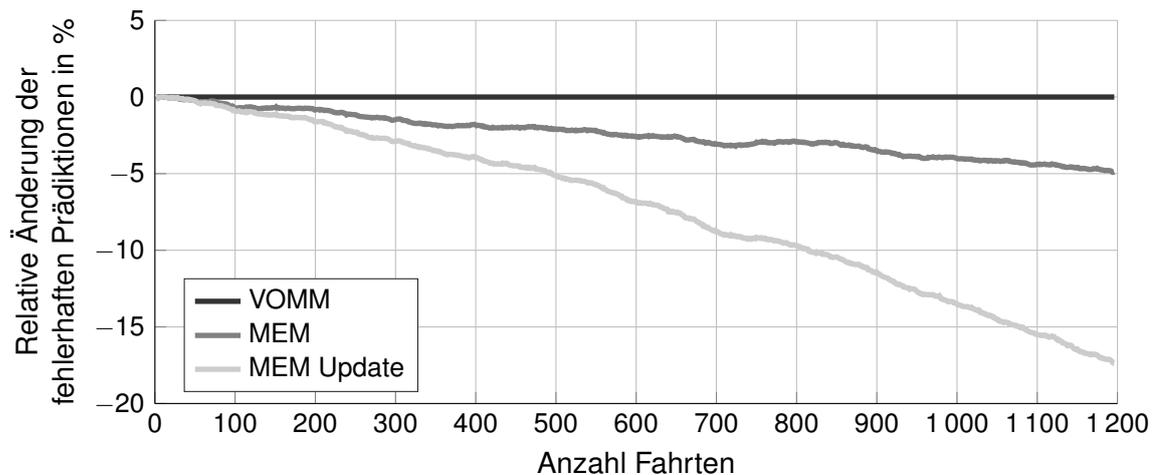


Abbildung 4.10 : Reduktion der Fehlprädiktionen im zeitlichen Verlauf.

Um den Anpassungsbedarf zu bewerten, wird die Häufigkeit und die Positionen der fehlerhaften Prädiktionen untersucht. Abb. 4.11 veranschaulicht die entsprechenden Orte. Hinterlegt sind dabei die bereits in Abb. 4.6 dargestellten Start- und Endpunkte der Nutzergruppe. Es zeigt sich, dass die Orte mit den Updatevorgängen mit den Regionen der häufigen Anfahr- und Halteorte und mit den zugehörigen Verbindungsstraßen korrelieren. Mehr als zwei Updates finden lediglich auf 5,3 % der von der Nutzergruppe befahrenen Segmente statt. Durch die Updates an Startpunkten und wichtigen Kreuzungen kann das nutzerspezifische Modell länger zusammenhängende Routenabschnitte identifizieren. Dies spiegelt sich auch in den Ergebnissen für den Prädiktionshorizont wider.

4.3.2 Prädiktionshorizont

Weiterführend wird der Prädiktionshorizont näher untersucht. Hierfür werden in den betrachteten Fahrten an beliebiger Position und an über 5 000 Stellen initiale Prädiktionen durchgeführt. Bei allen Untersuchungen werden zehn zurückliegende Segmente als Eingangsdaten verwendet. Der maximale Horizont beträgt zehn Segmente. Neben dem Referenzmodell werden das allgemeine MEM und die nutzerindividuelle Variante, die über die Durchführung hinweg angepasst wird, vorgestellt. Zudem zeigt eine weitere Variante (MEM nach Update) das Ergebnis der Prädiktion nachdem die Datenbasis vollständig auf den Nutzer angepasst worden ist. Dies entspricht der Situation, wenn alle 1 224 Fahrten in identischer Weise erneut durchgeführt werden, ohne dass ein weiteres Update erfolgt.

Abb. 4.12 zeigt das Ergebnis des Versuchs. Die Güte wird bezogen auf die Zahl der Segmente angegeben. Zwischen dem Referenzmodell und dem allgemeinen MEM sind im mittleren Bereich

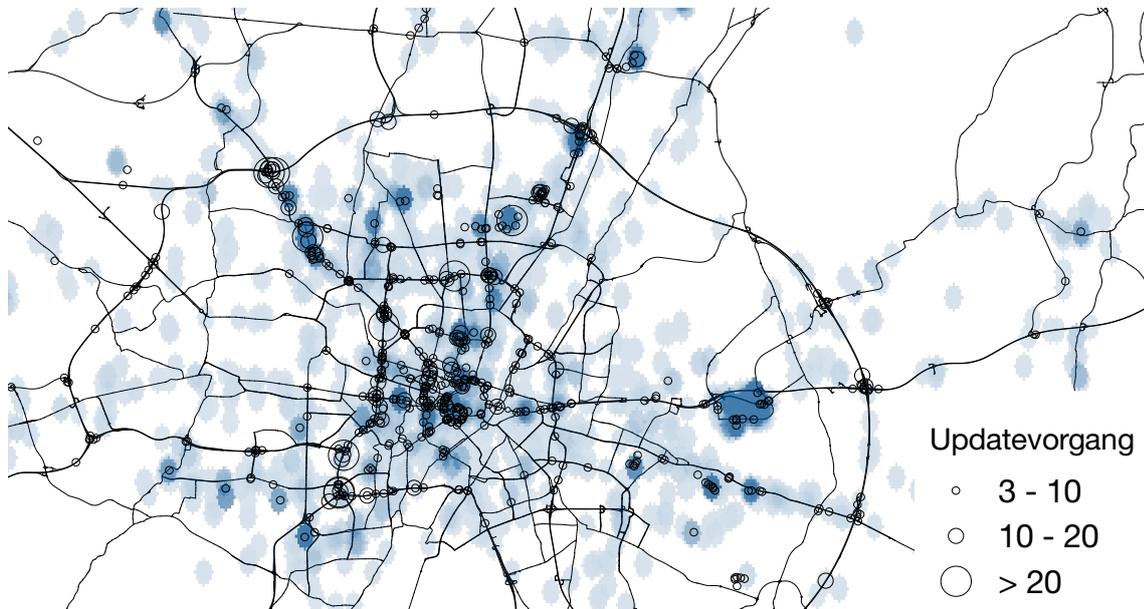


Abbildung 4.11 : Häufigkeit und Position der nutzerindividuellen Updates (Karte: OSM © ⓘ ⓘ)

Vorteile von ca. 2 % des letzteren ersichtlich. Das Modell MEM Update zeigt bereits für wenige Segmente einen Vorteil gegenüber der Referenz. Bei einem Horizont von fünf Segmenten übersteigt es den Referenzwert um ca. 10 %. Bei zehn Abschnitten liegt der Vorteil bei ca. 11 %. Das bereits angepasste Modell übersteigt den Wert des nutzerspezifischen Modells um bis zu 5 %. Gerade der Unterschied zu dem Referenzmodell macht deutlich, welches Potential in der Modellanpassung liegt.

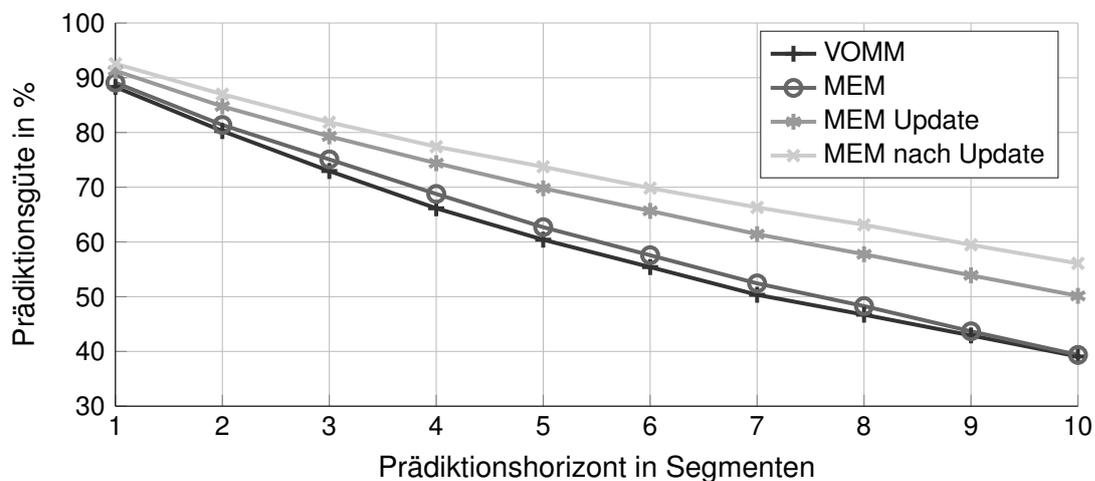


Abbildung 4.12 : Abhängigkeit der Prädiktionsgüte vom Horizont in Segmenten

Zudem wird der Horizont auf die Wegstrecke betrachtet. Dies ist in Abb. 4.13 bis 5 km dargestellt. Der abgebildete Wert muss zusammen mit der untersuchten Datenbasis betrachtet werden, weil die bei der Untersuchung verwendeten Fahrten unterschiedliche Längen aufweisen. Diesbezüglich ist die Prädiktionsgüte relativ zur Datenbasis in Abb. 4.14 dargestellt. Aus beiden Grafiken wird ersichtlich, dass alle Modelle bei einem kurzen Horizont sehr gute Vorhersagen treffen. Bereits ab 0,5 km wird allerdings der Vorteil der nutzerspezifischen Varianten deutlich.

Abb. 4.14 ist bezogen auf die Datenbasis und zeigt, dass die Güte der Prädiktion der angepassten Modelle nahezu unabhängig von der Straßensegmentlänge ist. Das weist darauf hin, dass Strecken mit längeren Segmenten, wie bspw. Autobahn- oder Überlandfahrten, mit gleicher Güte vorhergesagt werden wie Fahrten innerorts. Das allgemeine MEM sowie das Referenzmodell zeigen eine niedrigere Prädiktionsgüte bei Fahrten mit längeren Segmenten.

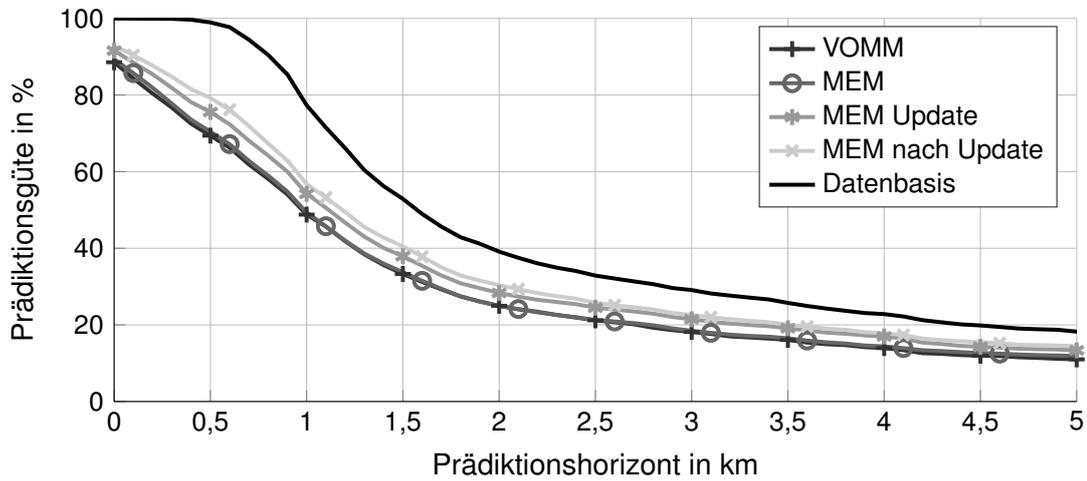


Abbildung 4.13 : Abhängigkeit der Prädiktionsgüte vom Horizont in km

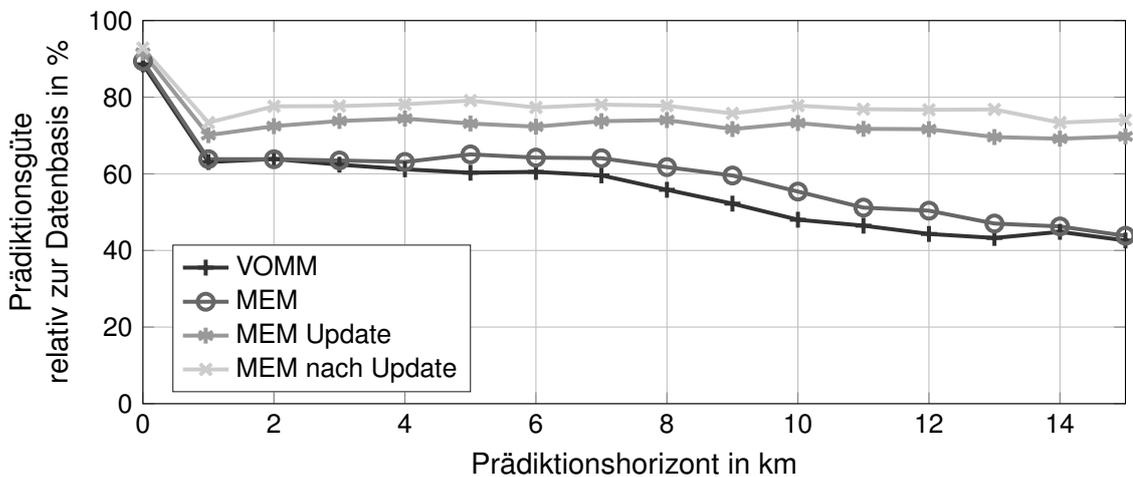


Abbildung 4.14 : Prädiktionsgüte bezogen auf die betrachtete Datenbasis

4.3.3 Beispielfahrt

Zuletzt wird eine Beispielfahrt veranschaulicht. Abb. 4.15 zeigt eine Fahrt, den Prädiktionshorizont und mögliche Fehler. Dabei wird das Referenzmodell mit dem personalisierten MEM verglichen. Dargestellt wird die tatsächliche Route und die Prädiktionsfehler der beiden Varianten.

Die Fahrt beginnt zentrumsnah, verläuft über den Autobahnring und endet in einem Vorort von München. Zu Beginn erfolgt jeweils eine initiale Prädiktion mit einem Segment und der Fahrtrichtung. Das VOMM liegt in diesem Gebiet mit der reduzierten Zahl der Eingangssegmente aus der aktuellen Fahrt in drei Fällen falsch. Das nutzerspezifische MEM kann die Fahrt zu Anfang nicht korrekt vorhersagen und zeigt in zwei Fällen eine Fehlprädiktion. Auf der weiteren Fahrt in Richtung Autobahn liegt die Prädiktion bei beiden Varianten zu großen Teilen richtig. Bei der Auffahrt zur Autobahn kann das personalisierte Modell die Richtung bereits richtig vorhersagen. Das Referenzmodell liegt hier falsch. Ebenso kann das angepasste Modell die Ausfahrt zum häufig angefahrenen Ziel korrekt bestimmen, während das VOMM den weiteren Weg auf der Autobahn Richtung Flughafen sieht. Letzteres ist ein von den Taxis häufig angefahrenes Ziel und daher bei den allgemeinen Modellen priorisiert.

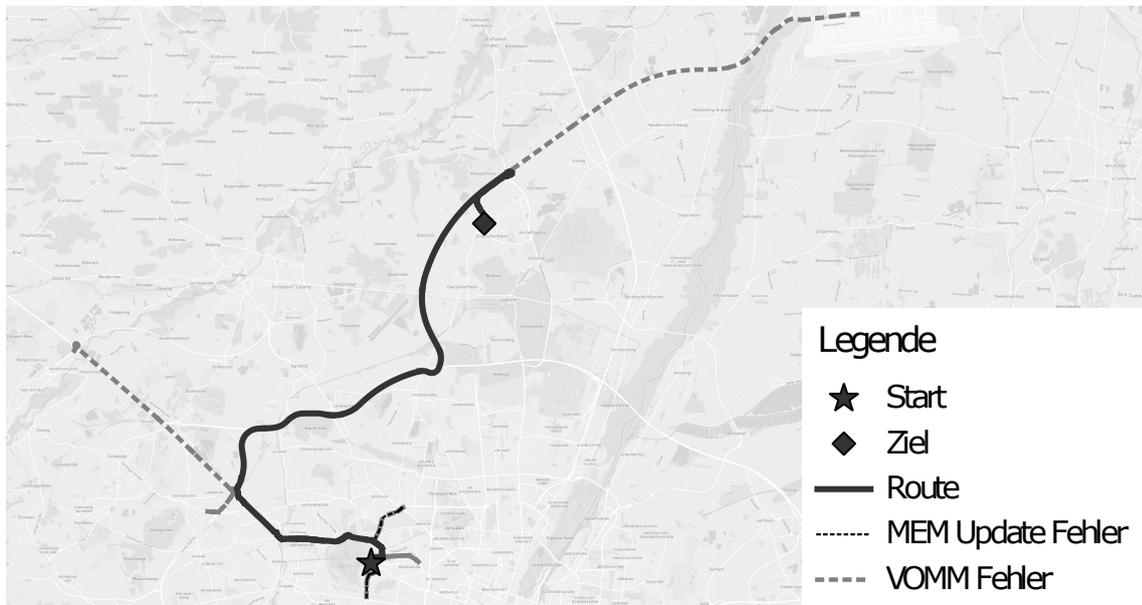


Abbildung 4.15 : Vergleich von VOMM und MEM an einer Beispielfahrt (Karte: OSM © ⓘ ⓘ)

4.4 Bewertung

Es folgt eine Bewertung und Diskussion zur Trip-Prädiktion und den Beiträgen dieser Arbeit. Dabei wird zu Beginn auf den untersuchten Stand der Forschung und die Datenbasis eingegangen. Anschließend wird die angewandte Methodik diskutiert. Abschließend findet eine Bewertung der Ergebnisse statt.

Stand der Forschung

Die vorgestellten Quellen geben einen ausführlichen Überblick über die bisherige und aktuelle Forschung auf dem Gebiet. Darüber hinaus werden Patente mitaufgeführt, die allerdings wenige Details über die Umsetzung der Methoden vorstellen. Eine inhaltliche Bewertung und Abgrenzung dieser Arbeit wurde in Kapitel 4.1.5 durchgeführt. Es kann festgehalten werden, dass ein breites Forschungsfeld existiert. Die Methoden verfolgen dabei unterschiedliche Aufgaben und Ziele. Die Abgrenzung zum Stand der Forschung hat allerdings ergeben, dass weiterhin ein Spielraum für fortführende Untersuchungen besteht. Diese wurden in dieser Arbeit aufgegriffen.

Datenbasis

Im Vergleich zu den vorgestellten Untersuchungen aus dem Stand der Technik ist die in dieser Arbeit verwendete Datenbasis umfangreich. Der breite Untersuchungsraum in Oberbayern schließt sowohl städtische als auch ländliche Gebiete mit ein. Der Datensatz zeichnet sich durch eine Vielzahl an Taxis aus. Die präferierten Routen können zum Teil durch diese Flotte geprägt sein. So werden neben POIs bevorzugt auch Taxistände oder deren Fahrzeugdepots angefahren. Allerdings wird durch den kontinuierlichen Betrieb der Fahrzeuge eine Fülle an zusammenhängenden Routenabschnitten im Großraum München befahren, die auf Grund der Kundenaufträge beliebig verteilt sein können. Zusammenfassend kann festgehalten werden, dass das Hauptstraßennetz und das Zentrum von München sehr gut im Datensatz abgedeckt sind. Dies trifft nicht für alle Nebenstraßen oder Wohngebiete zu. Zwar wurden diese grundsätzlich im Laufe der Flottenversuche befahren, manche davon weisen jedoch zu wenige Datensätze auf, um statistisch aussagekräftig verwendet zu werden. Diese wurden für die Betrachtung der Ergebnisse ausgeschlossen.

Der Abstraktionsgrad wird auf Straßensegmente gesetzt. Dies ist eine vertretbare Einschränkung der Genauigkeit, weil es das Ziel dieser Arbeit ist, einen mittelfristigen Prädiktionshorizont zu erzeugen. Trotz einer umfangreichen Vorverarbeitung der Daten können immer wieder einzelne Fehler auftreten. So wird durch das Mapmatching eine zusammenhängende Folge an Straßensegmenten erzeugt, die von den im Kartendatensatz angegebenen Kosten für das

Routing-Verfahren abhängen. Sonderregelungen für Fahrzeuge des öffentlichen Personennahverkehrs erlauben zum Teil eine Befahrung von Einbahnstraßen in entgegengesetzter Richtung. Durch Anpassungen in dem Mapmatching-Verfahren werden derartige besondere Bestimmungen berücksichtigt. Weitere ungewöhnliche Routenplanungen oder Verkehrsverstöße können allerdings auf Grund der großen Datenmenge nicht ausgeschlossen werden.

Methodik

Die Zielsetzung und Themenstellung dieser Arbeit sieht eine breite Anwendergruppe sowohl an Privatpersonen, Nutzergruppen oder Fahrzeugflotten vor. Dabei ist die Prädiktion des Fahrtziels bzw. der vollständigen Route nicht von entscheidender Bedeutung. Dies wäre in vielen Fällen als konträr anzusehen, weil dies nur für Privatpersonen oder eingeschränkte Nutzergruppen, über die Speicherung von Fahrtzielen sinnvoll umsetzbar wäre. Zudem ist es das Ziel, möglichst wenige nutzerspezifische Informationen zu verwenden. Dies schließt mit ein, dass der Bezug zu den Fahrtzielen aus den entwickelten Modellen nicht mehr direkt herzustellen ist.

In dieser Arbeit wird daher ein kombinierter Ansatz mit allgemeinen Modellen und einer Personalisierung dieser entwickelt. Im Speziellen wird ein neuronales Netz untersucht. Die Methode des maschinellen Lernens hat sich in anderen Anwendungsgebieten, wie bspw. dem Einsatz im Bereich von Q&A bereits bewiesen. Das Funktionsprinzip wird auf diese Fragestellung angewendet, weil Parallelen in den Datensätzen, der Abfolge von diskreten Wörtern bzw. Straßensegmenten, gezogen werden können.

Ergebnisse

Die entwickelte Methode wird mit einer Weiterentwicklung eines klassischen Verfahrens verglichen. Bei der Modellierung werden die gleichen Eingangsdaten, ein identischer Testdatensatz und gemeinsame Basisparameter verwendet. Das MEM zeigt Vorteile im Bereich von ca. 5 % gegenüber der Referenz. Durch die breite Datenbasis wird eine Vielzahl unterschiedlicher Fahrtziele und Routen betrachtet. Ähnliche Trips werden in den Daten von dem Modell gefunden, aber bei der Wahl des jeweils nächsten Segments besteht in vielen Fällen keine eindeutige Lösung. Das VOMM-Referenzmodell bildet diese Unterschiede in einem stochastischen Prozess klar definiert ab. Das MEM lernt diese über den Trainingsvorgang. Dieses hat den Vorteil dabei weitere Informationen zu lernen, ohne die Datenbasis zu reduzieren. Dies wird durch eine Anpassung der Gewichte vorgenommen. Speziell ist dies der Fall bei der Berücksichtigung der Tageszeit aller Trips. Des Weiteren dienen durch die gebietsweise Betrachtung zusätzlich weitere Fahrten, die nicht direkt mit den betrachteten Segmenten in Verbindung stehen, zur Anpassung der Gewichte. Über die Beziehung der Segmente in dem Kartendatensatz, wie bspw. der geographischen Fahrtrichtung oder Abbiegevorschriften wird dies bestärkt. Die grundsätzliche Verteilung der Abbiegewahrscheinlichkeit bleibt allerdings bestehen, so dass nur ein leichter Vorteil herausgearbeitet werden kann.

Die Personalisierung zeigt einen merklichen Effekt mit einer Verbesserung von ca. 17 % des Ergebnisses. Dies ist auch in dem Testdatensatz für eine breite Nutzergruppe mit mehreren Fahrzeugen gegeben. Über den Updatevorgang wird das MEM erweitert. Der kontinuierliche Lernvorgang passt dabei die Gewichte an. Bei einzelnen Updatevorgängen führt dies noch nicht notwendigerweise dazu, das prädizierte Segment zu wechseln, wenn die Segmentfolge im allgemeinen Modell nur selten in dieser Weise befahren wird. Das allgemeine Modell dient immer als Basis und kann verwendet werden, wenn keine individuellen Daten vorhanden sind. Dies ist ein klarer Vorteil gegenüber rein nutzerbezogenen Ansätzen.

5 Fahrprofilprädiktion

Nachfolgend wird erst auf den Stand der Forschung eingegangen. Anschließend wird die angewandte Methodik erläutert. Daraufaufgehend werden die erzielten Ergebnisse vorgestellt und abschließend bewertet.

5.1 Stand der Forschung

In diesem Bereich widmet sich die Forschung verschiedenen Zielen, die mit einer Fahrprofilprädiktion ermöglicht werden. Ein häufiger Fall ist die Verwendung der Prädiktion zur Optimierung des Energiemanagements bei Hybridfahrzeugen. Die Hybrid Electric Vehicle (HEV) und Plug-In Hybrid Electric Vehicle (PHEV) schöpfen das größte mögliche Potential aus, wenn das erwartete Fahrprofil und die Fahrstrecke im Voraus bekannt sind. Des Weiteren können prädizierte Fahrprofile zur Prognose des Energieverbrauchs dienen. Ein weiteres Ziel ist es, durch die streckenbezogenen Profile den Verkehrsfluss oder den Umwelteinfluss zu bewerten und damit das Routing der Fahrzeuge anzupassen.

Bei der Betrachtung ist eine Unterscheidung des Detaillierungsgrads der Prädiktion möglich. Für den Verkehrsfluss wird diese großräumig über mehrere Straßensegmente oder Stadtgebiete benötigt. Die Prädiktion wird damit unabhängig von einer einzelnen fahrt- oder von fahrerindividuellen Gegebenheiten. Ziel ist es einen Durchschnittswert für die aktuelle oder zukünftige Geschwindigkeit auf einem Straßenabschnitt zu bestimmen. Alternativ kann ein detaillierteres Geschwindigkeitsprofil erstellt werden. Dies ist dabei abhängig von einer einzelnen Fahrt und berücksichtigt das Beschleunigungs- und Verzögerungsverhalten der Fahrer. Zudem kann dies mit unterschiedlichen Datenquellen ergänzt werden und Umgebungsbedingungen berücksichtigen.

Zudem ist eine Unterscheidung anhand der Länge des Prädiktionshorizont möglich. Dies kann kurzfristig im Rahmen der Sichtweite des Fahrers und der Sensorik erfolgen oder über die Einbindung von LSA und einer kooperativen Längsdynamikregelung mit vorausfahrenden Fahrzeugen. Alternativ kann die Betrachtung über einen längeren Horizont bis zum Fahrtende oder für eine beliebige Trip-Distanz geschehen. Dabei resultieren für längere Abschnitte vermehrt Ungenauigkeiten auf Grund der Unvorhersehbarkeit von Verkehrseignissen, besonders im Stadtverkehr. Eine Aktualisierbarkeit der Prädiktion im Laufe der Fahrt kann dies verbessern.

Im Folgenden werden in Unterkapiteln mehrere Möglichkeiten vorgestellt, ein Geschwindigkeitsprofil zu prädizieren. In Kapitel 5.1.1 werden mathematische Modelle vorgestellt, die unabhängig von realen Fahrten das Fahrverhalten beschreiben. Dabei werden vor allem Umgebungsinformationen wie Verkehrsfluss, Kurvenkrümmungen oder Fahrbahnbreite berücksichtigt. Anschließend werden in Kapitel 5.1.2 fahrdatenbasierte Ansätze vorgestellt, die aus mehreren Fahrten ein zusammenhängendes Fahrprofil generieren. Zuletzt wird die Bewertung von Fahrprofilen in Kapitel 5.1.3 betrachtet, weil bei der Analyse der Eigenschaften von Fahrprofilen oder weiteren Zeitreiheninformationen Parallelen zu dieser Arbeit gezogen werden können. Mit Hilfe der Bewertung ist eine Bestimmung des Fahrzustandes, des Verkehrsflusses oder des Fahrverhaltens möglich. Abschließend wird der Stand der Forschung zusammengefasst und diskutiert.

5.1.1 Mathematische Modelle

Bei den mathematischen Modellen werden Ansätze betrachtet, bei denen das Verhalten des Fahrers im Zentrum steht. Anschließend werden Methoden zur optimierte Längsführung von Fahrzeugen aufgeführt, bei denen der Fahrer der Vorgabe des Modells folgen muss.

Fahrerverhaltensmodelle

In der Modellierung der Verkehrsflussdynamik wird das Intelligent Driver Model (IDM) von

Treiber [194] als zeitkontinuierliches Fahrzeugfolgemodell eingesetzt. Das Modell beschreibt das Verhalten von einzelnen Fahrern, vor allem in Bezug auf die Dynamik des Vorderfahrzeugs. Basierend auf Plausibilitätsbedingungen [126] werden Funktionen für die Geschwindigkeit und Beschleunigungen in Abhängigkeit einer Wunschgeschwindigkeit hergeleitet. Die Fahrstrategien beschreiben u.a. den Sicherheitsabstand zum Vorderfahrzeug und intelligente Bremsstrategien mit angepasster Verzögerung in Abhängigkeit davon, ob eine Situation eine stärkere Bremsung erfordert. Das Modell wird auch abseits der Verkehrssimulation eingesetzt. Ondruska [195] berechnet mit dem Modell ein Fahrprofil, um den Energieverbrauch eines Elektrofahrzeugs zu bestimmen und setzt dies für eine Reichweitenprädiktion ein. Schakel [196] nutzt das Modell zur Verbesserung des Verkehrsflusses durch einen Geschwindigkeitshinweis an den Fahrer.

Ein funktional ähnliches Fahrermodell stellt Back [197] in seiner Dissertation vor. Er setzt die Prädiktion der Beschleunigung und der Geschwindigkeit für eine energieoptimale Antriebsregelung ein. Dabei wird eine Zielgeschwindigkeit als Minimum aus der gesetzlich erlaubten Höchstgeschwindigkeit, der Höchstgeschwindigkeit in Kurven, einer umgebungsbedingten Höchstgeschwindigkeit (berechnet aus der Sichtweite) oder einer verkehrsbedingten Maximalgeschwindigkeit (bspw. über Radarsignale und dem Abstand zum Vorderfahrzeug) ermittelt. Für den Wechsel zwischen den diskreten Werten auf einer Strecke werden rampenförmige Übergänge mit konstanten und fahrerabhängigen Beschleunigungen und Verzögerungen angenommen.

In [198] wählt Biermann einen ähnlichen Ansatz. Ebenso werden nicht-fahrzeugspezifische Parameter dabei mit einbezogen. Notwendig für das System ist vor allem das voraus liegende Streckenprofil, sowie die Relativgeschwindigkeit und der Abstand zum Vorderfahrzeug. Zudem wird der tatsächliche Verkehrszustand mitbetrachtet. Es werden zwei Ansätze verfolgt. Der Fahrer kann die geforderte Antriebsleistung bestimmen und diese wird auf Basis der Prädiktion effizient umgesetzt. Alternativ kann das System die vollständige Regelung der Längsdynamik durchführen. Dies hat einen Kraftstoffvorteil von bis zu 5 % zur Folge. Unter anderem wird damit auch der Einsatz der Reibbremse reduziert.

Eine weitere Form eines synthetischen Geschwindigkeitsprofils generiert Minett [199] und vergleicht dies mit realen Fahrten auf einer fixen Route. Die mittlere Verkehrsgeschwindigkeit auf den Straßenabschnitten dienen dabei als Zielgeschwindigkeit. Die Übergänge werden mit einem Beschleunigungsmodell, das von einer fixen Distanz von 20 m für die Beschleunigungs- oder Verzögerungsvorgänge ausgeht, beschrieben. Begrenzt wird der Vorgang um eine maximale Beschleunigung von $3 \frac{\text{m}}{\text{s}^2}$. Als Fazit wird angemerkt, dass dies nicht ausreichend ist, um das Verhalten realitätsnah wiederzugeben. Eine Berücksichtigung von Umwelt und Infrastrukturparameter wird als erforderliche Erweiterung genannt.

Kabrowski [200] verwendet ebenso den streckenspezifischen Verkehrsfluss für die Erstellung eines Fahrprofils. Das Ziel ist dabei, den Energieverbrauch für ein Fahrzeug auf einer Fahrt zu prädizieren. Das Fahrermodell wechselt zwischen den Zielgeschwindigkeiten mit einer konstanten Beschleunigung oder Verzögerung. Zusätzlich werden Stopps am Ende eines Segments und deren durchschnittliche Haltedauer berücksichtigt.

Durch zusätzliche Sensorinformationen kann die Position des Vorderfahrzeugs in eine Echtzeitprädiktion mit aufgenommen werden. Themann [201] bindet neben den Umweltinformationen und Kartendaten auch C2C- und C2I-Informationen mit ein. Unter anderem wird die Schaltstrategie der LSA auf der vorausliegenden Strecke mit einbezogen. Die Daten werden in eine Erweiterung des PELOPS [202] Fahrerfolgemodell gespeist und prädizieren für die nächsten 1,5 km das Geschwindigkeitsprofil. Aufbauend darauf wurde eine energieeffiziente Längsführung umgesetzt [203].

Engstle [204] wählt einen vergleichbaren Ansatz wie [200]. Ebenso wird die Strecke und der Verkehr mit in das Modell aufgenommen. Für die Beschleunigung und Verzögerung werden mehrere Fahrercharakteristika (u.a. sportlich, komfortabel) verwendet, die nicht genauer spezifiziert werden. Der prädizierte Verlauf der Geschwindigkeit und der reale Verlauf zeigen laut Engstle eine brauchbare Übereinstimmung. Allerdings muss der Fahrer die Geschwindigkeitsvorgaben einhalten.

Schüller [32] identifiziert die Einflussfaktoren auf die Wahl der Geschwindigkeit im Straßenverkehr. Er beschreibt das Geschwindigkeitsverhalten auf Stadtstraßen und die Auswirkungen auf die Verkehrssicherheit. Mit Hilfe von Messungen der Geschwindigkeiten einzelner Fahrzeuge auf Straßensegmenten werden die Einflussfaktoren aus der Gestaltung, dem Betrieb und dem Umfeld der Straße vorgenommen. Untersucht werden 44 Einflussvariablen auf Tempo-50-Straßen in Dresden. Unter anderem wird die nutzbare Fahrbahnbreite, das Vorkommen von Bäumen am Straßenrand oder im Straßenraum sowie die Nutzungsart des Gebietes betrachtet. Es stellt sich heraus, dass mehrere signifikante Einflussgrößen aus der Stärke der Befahrung der Straße, der Querschnittsgestaltung, der Umfeldnutzung und der Fahrbahnoberfläche hervorgehen. Schüller stellt zudem multiple Regressionsmodelle auf und untersucht die verschiedenen Modelle auf die Eignung zur Bestimmung der mittleren Geschwindigkeit auf einem Straßenabschnitt.

Modelle zur optimierten Längsführung

In [205] beschreibt van Keulen ein mathematisches Modell für einen energetisch optimalen Geschwindigkeitsverlauf für Lkw. Ähnlich wie in den bereits erwähnten Fahrerverhaltensmodellen wird die mittlere Geschwindigkeit aus Verkehrsinformationen als Zielgeschwindigkeit zusammengesetzt. Zwischen den unterschiedlichen Höchstgeschwindigkeiten bestimmt die Optimalität die Vollastbedingung des Fahrzeugs. Bei der Verzögerung wird speziell für HEV eine Rekuperation des Fahrzeugs berücksichtigt.

Abhängig von der Verkehrslage auf Autobahnen optimiert Kohut [206] den Kraftstoffverbrauch mit Hilfe modellprädiktiver Regelung. Es wird der Verkehrsfluss 2 km vor dem Fahrzeug als bekannt angenommen. Kohut definiert eine obere und untere Grenze für die erlaubte Geschwindigkeit. Diese ist durch die zulässige Höchstgeschwindigkeit und eine maximale Reisezeit bestimmt. Die Optimierung des Kraftstoffverbrauchs erfolgt unter Berücksichtigung von akzeptablen Reisezeiten. Bis zu 7 % Ersparnis bei nur knapp 3 % längerer Reisezeit sind die Folge.

Daeheung [207] setzt auf Dynamische Programmierung (DP), um das Fahrprofil in Abhängigkeit von Verkehrsfluss und Fahrzeugeigenschaften zu wählen. Dabei werden Gewichtungsfaktoren für Straßensegmente abhängig von der Segmentlänge und der Durchschnittsgeschwindigkeit erstellt. Des Weiteren wird ein Faktor aus dem Verhältnis der Durchschnittsgeschwindigkeit und der erlaubten Höchstgeschwindigkeit hinzugezogen.

Markschläger [208] grenzt zur Bestimmung einer global optimalen Fahrstrategie, die neben der maximalen Höchstgeschwindigkeit auch die Streckentopologie und Krümmung mitberücksichtigt, die mögliche Geschwindigkeit in einem Band ein. Innerhalb dieses Bandes werden die möglichen Geschwindigkeiten diskretisiert und bilden somit den Optimierungsraum. Mit DP wird die optimale Vorgabe gefunden. Es zeigen sich deutliche Kraftstoffersparnisse von durchschnittlich 10 %. Radke [209] erweitert den Ansatz um eine Fahrstrategie hinsichtlich Energieeffizienz, Fahrdynamik, Fahrkomfort und Fahrsicherheit. Zudem wird die subjektive Akzeptanz des Systems im Realversuch getestet. Für eine schnelleren Berechnung setzt Wahl [210] auf Iterative Dynamic Programming. Wollaeger [211] lagert als Alternative die Berechnung in das Back-End aus. Einen ähnlichen Ansatz zur energieeffizienten Fahrweise verfolgt von Saerens. In [212] wird das Problem mit Hilfe der Theorie der optimalen Steuerung gelöst.

Tabelle 5.1 : Literaturübersicht mathematische Modelle

Quelle	Jahr	Methode	Datenbasis	Horizont	Ziel
[194]	2000	Physikalisch	-	K+L	F
[195]	2014	IDM	-	L	P
[196]	2012	IDM	-	K	F
[197]	2006	Physikalisch	-	L	P
[198]	2009	Physikalisch + Verkehr	-	L	P
[199]	2011	Physikalisch + Verkehr	-	L	P
[200]	2012	Physikalisch + Verkehr + Stopps	-	L	P
[201]	2013	PELOPS + C2C + C2I	-	L	P
[204]	2014	Physikalisch + Verkehr + Fahrer	-	L	P
[213]	2014	Kognitive Mechanismen	-	L	F
[32]	2009	Regressionsmodelle	Einzelne Straßen	K	F
[205]	2009	Physikalisch	-	L	P
[206]	2009	Modellprädiktive Regelung	-	L	P
[207]	2012	DP	-	L	P
[208]	2012	DP	-	L	P
[210]	2013	Iterative Dynamic Programming	-	L	P
[211]	2012	DP + Back-End	-	L	P
[212]	2012	Optimal Control	-	L	P

K = Kurzfristig, L = Langfristig, F = Fahrverhalten, P = Energie

5.1.2 Datenbasierte Ansätze

Bei den datenbasierten Ansätzen wird eine Unterteilung auf Basis der Länge des Prädiktionshorizonts vorgenommen. Bei den vorgestellten Varianten ist das Ziel die Beschreibung des Fahrverhaltens.

Kurzfristige Prädiktion

Mit ca. 200 Befahrungen einer Autobahnauffahrt und einer S-Kurve mit anschließender 90°-Kurve von 15 Fahrern untersucht Schroedl [214] die Aufzeichnungen statistisch. Speziell werden die 25-ten und 75-ten Perzentile der Geschwindigkeit ermittelt. Anschließend erfolgt eine Prädiktion der nächsten 200 m mit einem neuronalen Netz. Als Eingang dient dabei die aktuelle Geschwindigkeit und die ermittelten statistischen Parameter. Das beste Ergebnis zeigt ein Netz mit 15 Neuronen mit einer Standardabweichung von $4,8 \frac{\text{km}}{\text{h}}$.

Bartholomaeus [215] wählt einen vergleichenden Ansatz zur Prädiktion der Geschwindigkeit einer Trambahn auf einer festen Route von 360 m Länge. Das Ziel ist eine echtzeitfähige Prädiktion auf einem Steuergerät. Über eine Singular Value Decomposition (SVD) werden die 131 aufgezeichneten Fahrten hierfür in einen Parameterraum überführt und dieser anschließend diskretisiert. Der Prädiktionsalgorithmus sucht auf Basis der zurückgelegten Strecke die aufgezeichnete Fahrt mit der höchsten Übereinstimmung. Das Ergebnis verbessert sich über die zurückgelegte Strecke.

Ciccarese und Donato [216, 217] stellen einen Ansatz zur Echtzeitprädiktion der Geschwindigkeit mit Daten von Fahrzeugen aus der Umgebung vor. Neben Position und Geschwindigkeit werden weitere Statusinformationen übertragen. In der Verkehrssimulation SUMO [218] wird der Ansatz für einen Prädiktionshorizont von bis zu 60 s untersucht. Der Fehler ist dabei abhängig von der Zahl der Fahrzeuge, die ihre Routeninformationen teilen und vor allem, ob auch empfangene Nachrichten von den Fahrzeugen weitergeleitet werden.

Sun [30, 219] untersucht für HEV neuronale Netze zur Prädiktion des Geschwindigkeitsverlaufs auf Basis realer Fahrdaten. Die vergleichenden Tests werden an Fahrzyklen mit einem Prädiktionshorizont von 10 s durchgeführt. Die neuronalen Netze verwenden die letzten 10 s als

Eingang. Am besten schneidet ein Radial-Basisfunktion neuronales Netz ab. Der Fehler hängt stark von der Vorausschaulänge ab [220].

Ein Autoregressive-Model ist ein statistisches Verfahren für die Analyse von Zeitreihen. Rezaei [221] verwendet es für die Prädiktion der Geschwindigkeit der nächsten zehn Zeitschritte auf Basis der letzten 21 Zeitschritte. Ergänzt wird das Modell um zusätzliche Informationen über bevorstehende Ereignisse (u. a. Stoppschild, Kurve, LSA). Bei dem Test auf einen Fahrzyklus für einen Horizont von 10 s liegt der Fehler bei unter $9 \frac{\text{km}}{\text{h}}$.

Geng [222] modelliert das menschliche Fahrverhalten für den Einsatz in autonomen Fahrzeugen. Auf Basis von realen Fahrdaten, den Streckeninformationen, dem aktuellen Wetter, der Zielgeschwindigkeit am Ende des Prädiktionshorizonts und weiteren Größen wird ein dreilagiges neuronales Netz mit 20 Neuronen trainiert, um das Verhalten für eine Situation zu bestimmen. Prädiktionen im Bereich von 5 s weisen größere Fehler auf als bei einer kurzfristigen Betrachtung.

Langfristige Prädiktion

Kerper [29, 223–226] beschreibt die wirkungsorientierte Analyse von Fahrprofilen mit dem Ziel der Reduzierung des Kraftstoffverbrauchs. Dabei wird ein kraftstoffreduzierendes Fahrprofil auf Basis eines zukünftigen Fahrprofils erstellt. Auf einer Route von 15 km Länge werden zehn Fahrten durchgeführt. Die Strecke wird in Segmente unterteilt, in denen mit Verkehrseignissen zu rechnen ist. Mit mehreren Randbedingungen führt der strategische Ansatz zu Segmenten von 200–300 m. Die darin enthaltenen Fahrprofile werden über das Distanzmaß Dynamic Time Warping (DTW) verglichen. Diese werden mit einem agglomerativen hierarchischem Clusterverfahren mit maximalem Abstand der Elementpaare kategorisiert. Die Zahl der Cluster ist variabel, allerdings abhängig von einem festgelegten Schwellenwert. Das repräsentative Profil wird berechnet aus dem Median der Geschwindigkeiten in diesem Cluster. Das wahrscheinliche Profil für einen längeren Horizont über mehrere Segmente hinweg, setzt sich aus den für die letzten bereits zurückgelegten Segmente häufigsten Profilen zusammen.

Mauk [227] stellt einen selbstlernenden Ansatz mit einem Datenspeicher im Fahrzeug vor. Das Streckennetz wird dabei aus den gefahrenen Fahrten erlernt. Wird ein Streckenabschnitt erneut befahren, können gespeicherte Attribute wieder verwendet werden. Speziell kontinuierliche Größen wie bspw. die Geschwindigkeit und deren Verteilung werden über eine Wahrscheinlichkeitsdichtefunktion beschrieben. Mauk untersucht auch den Umgang mit veraltetem Wissen, der Anzahl der notwendigen Lernfahrten und die davon abhängige Güte der Prädiktion. Ziel ist eine Anwendung in der Reichweitenprädiktion oder der Betriebsstrategie für Hybridfahrzeuge.

Kabrowski [228] erstellt für PHEV synthetische Fahrprofile. Die geplante Route wird in Segmente mit gleicher Durchschnittsgeschwindigkeit bzw. zulässiger Höchstgeschwindigkeit und potentiellen Stopps unterteilt. Aus aufgezeichneten Fahrten werden über Markov Ketten Geschwindigkeitsprofile unter den gleichen Randbedingungen erzeugt. Diese entsprechen nicht zwingend realen Fahrten auf derselben Strecke.

Die Prädiktion des Fahrprofils für eine vollständige Route im urbanen Umfeld untersucht Park [229]. In dem Verfahren werden Streckendaten, individuelle Fahrdaten, Geschwindigkeitsbegrenzungen und der Verkehrsfluss berücksichtigt. Die Datenbasis stellen 92 Fahrten von zwei verschiedenen Fahrzeugen auf einer Route. Vorab erfolgt eine Segmentierung der Fahrt nach Verkehrsfluss und Straßendaten in vier Kategorien (freifließend, Stoppschild, LSA, Abbiegevorgang). Für jede Kategorie wird ein mehrlagiges Perzeptron erstellt. Für jedes Netz werden unterschiedliche Eingangsparameter verwendet. Für 31 Testfahrten liegt der ermittelte Fehler bei unter $6 \frac{\text{km}}{\text{h}}$. Wang [230] erweitert das Modell für Autobahnauf- und -abfahrten.

Einen ähnlichen Ansatz verfolgt Lemieux [231]. Bei einer Route, die von einem Fahrer mehr als 700 mal befahren wurde, werden Verkehrsinformationen und die Straßengeometrie aufgezeichnet. Das Modell besteht aus einem gestapelten Autoencoder. Der mittlere quadratische Fehler liegt bei $14,7 \frac{\text{km}}{\text{h}}$.

Tabelle 5.2 : Literaturübersicht datenbasierte Ansätze

Quelle	Jahr	Methode	Datenbasis	Horizont
[214]	2003	Neuronales Netz	>200 Manöver	K
[215]	2008	SVD + Nachbar	131 Manöver	K
[216]	2012	SUMO	-	K
[219]	2014	RBF-Neuronales Netz	Fahrzyklen	K
[221]	2015	Autoregressive-Model	Fixe Route 40 Fahrten	K
[222]	2016	Neuronales Netz	>40 000 km	K
[29]	2016	DTW + Hierarchisches Clustern	Fixe Route 10 Fahrten	L
[227]	2011	Selbstlernender Ansatz	Fixe Route 85 Fahrten	L
[228]	2013	Markov Ketten	> 1 000 Fahrten	L
[229]	2013	Mehrlagiges Perceptron	Fixe Route 92 Fahrten	L
[231]	2015	Stacked Autoencoder	Fixe Route > 700 Fahrten	L

K = Kurzfristig, L = Langfristig

5.1.3 Bewertung von Fahrprofilen

Bei der Bewertung von Fahrprofilen wird eine Unterteilung in Verkehrsfluss, Energieverbrauch und Emissionen sowie Methoden zur Fahrerabsichtserkennung vorgenommen.

Verkehrsfluss

Um den Verkehrsfluss aus GPS-Daten von einzelnen Fahrzeugen zu kategorisieren, betrachtet Yoon [232] die mittlere räumliche und mittlere zeitliche Geschwindigkeit für Streckensegmente. Bereits mit wenigen Datensätzen ist eine Aussage über den Zustand möglich. Die Klassifizierung der Daten ist auch nachträglich möglich.

Montazeri [233] bestimmt zur Klassifizierung des Verkehrsflusses aus Fahrdaten mehrere Parametern nach [28]. Anschließend erfolgt ein Clustern über den k-means Algorithmus. Mit drei Kategorien für den Verkehrsfluss und den Attributen mittlerer Beschleunigung sowie dem Anteil der Stillstandszeit auf dem Segment zeigt sich eine Übereinstimmung der Daten mit dem Verkehrsfluss zu 87 %.

Einen Ansatz zur Fahrzustandsdetektion für eine Zykluserkennung stellt Xi [234] vor. Die Zuordnung zu den Kategorien Autobahnfahrt, Überlandfahrt und Fahrt im urbanen Bereich (freifließend oder verkehrsreich) wird mit einer Support Vector Machine (SVM) bestimmt. Die Betrachtung erfolgt in einem Zeitraum von 150 s. Als wichtigste Parameter erweisen sich die Durchschnittsgeschwindigkeit, die maximale Beschleunigung sowie die maximale Verzögerung und der zeitlich prozentuale Anteil an Stopps.

Energieverbrauch und Emissionen

Wie in Kapitel 1.1 bereits eingeführt, beschreibt Ericsson [28] die Haupteinflussgrößen über eine Principal Component Analysis (PCA) für Emissionen und Verbrauch. Darüber hinaus wurden in [235] auch Einflussgrößen von Straßeneigenschaften und dem Verkehr auf das Fahrverhalten und damit auf den Schadstoffausstoß und den Kraftstoffverbrauch betrachtet.

Kang [236] entwickelt straßensegmentbasierte Emissionsmodelle. Dafür werden sechs Fahrtmuster auf den Abschnitten über Geschwindigkeit, Beschleunigung, Motorlast und den Verkehrsfluss klassifiziert. Aus diesen Klassen werden Verbrauchscharakteristiken für den Abschnitt ermittelt. Die Methode wird für ein Eco-Routing-Verfahren verwendet. Es wird ein Vergleich von verschiedenen Routen mit den Optimierungsgrößen Zeit, Distanz und Emission erstellt.

Eine ähnliche Fahrtmusteridentifikation erfolgt in [237] durch Fahrtsegmentierung aus Verkehrs- und Steigungsdaten. Die benötigte Antriebsleistung für die jeweiligen Muster wird bestimmt und eine Prädiktion des Energieverbrauchs auf der Route mit denselben Größen für eine beliebige Fahrt durchgeführt.

Hemmerle [24] betrachtet den erhöhten Kraftstoffverbrauch auf verkehrsreichen Straßen. Dafür werden die Fahrten auf einem Abschnitt manuell klassifiziert. Die 15 Klassen sind abhängig von der Durchschnittsgeschwindigkeit und der Zahl der Anhaltevorgänge. Es folgt eine empirische Bestimmung des Kraftstoffverbrauchs. Es zeigt sich ein mehr als doppelt so hoher klassenspezifischer Verbrauch bei hohem Verkehrsaufkommen.

Rahimi-Eichi [238] setzt mit Hilfe von Markov Ketten basierend auf den Geschwindigkeits- und Beschleunigungsübergängen ein Fahrprofil für eine Route zusammen. Dies bietet die Möglichkeit, den Energieverbrauch von Elektrofahrzeugen zu bestimmen. Zusätzliche statistische Kriterien, u.a. die Durchschnittsgeschwindigkeit und Beschleunigung sowie deren Standardabweichungen werden mitberücksichtigt. Eine weitere Einschränkung der Daten für die Erzeugung der Markov Ketten auf Fahrten auf der gleichen Route verbessert das Ergebnis der Prädiktion.

Grubwinkler [95] stellt ein karten- und datenbasiertes Verfahren für die Prädiktion des Energieverbrauchs zentralisiert und über mehrere Fahrzeuge hinweg vor. Mit Hilfe von statistischen Merkmale aus den Fahrprofilen nach [28] wird der Durchschnittsenergieverbrauch bestimmt. Zusätzlich werden die Abweichungen von dem Durchschnittswert mit einer SVM, den vorausliegenden Verkehrsfluss und dem persönlichen Fahrverhalten prädiziert. Das Ergebnis zeigt eine Abweichung zum realen Verbrauch von unter 5 %.

Fahrerabsicht

Anhand der Position des Fahrzeugs, der aktuellen Fahrtrichtung und dem Lenkradwinkel ermitteln Agamennoni [239] und Worrall [240] das Fahrverhalten im Verkehr. Mit Hilfe der Modellierung mit einem probabilistischen graphischen Modell werden die Kategorien *Beschleunigen*, *Fahren mit konstanter Geschwindigkeit* und *Verzögern* bestimmt.

Johnson [241] stellt eine Fahrmanövererkennung vor. Mit den Sensorinformationen eines Smartphones werden Abbiegevorgänge oder Spurwechsel erkannt. Das Verfahren basiert auf dem Vergleich der Zeitreiheninformationen der Sensoren mit klassifizierten Daten auf Basis des DTW Algorithmus. Besonders aggressives Fahrverhalten kann zuverlässig detektiert werden.

Die Detektion von Abbiegevorgängen an einer Straßenkreuzung betrachtet Liebner [242] mit Hilfe des IDM. Es werden während der Fahrt Simulationen durchgeführt zur Erkennung von Rechtsabbiegevorgängen oder Geradeausfahrten jeweils mit oder ohne Stopp an einer Haltelinie. Für die Parametrierung des IDM werden 50 Realfahrten mit dem k-means Algorithmus in drei Klassen aufgeteilt und die jeweils charakteristische Beschleunigung bestimmt.

Eine ähnliche Problemstellung untersucht Gross [243]. Der Ansatz ist allerdings datengetrieben ohne die Verwendung eines Fahrermodells. Für neun Kreuzungen werden die Fahrdaten mit dem Klassifikationsverfahren Random Forest untersucht.

Tabelle 5.3 : Literaturübersicht Bewertung von Fahrprofilen

Quelle	Jahr	Methode	Datenbasis	Horizont	Ziel
[232]	2007	Klassifizierung	40 Manöver	-	V
[233]	2011	k-means	Fixe Route > 6 Monate	-	V
[234]	2011	SVM	unbekannt	K	V
[28]	2001	PCA	5 Fzg, 18 945 km	K	P
[235]	2005	PCA	5 Fzg, 18 945 km	K	P
[236]	2011	Entscheidungsbaum	unbekannt	K	P
[237]	2012	Fuzzy c-means	unbekannt	L	P
[24]	2014	Manuelle Klassifikation	>1 000 Manöver	K	P
[238]	2015	Markov Ketten	>23 000 km	L	P
[95]	2017	SVM	>450 000 km	L	P
[239]	2011	Probabilistisches Modell	unbekannt	L	F
[241]	2011	DTW	>200 Manöver	K	F
[242]	2012	IDM	>200 Manöver	K	F
[243]	2016	Random Forest	>50 000 Manöver	K	F

K = Kurzfristig, L = Langfristig, F = Fahrverhalten, P = Energie, V = Verkehrsfluss

5.1.4 Zusammenfassung und Abgrenzung

Eine zusammenfassende Übersicht zu den einzelnen Quellen und eine Einordnung in die relevanten Bereiche ist in Tab. B.1 dargestellt. Ebenso sind in diesem Fall, ähnlich wie bei der Trip-Prädiktion, die vorgestellten Methoden meist nicht direkt miteinander vergleichbar.

Die mathematischen Modelle haben ihren Ursprung in der Modellierung für Verkehrssimulationen. Die Fokus liegt bei der Untersuchung des Fahrzeugfolgeverhaltens [194]. Es wird versucht, mit möglichst wenig Parametern das Verhalten allgemeingültig abzubilden. Das reale Verhalten ist jedoch von unzähligen Einflussfaktoren abhängig. Nach Schüller [32] spielen die Straßenbeschaffenheit und auch die örtlichen Gegebenheiten eine entscheidende Rolle. Viele dieser Parameter sind schwer, großräumig zu erheben. Eine ortsspezifische Modellierung des Verhaltens ist daher beschränkt möglich. Viele Modelle der vorgestellten datenbasierten Ansätze nutzen den vollständigen Datenbestand ortsunabhängig und sind daher ebenso auf ein verallgemeinertes Fahrverhalten gerichtet, was mit einer begrenzten Datenbasis zu begründen ist. Nur wenige Ansätze berücksichtigen eine ortsspezifische Modellierung und demonstrieren dies an einzelnen Routen oder für einen eingeschränkten zeitlichen oder räumlichen Bereich.

Ziel dieser Arbeit ist, wie in Kapitel 2.3.1 beschrieben, das Geschwindigkeitsprofil zu präzisieren. Die vorgestellten Methoden zur energieoptimierten Längsführung zielen auf eine aktive Beeinflussung des Fahrverhaltens. Allerdings zeigt die Herangehensweise zur Bestimmung des Optimierungsraums Möglichkeiten, das vorausliegende Geschwindigkeitsprofil regelbasiert einzugrenzen. Ebenso ist die Eigenschaftsidentifikation durch die Fahrprofile zwar nicht direkt mit dem Ziel dieser Arbeit vergleichbar, aber es werden vielversprechende Ansätze vorgestellt, den Fahrzustand zu identifizieren.

Im Rahmen dieser Arbeit ist eine Prädiktion für die nächsten Straßensegmente für einen mittel- bis langfristigen Horizont von Bedeutung. Speziell die vorgestellten Methoden für eine kurzfristige Prädiktion weisen hohe Abweichungen auf, wenn diese für eine langfristige Prädiktion angewendet werden. Die datenbasierten Methoden sind hauptsächlich Varianten des maschinellen Lernens. Es werden mehrere Möglichkeiten untersucht, mit denen diese Fragestellung angegangen werden kann.

Wie bereits bei der Trip-Prädiktion werden die in Kapitel 3.1.1 beschriebenen Flottenversuche verwendet. Im Vergleich zu den meisten vorgestellten Ansätzen wird eine breitere Datenbasis verwendet, die sowohl Stadtstraßen, ländliche Gebiete und Autobahnen auffasst. Für diese

sollen anhand realer Fahrprofile langfristige Prädiktionen ermöglicht werden. Aufbauend auf dem Geschwindigkeitsverlauf werden weitere Fahrzustandsgrößen abgeschätzt. Darüber hinaus sollen weitere Eigenschaften, wie der orts- und situationsspezifische Energieverbrauch ermittelt werden können.

5.2 Methodik

Im folgenden Abschnitt wird auf die gewählte Methodik eingegangen. Zuerst wird der Systemansatz vorgestellt. Anschließend werden die einzelnen Schritte daraus hervorgehoben.

5.2.1 Systemansatz

Die Forschungsfrage sowie der Lösungsansatz der Fahrprofilprädiktion werden in Kapitel 2.3.2 bereits eingeführt. Die grundsätzliche Zielsetzung ist es, ortsspezifische Muster in dem Bewegungsverhalten aller Fahrzeuge zu erkennen und diese für eine vorausliegende Fahrt vorherzusagen. Dabei ist neben dem Geschwindigkeitsprofil auch die Gierrate, bzw. die Beschleunigungen des Fahrzeugs von Interesse. Zusätzlich kann die Aussagekraft über weitere Metadaten, wie bspw. den Energieverbrauch erweitert werden.

In Abb. 5.1 wird das Verfahren der Fahrprofilprädiktion vorgestellt. Der Ansatz kann in vier wesentliche Schritte unterteilt werden, die in den darauf folgenden Kapiteln detailliert betrachtet werden. Die Resultate der einzelnen Schritte sind jeweils auf der rechten Seite dargestellt.

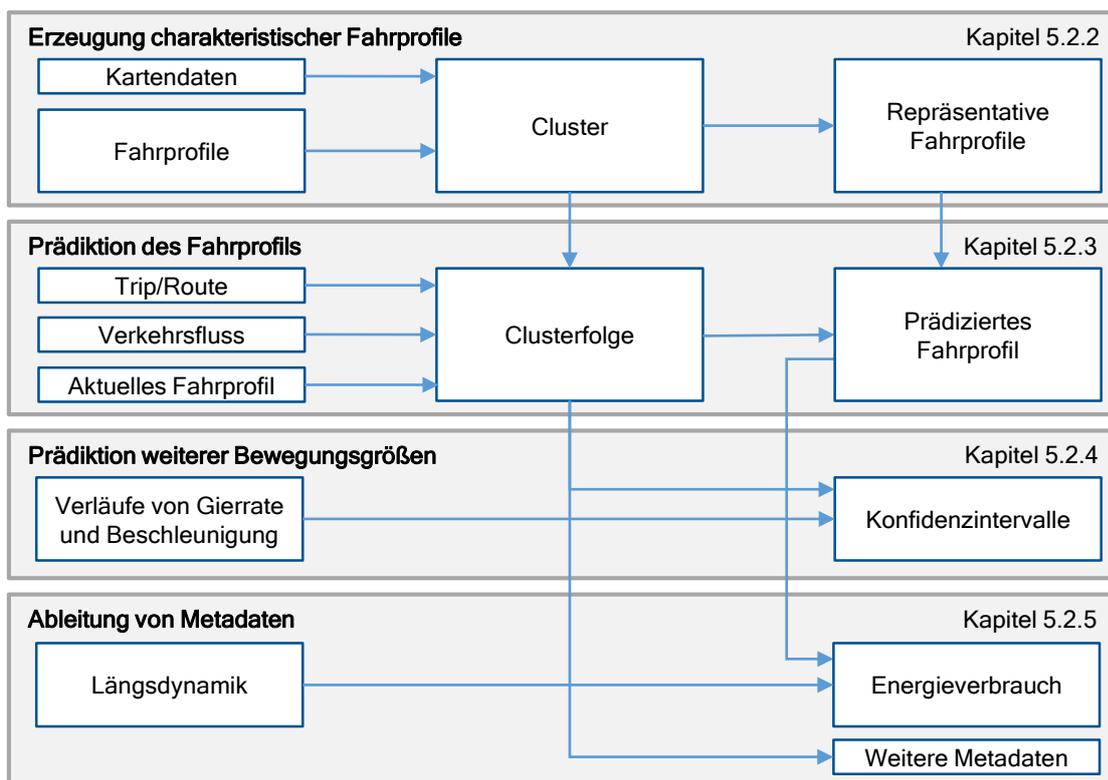


Abbildung 5.1 : Ansatz der Fahrprofilprädiktion

Zu Beginn werden die in Kapitel 3.1.1 vorgestellten Fahrprofile näher untersucht. Diese werden nach den einzelnen Straßensegmenten mit den Informationen aus den Kartendaten (Kapitel 3.1.2) verknüpft. Um die Ähnlichkeit zwischen den einzelnen Fahrprofilen zu ermitteln, wird das Distanzmaß Fast Dynamic Time Warping (FastDTW) verwendet. Darauf aufbauend dient ein spektrales Clusterverfahren zur Identifikation und Zuordnung der Fahrten zu bestimmten Kategorien. Diese Cluster sind ein wesentlicher Bestandteil des darauf aufbauenden Prädikti-

onsmodells. Für die einzelnen Straßenabschnitte und die jeweiligen, dazugehörigen Kategorien werden dafür charakteristische Fahrprofile ermittelt.

Als nächster Schritt erfolgt die Prädiktion des Fahrprofils. Als Basis dient die Methode der multinomialen logistischen Regression. Für jedes Segment wird ein Prädiktionsmodell erstellt, das neben den zuvor befahrenen Straßensegmenten und den jeweilig zugeordneten Cluster der Fahrprofile den Verkehrsfluss mitberücksichtigt. Mit den Daten der aktuellen Fahrt auf den vorangegangenen Segmenten und dem momentanen Verkehrsfluss wird das zu erwartende Cluster auf dem nächsten Segment prädiziert. Dieses Verfahren wird für die Zahl der Segmente auf der vollständigen Route bzw. im Rahmen des zuvor ermittelten Trips wiederholt. Damit wird für den Horizont eine Folge von Clusterzugehörigkeiten bestimmt, welche die zu erwartende Fahrt am besten beschreibt. Mit den zuvor definierten repräsentativen Profilen kann ein vollständiges Fahrprofil über einen längeren Abschnitt berechnet werden.

Neben der Geschwindigkeit wurden für die aufgezeichneten Fahrten in der Datenbasis auch Beschleunigungen und Drehraten hinterlegt. Mit der im vorherigen Abschnitt bestimmten Clusterfolge können die Zeitreihen für die weiteren Bewegungsgrößen selektiert werden, die einen ähnlichen Geschwindigkeitsverlauf wie die zu erwartende Fahrt aufweisen. Für das jeweilige Straßensegment werden die entsprechenden Zeitreihen genauer betrachtet und Konfidenzintervalle für die vorausliegende Fahrtstrecke bestimmt. Besonders in Kurvenfahrten können diese Informationen zur Bewertung der Fahrstabilität von Nutzen sein.

Zuletzt können weitere Metadaten oder Zeitreiheninformationen abgeleitet werden. So kann unter anderem der für den Horizont notwendige Energieverbrauch des Fahrzeugs aus dem prädizierten Fahrprofil mit Hilfe einer Längsdynamiksimulation abgeleitet werden. Alternativ können analog zu der Gierrate anderweitige Zeitreiheninformationen prädiziert werden. Generell können aus den ermittelten Clusterfolgen für den Abschnitt weitere Metadaten, wie bspw. die mittlere Standzeit oder die Parameter in Kapitel 1.1 ausgewertet werden.

5.2.2 Erzeugung charakteristischer Fahrprofile

Für die Prädiktion des Fahrprofils wird ein datenbasierter Ansatz verfolgt. Dabei wird von einer straßensegmentspezifischen Betrachtung der Fahrten ausgegangen. Mit Hilfe von FastDTW werden die Profile miteinander verglichen. Die Fahrprofile werden über die resultierende Ähnlichkeitsmatrix mit einem spektralen Clusterverfahren kategorisiert. Aus diesen Kategorien dient jeweils ein Profil als charakteristisches Fahrprofil. Das Vorgehen ist angelehnt an der Arbeit von Kerper [226]. Allerdings erfolgte dabei keine kartenbasierte Betrachtung. In diesem Fall wird die Länge eines Segments nicht durch die Methode vorgegeben. Zudem wurden in [29] nur wenig Fahrten betrachtet und keine großflächige Untersuchung vorgenommen. Hierfür wird das Verfahren zur Optimierung der Rechenzeit für die großflächige Betrachtung angepasst. Die Ergebnisse diesbezüglich werden in [244] näher erläutert. Im Folgenden werden die einzelnen Schritte kurz vorgestellt.

Kartenbasierte Betrachtung

Die in Kapitel 3.1.1 beschriebenen Daten werden nachfolgend ausschließlich für den Raum Oberbayern ausgewertet. Wie bereits vorgestellt, zeichnen sich die Daten durch die GPS-Koordinaten, die Fahrgeschwindigkeit v_t und den Zeitstempel t aus. Gemäß der Vorverarbeitung durch das Mapmatching aus Kapitel 3.2.1 werden die Karteninformationen aus Kapitel 3.1.2 ergänzt. Aus der zweidimensionalen räumlichen Betrachtung wird damit eine eindimensionale, straßensegmentspezifische Betrachtung für jeweils einen Straßenabschnitt.

Damit haben alle Fahrten auf einem Segment die gleiche Distanz zurückgelegt. Der zurückgelegte Weg s wird über eine numerische Integration der Geschwindigkeit v_t über die Trapezregel [245, S. 828–831] berechnet. Aus der zeitlich aufgelösten Geschwindigkeit v_t wird die räumliche Geschwindigkeit v_s ermittelt. Der Unterschied zwischen v_t und v_s ist in Abb. 5.2 dargestellt. Der zeitliche Bezug geht dabei zum Teil verloren, weil die Dauer einer Stillstandphase aus v_s nicht bestimmt werden kann. Für eine gleiche Auflösung für alle Fahrten erfolgt eine lineare

Interpolation für jeden Meter. Damit sind weiterhin variable Haltepunkte bspw. vor Kreuzungen oder an LSA möglich.

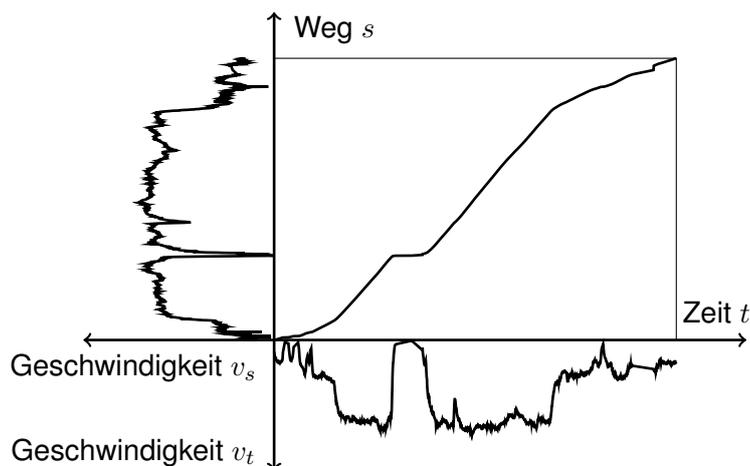


Abbildung 5.2 : Geschwindigkeitsprofil und Geschwindigkeitsganglinie nach Leutzbach [246, S. 32]

Die gefahrene Route hat einen entscheidenden Einfluss auf das Geschwindigkeitsprofil. Besonders Abbiegevorgänge vor oder nach dem betrachteten Straßensegment weisen deutliche Unterschiede im Verlauf des Fahrprofils auf. Daher wird die kartenbasierte durch eine routenabhängige Betrachtung erweitert. Es wird bei der Kategorisierung der Fahrprofile eine Unterscheidung in Ursprung und Ziel des Straßensegments vorgenommen. Die Profile auf unterschiedlichen Abschnitten oder auf identischen Abschnitten und unterschiedlichen Ursprungs-Ziel-Kombinationen können zwar Ähnlichkeiten aufweisen, um die straßentypische Betrachtung beizubehalten wird hierbei keine Verallgemeinerung vorgenommen.

Um den Prozess der Klassifizierung der Daten zu beschleunigen wird eine weitere Vorverarbeitung durchgeführt. Dabei wird auf ein offensichtliches Merkmal hin untersucht. Es erfolgt eine Unterscheidung der Fahrprofile anhand der Existenz von Haltevorgängen. Dabei wird untersucht, ob auf der zweiten Hälfte des vorherigen Segments und/oder im aktuellen Segment ein Stopp eingetreten ist. Beide Kriterien führen zu einer deutlichen Unterscheidung der Fahrprofile und können somit vorab in vier Klassen sortiert werden. Als Ergebnis aus der kartenbasierten Vorverarbeitung liegt eine Vielzahl an Fahrprofilen mit gleicher Länge für jedes Streckensegment in den vier Klassen der Vorsortierung und einer routenspezifischen Zuweisung vor.

Vergleich von Zeitreihen: Dynamic Time Warping

Zur Kategorisierung der Daten ist ein Ähnlichkeits- oder Distanzmaß erforderlich. Dieses muss gemäß der Daten geeignet gewählt werden. Die Geschwindigkeitsprofile können als Zeitreihen – in diesem Falle eine Reihe mit der Basis des zurückgelegten Wegs als stetig aufsteigende Größe – betrachtet werden. Generell können mehrere Distanzmaße verwendet werden, bspw. die euklidische Distanz oder die Hamming Distanz. Allerdings können diese bei einer Verschiebung der Signale in der Zeitdomäne zu hohen Fehlern führen [247]. Dies ist bei Fahrprofilen der Fall, weil bspw. für Haltevorgänge an Kreuzungen oder Haltelinien unterschiedliche Positionen eintreten können. Dies führt in diesem Bereich zu hohen euklidischen Distanzen beim Vergleich der Signale und zu einer mangelnden Vergleichbarkeit der Zeitreihen [226]. DTW kann mit derartigen Verschiebung im Zeitbereich umgehen und dies erfolgreich lösen [248, 249]. Die Methode ist nach Berndt [247] folgendermaßen definiert:

Es werden zwei Vektoren q und c miteinander verglichen. Wie zuvor beschrieben handelt es sich um zwei Fahrprofile auf einem Segment mit der gleichen Länge S .

$$q = q_1, q_2, \dots, q_i, \dots, q_S \quad (5.1)$$

$$c = c_1, c_2, \dots, c_j, \dots, c_S \quad (5.2)$$

Es wird eine Matrix aus den Elementen der Vektoren und einem generellen Distanzmaß aufgebaut. Dabei sind nach [247] verschiedene Funktionen möglich. In diesem Falle wird der Betrag der Differenz $d(q_i, c_j)$ der Elemente nach

$$d(q_i, c_j) = |q_i - c_j| \quad (5.3)$$

gewählt. Ziel ist es, einen Warping Pfad w zu ermitteln, der die Elemente von q und c anhand der minimalen Distanz zwischen den Elementen ausrichtet. Der Pfad

$$w = w_1, w_2, \dots, w_n, \dots, w_N \text{ mit } w_n = (i, j)_n \quad (5.4)$$

unterliegt in diesem Falle den Randbedingungen für $w_1 = (1, 1)$ und $w_N = (S, S)$. Des Weiteren werden nach [247] die Bedingungen der Stetigkeit, welche die Schritte auf benachbarte und diagonale Zellen begrenzt, und der Monotonie, das ein monotonen Wachsen von w nach Gl. 5.4 in i oder j verlangt, gefordert. Daraus folgt der optimale Pfad

$$DTW(q_i, c_j) = \min_w \left[\sum_{n=1}^N d(w_n) \right], \quad (5.5)$$

Der Pfad DTW ist derjenige Weg durch die Matrix, der die Kosten dabei minimiert. Um diesen Pfad zu finden, wird die kumulative Distanz γ über dynamische Programmierung rekursiv ermittelt.

$$\gamma(i, j) = d(q_i, c_j) + \min \begin{cases} \gamma(i-1, j) \\ \gamma(i, j-1) \\ \gamma(i-1, j-1) \end{cases} \quad (5.6)$$

Die Distanz in der letzten Zelle entspricht dem Distanzmaß für die beiden Zeitreihen. Mit diesem Verfahren können diese folglich miteinander verglichen werden.

Fast Dynamic Time Warping

Aufgrund der aufwendigen Berechnung des DTW-Pfades zeigt die Methode einen hohen Rechenaufwand. Bei einer Länge des Segments von S resultiert dies in einem Komplexitätsgrad von $O(S^2)$, bei der Betrachtung von zwei Zeitreihen. Zusätzlich ist es für die Erstellung der Cluster erforderlich, dass alle, bzw. möglichst viele Zeitreihen mit diesem Verfahren miteinander verglichen werden. Bei N Zeitreihen für ein Segment liegt dies bei $\frac{N(N-1)}{2}$ Operationen. Für die Betrachtung über den gesamten Beobachtungsraum führt das zu einem erheblichen Rechenaufwand.

Es gibt verschiedene Ansätze, die versuchen, die Rechenzeit des Verfahrens zu reduzieren oder das Ergebnis zu approximieren. Keogh [250] nennt unter anderem Methoden, welche die Auswertung der Distanzmatrix einschränken. Bspw. kann zur Berechnung der DTW-Matrix ein Band eingesetzt werden, das den Raum der Auswertung um die Diagonale und einem zusätzlichen Radius-Parameter begrenzt. Weitere Möglichkeiten sind die Abstraktion der Daten oder deren Indizierung.

Salvador [251] führt die Erweiterung FastDTW ein. Drei Schritte müssen dabei erfolgen:

- Reduktion der Auflösung:
Die Auflösung der Daten wird um ein Vielfaches reduziert. Dies kann bspw. über das *Piecewise-Aggregate-Approximation*-Verfahren [252] erfolgen. Anschließend wird der DTW-Pfad auf der niedrigeren Auflösung ermittelt.
- Projektion:
Der zuvor ermittelte Pfad wird auf die nächst höhere Auflösung projiziert.
- Verfeinerung:
Der Pfad der niedrigeren Auflösung erstreckt sich nun über mehrere Zellen in der Matrix. Dies wird als Band für die Begrenzung der Auswertung eingesetzt. Zusätzlich kann ein Radius-Parameter r verwendet werden, der das Band um zusätzliche Zellen weitet. Anschließend

wird erneut DTW für die höhere Auflösung unter der Begrenzung durchgeführt und der DTW-Pfad somit verfeinert.

Es handelt sich dabei um ein iteratives Verfahren. Für eine Aussage äquivalent zu DTW muss das Verfahren bis auf die volle Auflösung erhoben werden. Das Vorgehen wird in Abb. 5.3 dargestellt. Es wird deutlich, dass nur ein begrenzter Ausschnitt in der vollen Auflösung durchsucht werden muss. Nach [251] reduziert sich die Komplexität auf $O(S)$ falls $r < S$.

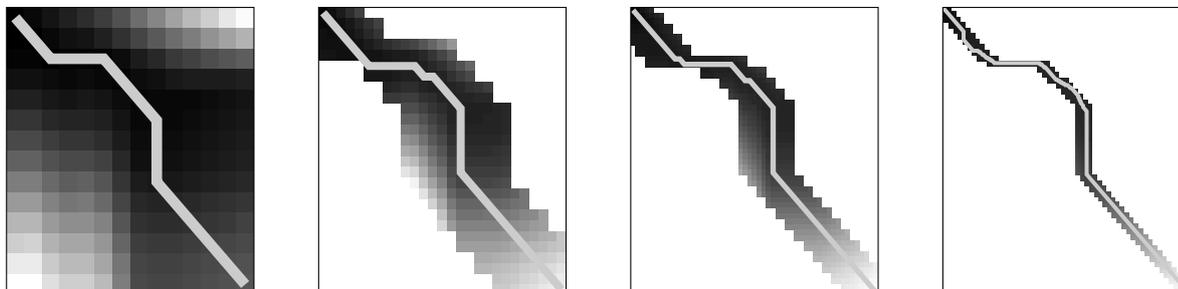


Abbildung 5.3 : Projektion des optimalen Pfades aus niedriger Auflösung (links) in höhere Auflösungen (rechts) nach Salvador [251]

In [244] wurde der Einfluss des Radius-Parameters und der Grad der Approximation, bis zu der das Verfahren durchgeführt wird, auf den Fehler bei der Clusterzuordnung mit einem spektralen Clusterverfahren untersucht. Da nur die Zuordnung der unterschiedlichen Fahrprofile zu den Clustern und nicht die absolute DTW-Distanz entscheidend ist, kann die Rechenzeit weiter reduziert werden. Das Ergebnis dieses Schritts ist eine symmetrische Distanzmatrix Y mit einem paarweisen Vergleich der Zeitreihen pro Straßensegment.

Ausreißer

Fehler in der Datenaufzeichnung, oder ein Fahrverhalten, das nur in Ausnahmefällen erfolgt (Notbremsung, Halten am Straßenrand bspw. bei Taxis) kann zu Ausreißern in der Datengrundlage führen. Diese müssen gemäß dem Vorgehensmodell in Kapitel 2.3.3 bereinigt werden. Ansonsten kann es dazu führen, dass das nachgelagerte Clusterverfahren diese als einzelne Gruppen identifiziert. Zur Bereinigung wird die erzeugte Distanzmatrix Y analysiert. Zuerst wird der durchschnittliche Abstand zu den Fahrten in der näheren Umgebung bestimmt. Dafür wird das 1-Perzentil der Distanz der Fahrten zu den übrigen Fahrten ermittelt. Weiter wird das 95-Perzentil davon als Grenze zur Definition eines Ausreißers bestimmt. Folglich sind Profile mit wenigen Nachbarn und einem großen Abstand zu anderen Fahrten exponiert. Die Ausreißer werden in den nachfolgenden Schritten nicht näher berücksichtigt.

Spektrales Clusterverfahren

Eine Übersicht über verschiedene Clusterverfahren, die zur Klassifizierung der Daten eingesetzt werden können, gibt Xu [253]. In [244] zeigt das spektrale Verfahren reproduzierbare Ergebnisse, auch wenn das FastDTW Verfahren nicht bis zum vollständigen Abstraktionsgrad angewendet wird. Dadurch lässt sich die Rechenzeit optimieren. Bei dem Verfahren werden die Eigenwerte und Eigenvektoren einer Ähnlichkeitsmatrix \hat{A} [254] verwendet. Dieses wird über die Elemente y_{qc} der Distanzmatrix Y , welche die Distanz zwischen zwei Profilen ausdrücken und Gl. 5.7 bestimmt. Das generelle Vorgehen zum spektralen Clustern erläutert Ng [255]. Die Methode wird um einen selbst-parametrierenden Skalierungsfaktor mit lokalen Gewichten nach Zelnik-Manor [256] ergänzt.

$$\hat{A}_{qc} = \exp\left(-\frac{y_{qc}^2}{2\sigma_q\sigma_c}\right) \quad (5.7)$$

Die lokalen Gewichte σ_q und σ_c werden als mittlere Distanz zu den anderen Profilen bestimmt [257]. Anschließend wird die Diagonalmatrix D definiert zu

$$D_{qq} = \sum_{q=1}^z \hat{A}_{qc} \quad (5.8)$$

und die normierte Ähnlichkeitsmatrix L mit

$$L = D^{-1/2} \hat{A} D^{-1/2} \quad (5.9)$$

ermittelt. Die ersten l Eigenvektoren von L bilden den Merkmalsraum T . Über l wird dabei auch die Anzahl der Cluster bestimmt. Diese wird nicht fest vorgegeben, sondern über den Abstand der Eigenwerte nach [254] ermittelt. Die größte Differenz der Eigenwerte setzt die Zahl der Cluster l_b fest. Zudem wird l pro Ursprung-Ziel-Kombination und der beschriebenen Vorsortierung auf maximal fünf Cluster begrenzt. Es folgt eine Normierung der Zeilenvektoren von T . Anschließend wird der k-means Algorithmus mit der Vorgabe von l_b Clustern auf T angewendet. Zur Veranschaulichung des Ergebnisses wird in Abb. 5.4 eine zweidimensionale Darstellung der Distanzmatrix und der resultierenden Cluster mit Multidimensionaler Skalierung (MDS) [258, S. 163–175] für einen exemplarischen Straßenabschnitt dargestellt. Zudem wird die exponierte Lage der Ausreißer deutlich gemacht.

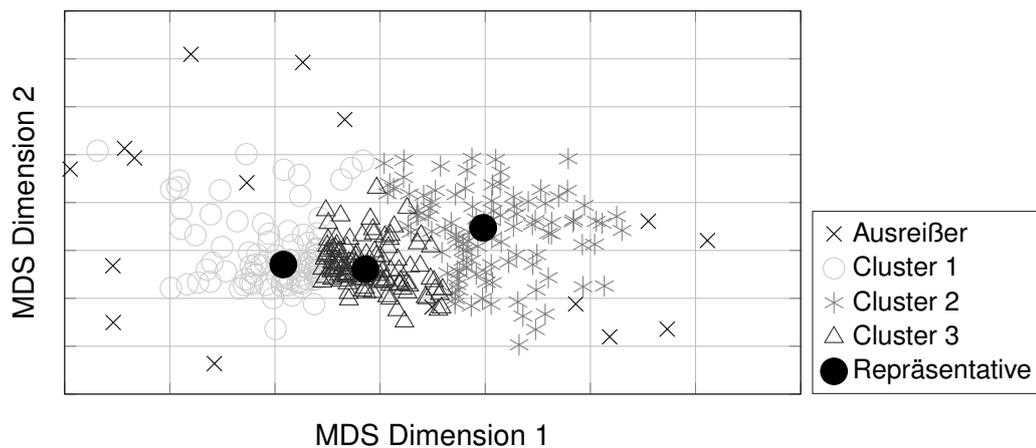


Abbildung 5.4 : Skalierung der Distanzmatrix in den zweidimensionalen Raum

Repräsentative Profile

Abschließend werden die repräsentativen Profile aus der Distanzmatrix Y bestimmt. Das Profil mit der kleinsten mittleren Distanz zu den anderen Profilen aus dem selben Cluster bestimmt das repräsentative Profil. Wie in Abb. 5.4 ersichtlich befindet es sich dadurch in einem zentralen Bereich des Clusters und ist daher charakteristisch dafür. In Abb. 5.5 wird das Ergebnis der Klassifikation, sowie die resultierenden Fahrprofile dargestellt.

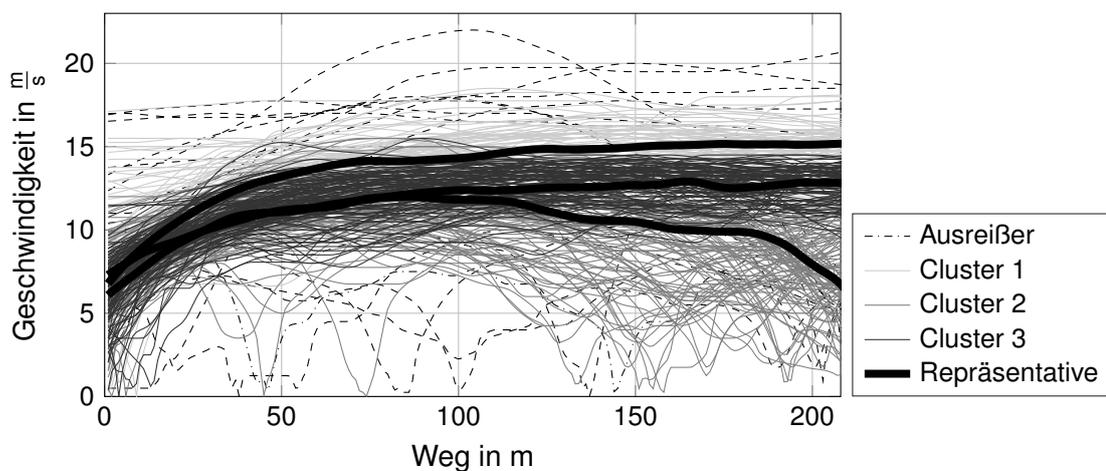


Abbildung 5.5 : Geschwindigkeitsprofile und Clusterzugehörigkeit nach [244]

Einfluss der Anzahl an Fahrten

Das vorgestellte Verfahren clustert die zugrunde liegenden Fahrten der Datenbasis. Wie in Kapitel 3.2.1 beschrieben, existieren dabei für die Straßensegmente jeweils unterschiedlich

viele Aufzeichnungen. Um zuverlässig Muster in einem Datensatz identifizieren zu können, ist es erforderlich, eine ausreichend große Auswahl an verschiedenen Fahrten zu evaluieren. Um den Einfluss der Anzahl an Fahrten für die Methode festzustellen, wird in Abb. 5.6 ein Straßenabschnitt näher betrachtet. Als Referenz dienen 1 000 von insgesamt über 2 000 Fahrten des Abschnitts zur Erstellung der Cluster und der charakteristischen Fahrprofile. Abgebildet sind die repräsentativen Profile in einer zweidimensionalen MDS-Darstellung. In je fünf zufälligen Stichproben werden unterschiedlich viele Fahrten als Eingang für das Verfahren ausgewählt. Die dargestellten Proben entsprechen den repräsentativen Profilen des erzeugten Clusters.

Cluster, die sich deutlich voneinander abgrenzen, können auch mit bereits 50 Fahrten in dem Clusterverfahren identifiziert werden. Mit steigender Zahl an Fahrten lässt sich eine Spreizung und damit Unterschiede in den Fahrprofilen erkennen. Ein möglichst breites Spektrum an Profilen ist zur Erstellung der Cluster daher sinnvoll. Wenn ähnlich viele Fahrten wie in der Referenz verwendet werden, stellt sich eine starke Tendenz für die zufällig gewählten Proben heraus, vergleichbare Resultate zu erzielen. Es kann daher auch nur eine zufällig gewählte Teilmenge der Daten für die Clusterbildung eingesetzt werden, um diese geeignet zu kategorisieren.

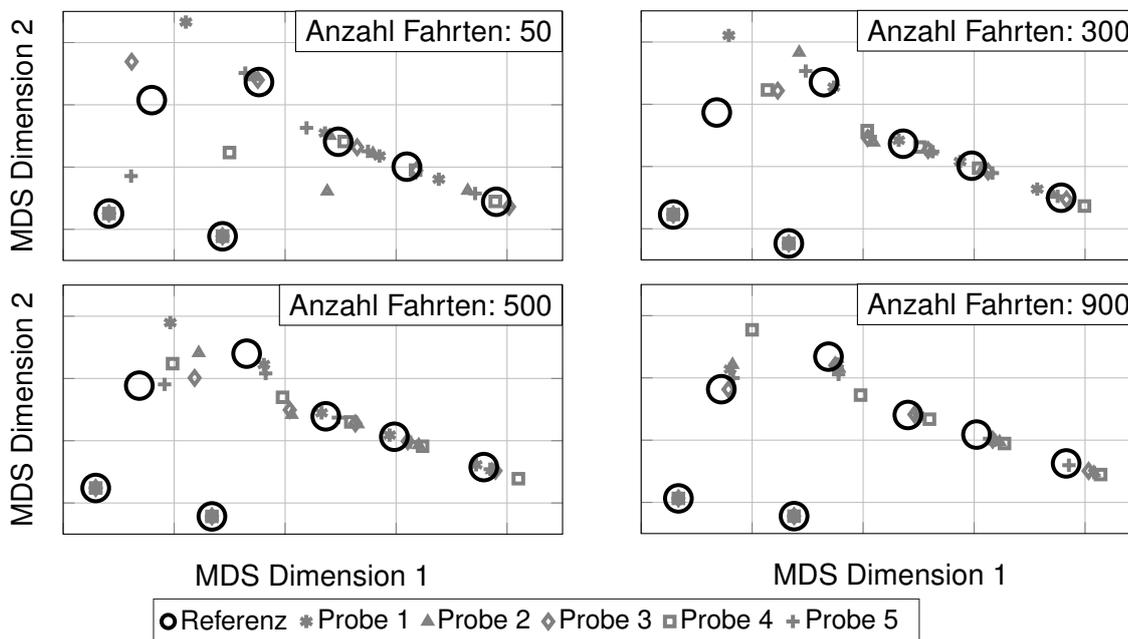


Abbildung 5.6 : Einfluss der Anzahl der Fahrten auf die Clustergenerierung

Klassifizierung weiterer Fahrten

Für die Verwendung in einer Online-Applikation, bei der kontinuierlich neue Fahrten zur Datenbasis hinzustoßen, ist eine sofortige Neuberechnung der Clusterzuordnung nicht zielführend, weil die Distanzmatrix für jede neue Fahrt ergänzt werden müsste. Daher wird ein Klassifizierungsverfahren angewandt, das neue Fahrten den bisherigen Clustern zuordnet. Hierfür wird *k-Nearest-Neighbor* auf Basis der *Piecewise-Aggregate-Approximation*-Informationen verwendet. Das am häufigsten zugeordnete Cluster der fünf nächsten Nachbarn wird für das neue Profil übernommen. Bei diesem Verfahren werden die repräsentativen Profile nicht aktualisiert. Die Fahrten können auch als Ausreißer kategorisiert werden, falls hauptsächlich exponierte Fahrten unter den nächsten Nachbarn sind.

Als Beispiel wird eine Autobahnausfahrt untersucht, bei der die Clustergrenzen nahe beieinander liegen und keine offensichtliche Trennung der Daten vorliegt. In Abb. 5.7 werden die repräsentativen Profile, sowie alle Fahrten im zweidimensionalen Raum dargestellt. Es zeigt sich, dass nachträglich klassifizierte Daten innerhalb der häufigsten Fahrten korrekt zugeordnet werden können. Fahrten an den Clustergrenzen werden vereinzelt falsch kategorisiert.

Eine vollständige Neuberechnung der Clusterzuordnung ist über größere zeitliche Intervalle, oder bei wechselnden Umweltbedingungen sinnvoll. Eine langfristige Baustelle führt bspw. zu einem veränderten Fahrverhalten und resultiert in anderen repräsentativen Profilen. Sind

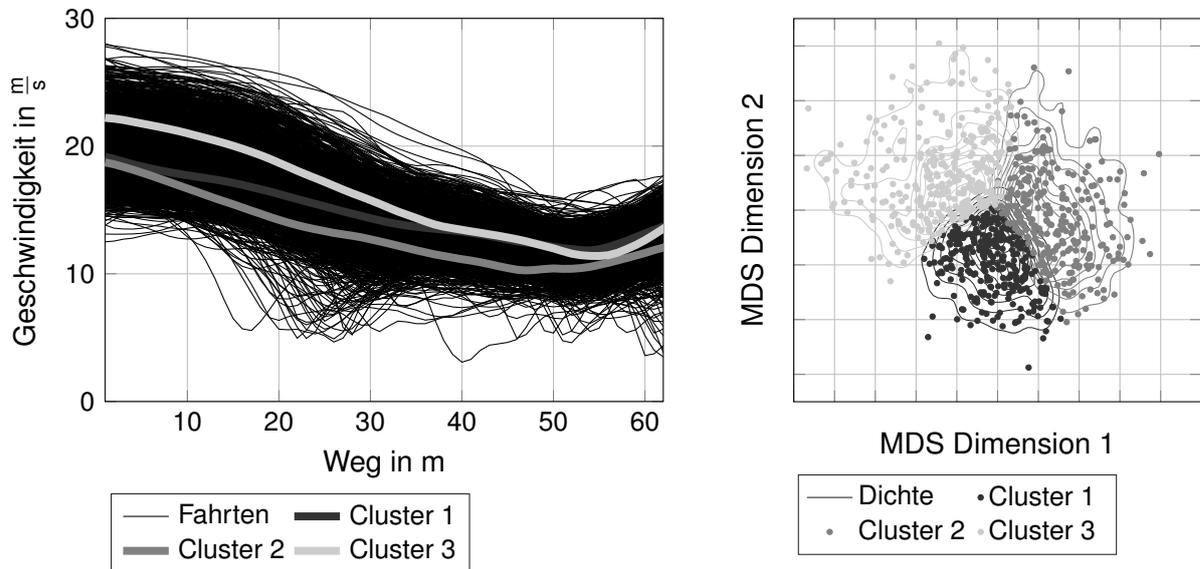


Abbildung 5.7 : Einfluss der Anzahl der Fahrten auf die Clustergenerierung

ausreichend Daten zu unterschiedlichen Witterungsbedingungen vorhanden, können diese gezielt zur Clusterbildung mit herangezogen werden.

5.2.3 Prädiktion des Fahrprofils

Ziel des Prädiktionsmodells ist die Zusammensetzung eines vollständigen Fahrprofils aus einzelnen repräsentativen Profilen aus dem vorangegangenen Kapitel in Abhängigkeit mehrerer Einflussfaktoren. Dabei soll das Ergebnis die bevorstehende Fahrt in Hinblick auf das Beschleunigungs- und Verzögerungsverhalten sowie der Geschwindigkeiten möglichst detailliert beschreiben. Zu den Einflussfaktoren zählen neben dem Verkehrsfluss auch das aktuelle Fahrverhalten des Fahrers. Daher wird die Prädiktion aktiv während der Fahrt angepasst, um auf Änderungen in der Situation zu reagieren. Nachfolgend werden die Modellschnittstellen näher erläutert. Anschließend wird das Prädiktionsmodell, sowie das Trainingsverfahren vorgestellt. Abschließend werden die vorgenommenen Modellanpassungen und Erweiterungen beschrieben.

Modelleingänge und Modellausgang

Die verwendeten Modelleingänge sind in Abb. 5.1 im zweiten Abschnitt dargestellt. Die Information zu der zu erwartenden Route oder den nächsten Streckenabschnitten nach Kapitel 4, dient der Zuordnung der entsprechenden Datenbasis für die jeweiligen Segmente. Zudem wird die mittlere Verkehrsgeschwindigkeit nach Kapitel 3.2.3 berücksichtigt. Alternativ kann die Tageszeit, der Wochentag bzw. Feiertage als Maß dienen, um das zu erwartende Verkehrsaufkommen oder den Verkehrsfluss einzubeziehen. Des Weiteren werden die Fahrprofile auf den zurückliegenden Streckenabschnitten hinzugezogen. Zum einen dienen diese zur Zuordnung zu den Clustern auf diesen Segmenten nach Kapitel 5.2.2 (Spektrales Clustering und Klassifizierung weiterer Fahrten), zum anderen können weitere Kennzahlen generiert werden. So kann das persönliche Fahrverhalten mit der durchschnittlichen Beschleunigung stärker gewichtet werden. Dazu werden die bereits zurückgelegten Segmente näher betrachtet. Es werden Beschleunigungs- und Verzögerungsvorgänge, die außerhalb der Standardabweichung (σ_{ap} und σ_{an}) für die Fahrten dieses Clusters liegen, gekennzeichnet. Treten diese mindestens einmal in den betrachteten Segmenten auf, wird dies bei der Prädiktion berücksichtigt. Die Werte für die Beschleunigung bzw. die Verzögerung werden getrennt voneinander betrachtet. Als Modellausgang wird für ein Segment das zu erwartende Cluster bestimmt. Über mehrere Segmente hinweg wird die Folge der Cluster berechnet, die nach Abb. 5.1 den Ausgang für weitere Untersuchungen darstellt. Aus den Clustern kann folglich das repräsentative Profil abgeleitet werden.

Um ein geeignetes Modell für die Prädiktion zu wählen, werden die Ein- und Ausgänge nach dem jeweiligen Skalenniveau [111] unterschieden: Bei der Nominalskala handelt es sich um Kategorien, bei denen keine Ordnung oder Wertung in den Daten ersichtlich ist (bspw. Postleitzahlen). Die Ordinalskala beschreibt Merkmale, bei denen eine Rang- oder Reihenfolge erkennbar und der Abstand dazwischen vergleichbar ist (bspw. Schulnoten). Weiter werden Werte einer Kardinalskala verwendet. Dabei handelt es sich um ein metrisches System mit interpretierbaren Abständen zwischen den Werten (bspw. Temperatur) [111].

Die oben genannten Modellgrößen werden gemäß der Skalenniveaus in Tab. 5.4 eingeordnet. Für die fahrer- und fahrtabhängige Clusterzuordnung der Fahrprofile liegt eine nominalskalierte Ausprägung vor. Bei den zusätzlich aus dem Fahrprofil erzeugten Kennzahlen wird eine nominalskalierte Ordnung angenommen. Der Verkehrsfluss TI weist eine Verhältnisskala auf. Hier ist eine metrische Betrachtung möglich. Falls ein anderer Verkehrsfluss (bspw. stufenweise von freifließend über stockenden Verkehr bis zu Stau) als Eingang verwendet wird, müsste diese als ordinalskalierte Ausprägung berücksichtigt werden. Die Zielgröße ist als Cluster des nächsten Straßenabschnitts ebenso ein nominalskalierter Wert.

Tabelle 5.4 : Übersicht über die Modellausprägungen

Typ	Name	Skala	Formelzeichen	Beispiele		
Fahrer	Cluster vorherige Segmente	nominal	CL	12	20	34
Fahrer	Kennzahlen Beschleunigung	nominal	A_p	$< \sigma_{ap}$	$> \sigma_{ap}$	
Fahrer	Kennzahlen Verzögerung	nominal	A_n	$< \sigma_{an}$	$> \sigma_{an}$	
Verkehr	Verkehrsfluss	metrisch	TI	$20 \frac{m}{s}$	$22 \frac{m}{s}$	$24 \frac{m}{s}$

Für die maschinelle Verarbeitung der nominalskalierten Ein- bzw. Ausgänge, ist eine Kodierung der Daten erforderlich. Da keine Rang- oder Reihenfolge vorliegt, können die Werte nicht nummeriert als eine Variable betrachtet werden. Um eine binäre Repräsentation der Daten zu erhalten, wird eine One-Hot-Kodierung (1-aus-n-Code oder Dummyvariable) angewendet [259]. Je nach Anzahl K der verschiedenen Kategorien wird die Variable in ebenso viele binäre Variablen überführt. Als Beispiel ergibt sich für eine Variable mit drei Kategorien folgende Möglichkeit:

Kategorie 1:[0, 0, 1] ,

Kategorie 2:[0, 1, 0] ,

Kategorie 3:[1, 0, 0]

Im Abschnitt 5.2.2 wird auf mögliche Fehler in der Datenaufzeichnung aufmerksam gemacht. Bei der Erzeugung der charakteristischen Fahrprofile und insbesondere der Erstellung der Cluster werden Ausreißer definiert, um diese nicht als eigenständige Kategorien mit aufzunehmen. In Folge dessen wird den Ausreißern kein Cluster zugeordnet. Bei der folgenden Erstellung des Prädiktionsmodells weisen diese Fahrten also, die auf den jeweiligen Abschnitt als nicht zuordenbar deklariert werden, fehlende Informationen auf. Es existieren verschiedene Methoden, um mit diesen fehlenden Daten umzugehen. In [260] werden die Varianten *Löschen des Eintrags* (Case Deletion) und *k-Nearest-Neighbor* als mögliche Lösungen vorgestellt. Die Ergebnisse sind abhängig von der Datenbasis und dem Anteil der fehlenden Daten. In dieser Arbeit werden die Fahrten mit fehlender Clusterzuordnung von der Betrachtung ausgeschlossen.

Multinomiale logistische Regression

Die Zielgröße unterscheidet in der Regel mehr als zwei Kategorien. Zur Prädiktion der Kategorie für das nächste Segment eignet sich eine Multinomiale Logistische Regression (MLR) und stellt eine Erweiterung der logistischen Regression dar. Letztere wird kurz eingeführt.

Bei der logistischen Regression erfolgt eine Analyse von dichotomen abhängigen Variablen, einer binären Entscheidung [261]. Der binäre Ereignisraum, im einfachsten Fall eine Ja/Nein-

Entscheidung, setzt voraus, dass die Summe der Wahrscheinlichkeiten eins ergibt [261]. Dies beschreibt Gl. 5.10 mit den N unabhängigen Variablen $\mathbf{x}_i = [x_1, x_2, \dots, x_N]$.

$$P(Y = 1|\mathbf{x}_i) + P(Y = 0|\mathbf{x}_i) = 1 \quad (5.10)$$

Die Kenngrößen und Beziehungen der logistischen Regression sind nach Backhaus [262] in Abb. 5.8 dargestellt.

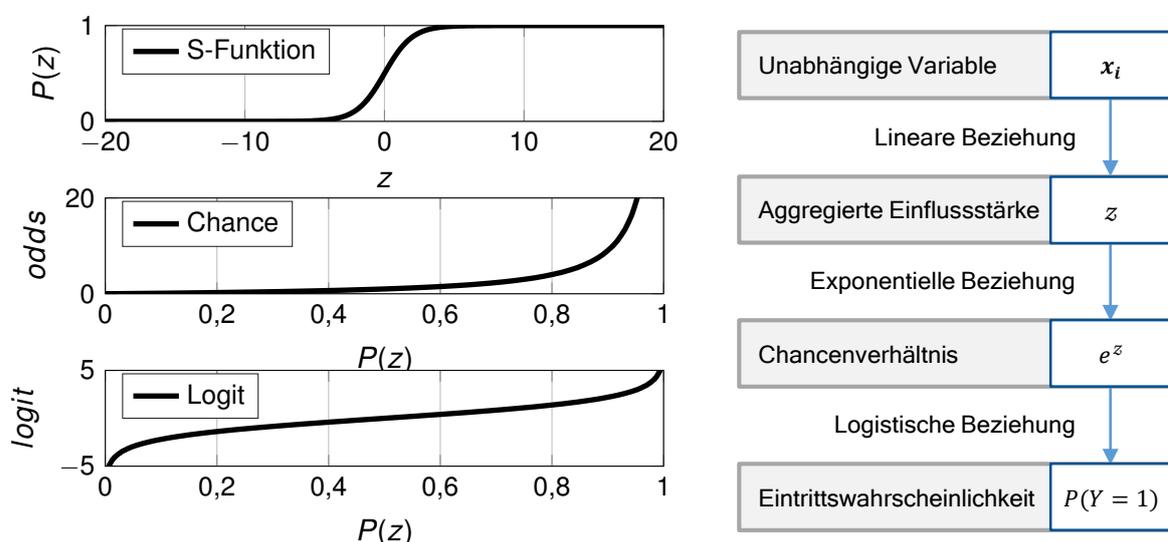


Abbildung 5.8 : Kenngrößen und Beziehungen der logistischen Regression

Das Modell basiert auf der logistischen Funktion [263, S. 4–28]

$$f(z) = \frac{1}{1 + \exp(-z)} \quad (5.11)$$

Wie in Abb. 5.8 dargestellt, handelt es sich um eine S-förmige Kurve mit den Grenzwerten 0 und 1. Die Funktion kann folglich das gewünschte Verhalten abbilden und ist abhängig von der aggregierten Einflussstärke z . Diese wird über einen linearen Zusammenhang aus der Summe der unabhängigen Variablen mit den Regressionskoeffizienten β gebildet.

$$z = \beta_0 + \sum_{n=1}^N \beta_n x_n \quad (5.12)$$

Der Unterschied zur linearen Regression liegt in der Betrachtung der abhängigen Variablen. Damit sich die Wahrscheinlichkeit für das Eintreten eines Ereignisses in dem Bereich zwischen 0 und 1 bewegt, werden zwei Transformationen vorgenommen. Anstatt der Eintrittswahrscheinlichkeit wird das Chancenverhältnis *odds* (5.13) betrachtet. Mit dem *odds* wird die Chance bestimmt mit welcher Wahrscheinlichkeit ein Ereignis eintritt im Gegensatz dazu, dass es nicht eintritt. Damit ist der Wert, wie in Abb. 5.8 zu sehen, allerdings noch nach unten beschränkt. Als weiteren Schritt wird der Logit-Wert *logit* (5.14) als Logarithmus des Chancenverhältnisses gebildet, um auch negative Werte zuzulassen [262, S. 283–356].

$$odds(z) = \frac{P(z)}{1 - P(z)} \quad (5.13)$$

$$logit(P(z)) = \ln \frac{P(z)}{1 - P(z)} \quad (5.14)$$

In diesem Raum kann eine Geradengleichung gefunden werden, welche die Eintrittswahrscheinlichkeit beschreibt. Die Regressionskoeffizienten können bspw. mit dem Maximum-Likelihood-Verfahren geschätzt werden. Für die Wahrscheinlichkeit des Eintretens des Ereignisses ergibt sich folglich

$$P(Y = 1|\mathbf{x}_i) = \frac{1}{1 + \exp(-(\beta_0 + \sum_{n=1}^N \beta_n x_n))}. \quad (5.15)$$

Bei der Erweiterung zur MLR wird die Regression für polytome, mehrteilige Variablen erweitert, weil oftmals mehr als zwei, also K Kategorien bei der Prädiktion möglich sind. Hierfür wird eine Referenzkategorie festgelegt und das *logit* der potentiellen Kategorie gegenüber der Referenzkategorie ermittelt. Für die verschiedenen Kategorien wird bis auf die Referenz jeweils ein Modell gebildet. Es resultiert für K Kategorien und N unabhängige Variablen eine Matrix mit $(K - 1)(N + 1)$ Regressionskoeffizienten [263, S. 450–452]. Die Wahrscheinlichkeit für das Eintreten der Referenzkategorie wird durch Gl. 5.16 beschrieben. Für das Eintreten jeder weiteren Kategorie gilt analog zu Kategorie 1 Gl. 5.17

$$P(Y = Ref|\mathbf{x}_i) = \frac{1}{1 + \sum_{k=1}^{K-1} \exp(-(\beta_{0k} + \sum_{n=1}^N \beta_{nk} x_n))} \quad (5.16)$$

$$P(Y = 1|\mathbf{x}_i) = \frac{\exp(-(\beta_{01} + \sum_{n=1}^N \beta_{n1} x_n))}{1 + \sum_{k=1}^{K-1} \exp(-(\beta_{0k} + \sum_{n=1}^N \beta_{nk} x_n))} \quad (5.17)$$

Training mit Maximum Likelihood

Die Schätzung der Regressionsparameter β erfolgt mit der Maximum-Likelihood-Methode [264] für nichtlineare Modelle. Dies zielt darauf ab, die kombinierte Wahrscheinlichkeit für alle Beobachtungen als Produkt der einzelnen Wahrscheinlichkeiten zu maximieren. In Folge dessen soll im binären Fall der Ausdruck

$$L(\beta|\mathbf{Y}) = \prod_{i=1}^I P(\mathbf{x}_i)^{Y_i} [1 - P(\mathbf{x}_i)]^{1-Y_i} \quad (5.18)$$

maximiert werden [262, S. 306]. Zur Vereinfachung dieser Berechnung wird der Wert logarithmiert, was auf Grund der streng monoton steigenden Funktion keine Auswirkung auf die Maximierung hat.

$$LL(\beta|\mathbf{Y}) = \sum_{i=1}^I \ln [P(\mathbf{x}_i)]^{Y_i} + \ln [1 - P(\mathbf{x}_i)]^{(1 - Y_i)} \quad (5.19)$$

Für das Maximum folgt, dass die erste Ableitung von LL gleich Null und die zweite Ableitung negativ sein muss. Die Regressionsparameter β werden folglich über alle I Beobachtungen ermittelt. Für das multinomiale Modell wird auf Czepiel [264] verwiesen.

Plausibilisierung

Nach der Schätzung der Parameter folgt die Überprüfung des Modells mit geeigneten Güte- oder Informationsmaßen [262]. Für das Trainingsverfahren wird zur Schätzung der Parameter die Maximum-Likelihood-Methode angewendet. Die Devianz (DEV) DEV bestimmt mit

$$DEV = -2(LL - LL_{sat}) \quad (5.20)$$

nach [262] einen weit verbreiteten Standardwert zum Vergleich verschiedener Modelle mit der gleichen Datenbasis. Dabei stellt LL_{sat} den Log-Likelihood Wert eines saturierten Modells dar. Dieser zeigt also das bestmögliche Modell (mit dem Extremfall $LL_{sat} = 0$). Dieser kann herangezogen werden, um die Güte mit anderen Modellen zu vergleichen. Der Wert ist abhängig von der Größe der Stichprobe. Je kleiner die Devianz, desto besser zeigt sich die Übereinstimmung des Modells mit der Datenbasis [262].

Die Devianz sinkt auch, wenn weitere unabhängige Variablen hinzugefügt werden, unabhängig von der Signifikanz. Daher werden zunehmend Informationsmaßzahlen eingesetzt, die versuchen dies zu korrigieren. Das Akaikes Informationskriterium (AIC) sowie das Bayessches Informationskriterium (BIC) berücksichtigen die Zahl der unabhängigen Parameter N . In Gl. 5.21 korrigiert das AIC die Devianz um die doppelte Zahl an unabhängigen Variablen. Der Wert wurde ursprünglich noch zusätzlich auf die Zahl der Beobachtungen normiert [265]. Das BIC korrigiert den Wert in Abhängigkeit der Stichprobenzahl deutlich stärker als das AIC. Beide Informationszahlen werden verwendet, um die Aufnahme von unterschiedlichen Prädiktoren in die Modelle zu bewerten.

$$AIC = DEV + 2N \quad (5.21)$$

$$BIC = DEV + N \ln(K) \quad (5.22)$$

Zur Bewertung der unabhängigen Variablen auf Signifikanz wird darüber hinaus der Wald-Test W unter der Betrachtung der p-Werte durchgeführt [265]. Mit den jeweiligen Regressionskoeffizienten β_{nk} , sowie deren Standardabweichung $\sigma_{\beta_{nk}}$ folgt:

$$W = \left(\frac{\beta_{nk}}{\sigma_{\beta_{nk}}} \right) \quad (5.23)$$

Für die p-Werte werden die Signifikanzniveaus $\alpha_5 = 5\%$, $\alpha_1 = 1\%$ und $\alpha_{01} = 0,1\%$ nach [266, S. 740] für *signifikant*, *sehr signifikant* und *hoch signifikant* verwendet.

Zur direkten Bewertung der Güte des Modells werden darüber hinaus Klassifikationstabellen eingesetzt. Als Beispiel ist Tab. 5.5 für drei Kategorien dargestellt. Diese stellen den beobachteten Werten die vorhergesagten Werte gegenüber. Daraus kann die Wahrscheinlichkeit für eine korrekte Prädiktion bestimmt werden. Die Trefferquote TQ folgt nach Gl. 5.24. Dieser Wert berücksichtigt jedoch nicht die ursprüngliche Verteilung der Beobachtungen. Daher sollte über die individuelle Trefferquote TQ_k (Gl. 5.25) die Prädiktionsgüte für jede Kategorie separat hinzugezogen werden.

Tabelle 5.5 : Klassifikationstabelle

		Prädiziert			Summe
		$K = 1$	$K = 2$	$K = 3$	
Beobachtet	$K = 1$	I_{11}	I_{12}	I_{13}	I_{1-}
	$K = 2$	I_{21}	I_{22}	I_{23}	I_{2-}
	$K = 3$	I_{31}	I_{32}	I_{33}	I_{3-}
	Summe	I_{-1}	I_{-2}	I_{-3}	I

$$TQ = \frac{\sum_{k=1}^K I_{kk}}{I} \quad (5.24)$$

$$TQ_k = \frac{I_{kk}}{I_{k-}} \quad (5.25)$$

Des Weiteren muss zur korrekten Erstellung der Modelle der Datensatz in einen Trainings- und Testdatensatz aufgeteilt werden. Die Trefferquote kann überhöht sein, falls für den Test des Modells der selbe Datensatz wie für das Training verwendet wird, weil bei der Erstellung des Modells versucht wird, eben diese zu maximieren [262, S. 318–319]. Dieser Effekt wird kleiner bei einer großen Anzahl an Trainingsdaten. Für kleine Datensätze wird die Leave-one-out-Methode eingesetzt. Dabei findet eine Kreuz-Validierung statt. Für jedes Element des Datensatzes wird

das Modell mit allen übrigen Daten neu trainiert und nur dieses Element wird zum Test verwendet. Dies hat den Vorteil, dass auch für kleine Datensätze die verwendeten Modellgrößen überprüft werden können.

Prädiktionshorizont

Die Prädiktion des Fahrzustandes erfolgt analog zu Kapitel 4.2.5 sequentiell in Abhängigkeit der vorgegebenen Route bzw. des Trips. Mit dem vorgestellten Verfahren der MLR wird jeweils für ein Straßensegment ein Prädiktionsmodell gebildet. Um einen längeren Prädiktionshorizont aufzuspannen, wird das Ergebnis der Prädiktion als Eingang für das darauffolgende Modell verwendet. Das Vorgehen ist in Abb. 5.9 dargestellt. Um den Fahrzustand, bzw. das entsprechende repräsentative Fahrprofil auf einem Straßensegment zu prädizieren, wird die Clusterzuordnung prädiziert. Als Eingang wird der Verkehrsfluss auf dem Straßensegment, die Clusterzuordnung der vorausgegangenen Segmente (in Abb. 5.9 sind dies zwei Segmente), sowie weitere Kennzahlen verwendet. Das ermittelte Cluster wird für die Prädiktion des nächsten Abschnitts als Eingang verwendet. Somit wird das vollständige Fahrprofil stückweise zusammengesetzt. Folgt das Fahrzeug der vordefinierten Route, können im Laufe der Fahrt einzelne Segmente mit diesem Ansatz angefügt werden, um die Länge des Prädiktionshorizonts aufrecht zu erhalten. Wird die Route geändert, oder weicht das tatsächliche Fahrprofil von der Prädiktion merklich ab, muss die Prädiktion für den vollständigen Horizont erneut ausgeführt werden (vgl. Kapitel 4.2.5). Auf Grund der Methode zur Erstellung der charakteristischen Profile kann es zu Sprüngen an den Segmentgrenzen kommen. Um ein zusammenhängendes Fahrprofil zu erstellen, wird ein gleitender Mittelwert an diesen Stellen eingesetzt.

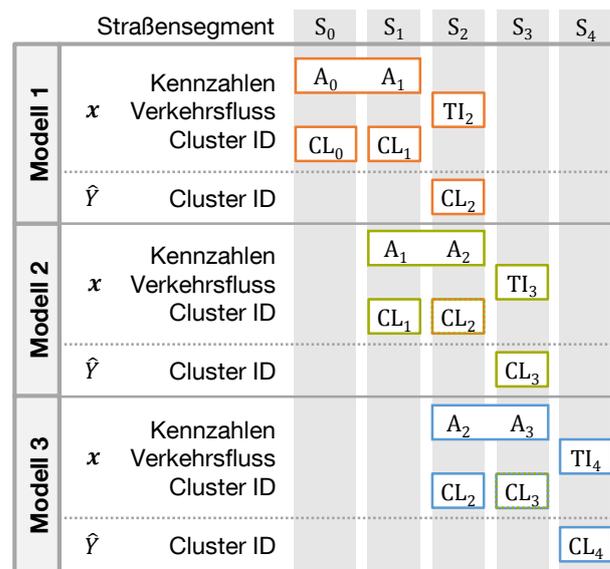


Abbildung 5.9 : Sequentieller Ansatz der Prädiktion

5.2.4 Prädiktion weiterer Bewegungsgrößen

Im Rahmen der Datenerfassung wurden mit den Smartphones, neben der GPS-Position, die Beschleunigungen und Drehraten in alle Raumrichtungen aufgezeichnet. Es geht darum, die zu erwartenden Bewegungsgrößen für die bevorstehenden Fahrtabschnitte abzuschätzen.

Die Längsbeschleunigung wird über die Ableitung der Fahrgeschwindigkeit ermittelt. Die Querbewegung und die Gierrate sind besonders für die Fahrzeugbewegung in Kurvenfahrten und damit die Fahrstabilität und Fahrsicherheit relevant. Die Beschleunigung in z-Richtung, bzw. die Nickbewegung des Fahrzeugs ist bei der Prädiktion von Fahrbahnunebenheiten von Interesse.

Methodik

Die Methodik zur Prädiktion der zusätzlichen Bewegungsgrößen ist in Abb. 5.1 dargestellt. Über das Clusterverfahren werden ähnliche Geschwindigkeitsverläufe kategorisiert. Der Vorgang erfolgt basierend auf Straßensegmenten. Die Fahrten weisen daher die identische Fahrtstrecke

und für die Fahrbahn ähnliche Bewegungsmuster auf. Fahrten, die dem gleichen Cluster zugeordnet werden, haben einen vergleichbaren Geschwindigkeitsverlauf und infolge dessen werden auch die weiteren Bewegungsgrößen Ähnlichkeiten zueinander aufweisen.

Es ist jedoch keine exakte Aussage über die Kurvenwahl oder die Fahrspur möglich, weil keine kamerabasierte Erfassung erfolgt ist. Die Kurvenwahl kann unter anderem von entgegenkommenden Fahrzeugen beeinflusst werden [267]. Es existieren Unterschiede durch den Fahrzeugtyp bzw. den Beladungszustand, was sich wiederum auf die Fahrdynamik auswirkt. Die Prädiktion der Bewegung wird in diesem Fall auf eine stationäre Betrachtung des Fahrzeugs beschränkt. Nach dem Einspurmodell [268, S. 97] folgt dann, dass die Querdynamik über die Beziehung

$$a_y = v(\dot{\psi} - \dot{\beta}) \quad (5.26)$$

ausgedrückt werden kann. Die Querbewegung a_y des Fahrzeugs ist dabei abhängig von der Gierrate $\dot{\psi}$ und der Schwimmwinkelgeschwindigkeit $\dot{\beta}$. Schultze [110] zeigt, dass für ein normales Fahrverhalten die Schwimmwinkelgeschwindigkeit vernachlässigt werden kann. Mehrjerdian [269] und Ammon [270] zeigen, dass bei konventionellen Fahrten im urbanen Bereich die maximal gefahrene Querbewegung bei ca. $5 \frac{m}{s^2}$ liegt. Mit diesen Annahmen ist eine stationäre und fahrzeugunabhängige Betrachtung von Gierrate und Querbewegung möglich.

Abschätzung der Datenqualität

Die Methode zur Datenerfassung mit Smartphones wird in Kapitel 3.1.1 erläutert. Hierbei wird auch die Datenqualität der Smartphones mit einer Referenzmesstechnik dargestellt. Im Flottenversuch ist keine exakte Positionierung der Smartphones im Fahrzeug im Vergleich zu interner Sensorik möglich. Hierfür ist eine aufwändige Nachbearbeitung der Daten wie bei Ritter [107] erforderlich, um auf eine exakte Aussage zu kommen. Diese Methodik ist jedoch nicht für den Einsatz mit großen Datenmengen geeignet. Im breiten Einsatz der Smartphones ist zudem nicht gewährleistet, dass diese dauerhaft und fest im Fahrzeug verbaut sind.

Um eine Abschätzung geben zu können, wie die Abweichung der Querbewegung und der Gierrate bei Geradeausfahrt ist, wird die Verteilung des Offsets über mehrere Fahrzeuge untersucht. Abb. 5.10 zeigt die relative Häufigkeit der Gierrate, sowie der Querbewegung für über 1 550 Fahrten auf einem einspurigen, geraden Straßensegment mit einer Länge von ca. 260 m. Dabei liegt der Mittelwert der Gierrate bei $-0,48 \frac{^\circ}{s}$, sowie die Standardabweichung bei $1,35 \frac{^\circ}{s}$. Die Gierrate ist symmetrisch und annähernd normalverteilt. Für die Querbewegung ergibt sich ein ähnliches Bild. Der Mittelwert liegt bei $0,08 \frac{m}{s^2}$. Die Standardabweichung bei $0,67 \frac{m}{s^2}$. Die Verteilung ist ebenso symmetrisch und annähernd normalverteilt. Die Aussagekraft der Untersuchung beschränkt sich auf die ermittelte Standardabweichung. Um Messfehler oder eine nicht fixierte Halterung des Smartphones auszuschließen, ist eine Vorsortierung der Fahrten über unplausible Verläufe und Größenwerte notwendig.

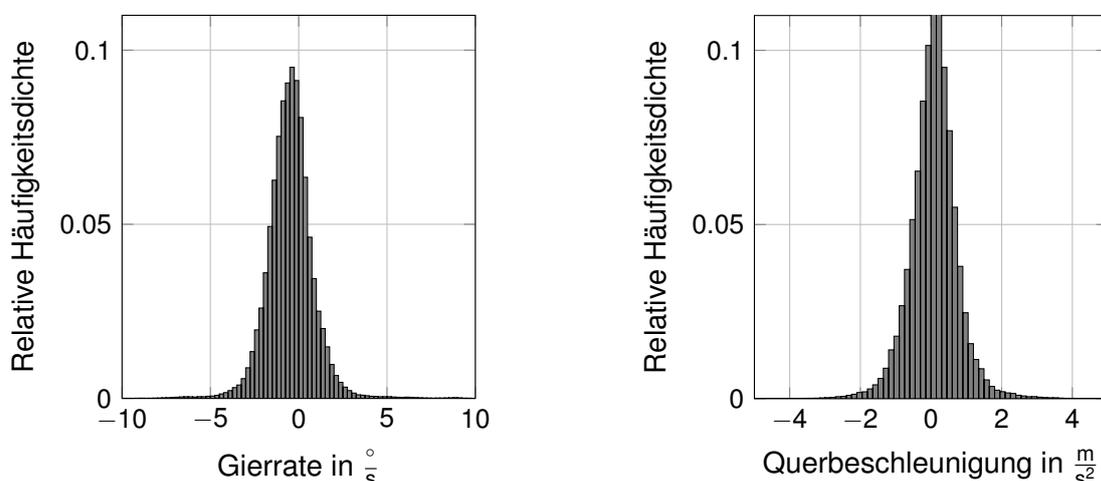


Abbildung 5.10 : Messwerte der Gierrate und der Querbewegung bei Geradeausfahrt

Auf Grund der möglichen Fehler und der Streuung der Messwerte ist keine exakte Aussage über das bevorstehende Bewegungsprofil möglich. Allerdings ist eine qualitative und über die Angabe von Konfidenzintervallen eine quantitative Bewertung möglich. Für die Auswertung der Daten werden folgende Annahmen getroffen: Das Display des Smartphones zeigt zum Fahrer und die Achsen der Koordinatensysteme des Fahrzeugs und des Smartphones sind parallel ausgerichtet. Für die Ermittlung der Orientierung des Smartphones wird die Erdbeschleunigung verwendet. Eine genauere Schätzung der Einbaulage, bzw. der Neigung des Smartphones wird nicht vorgenommen.

Gierrate

Die Erstellung der Konfidenzintervalle erfolgt einmalig pro Cluster mit den für die Clusterbildung verwendeten Fahrdaten. Neben der Geschwindigkeit ist für diese Datensätze eine vollständige Aufzeichnung von Gierrate und Querschleunigung erforderlich. Die zugehörigen Beschleunigungen und Gierraten werden gemäß der Orientierung ausgewählt. Anschließend werden die Daten um Ausreißer bereinigt. Anhand eines zweidimensionalen Histogramms wird die am häufigsten aufgetretene Gierrate ermittelt. Nach [268]

$$\dot{\psi} = \frac{v}{R} \quad (5.27)$$

folgt, dass bei gleichem Kurvenradius R (Abb. 5.11) die auftretende Gierrate allein von der Geschwindigkeit abhängig ist. Über das Verhältnis aller aufgezeichneten Geschwindigkeiten des Clusters zu dem charakteristischen Profil werden zu erwartende Ober- und Untergrenzen ermittelt. Zusätzlich wird die bereits beschriebene Standardabweichung um die Nulllage berücksichtigt. Mit den Grenzen werden unplausible Daten aussortiert. Auf den verbleibenden Datensatz wird ein Gauß-Filterverfahren [271, S. 108] angewandt. Zudem wird die Standardabweichung der Gierrate über die Distanz ermittelt und dient als Konfidenzintervall.

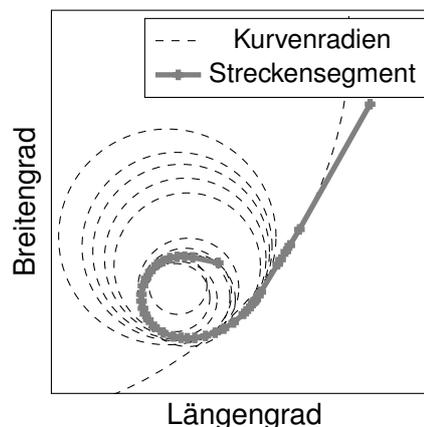


Abbildung 5.11 : Kurvenradien einer Autobahnausfahrt (Karte: OSM © ⓘ ⓘ)

Zur Veranschaulichung wird ein Cluster eines Straßensegments (Autobahnausfahrt, vgl. Geschwindigkeitsverlauf CL 1 in Abb. 5.18) mit über 600 Fahrten von 621 m Länge untersucht. Das resultierende Profil ist in Abb. 5.12 dargestellt.

Als Vergleich wird zudem in Abb. 5.12 die Gierrate abgebildet, die nach Gl. 5.27 über die repräsentative Geschwindigkeit und den aus den Kartendaten extrahierten Kurvenradien berechnet wurde. Auf Grund der deutlichen Abweichung von den aufgezeichneten Daten zeigt sich, dass die Karteninformationen nicht ausreichen, um eine qualitative Abbildung der Realität darzustellen.

Querschleunigung

Für die Bestimmung der Querschleunigung wird der identische Ansatz gewählt. Die zuvor getroffene Annahme mit der Vernachlässigbarkeit der Schwimmwinkelgeschwindigkeit bestätigt sich in Abb. 5.13. Mit Gl. 5.26 wird die Beschleunigung über die zuvor ermittelte Gierrate bestimmt. Für beide Kurven zeigt sich eine Übereinstimmung von $R^2 = 0,99$.

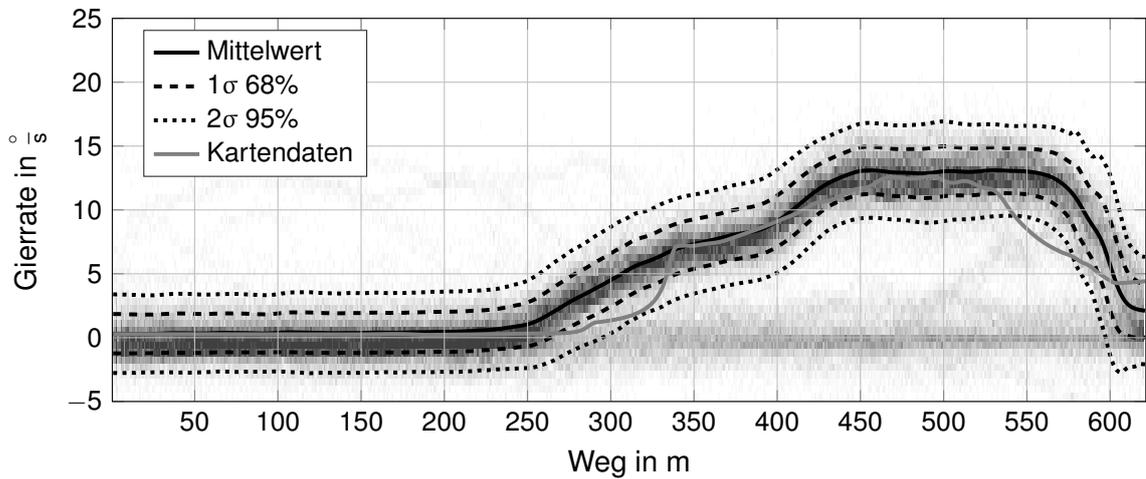


Abbildung 5.12 : Gierrate bei einer Autobahnausfahrt

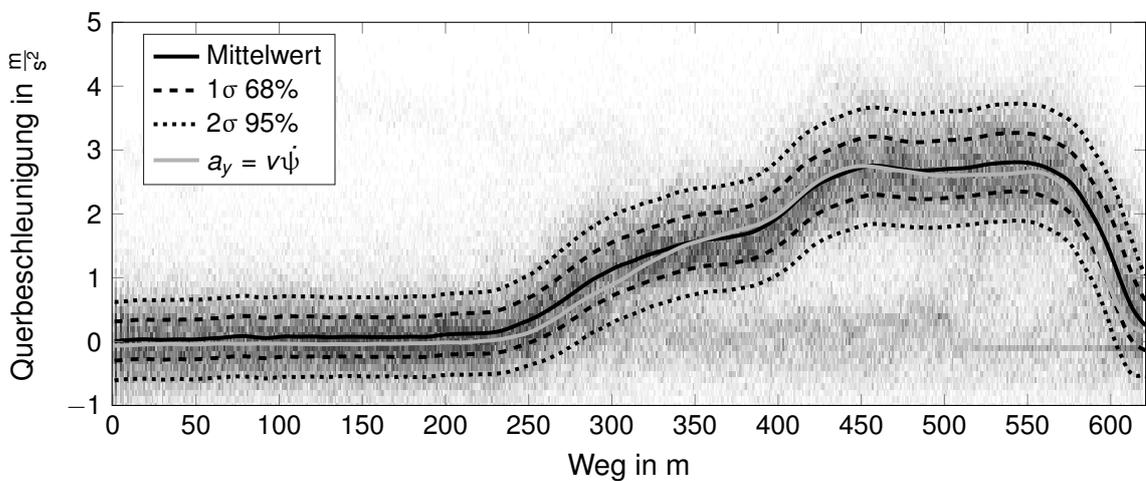


Abbildung 5.13 : Querbeschleunigung bei einer Autobahnausfahrt

Weitere Größen

Generell ist die Verwendung von internen Fahrzeugdaten mit eingebauter Messtechnik präziser. Eine erste Abschätzung ist aber auch mit dieser Methode möglich. Als weitere Einschränkung ist festzuhalten, dass keine Fahrspurunterscheidung möglich ist. Spurwechselvorgänge werden nicht betrachtet. Mit einer ausreichend großen Datenbasis und exakter Positionierung der Fahrzeuge wäre dies jedoch denkbar. Generell kann die vorgeschlagene Methode auch auf weitere Größen ausgeweitet werden, falls entsprechende Daten vorhanden sind. Bspw. könnten fahrzeugspezifische Werte wie der Lenkradwinkel, Fahrpedalstellungen oder das Motormoment mit betrachtet werden.

5.2.5 Ableitung von Metadaten

Zusätzlich zu den Bewegungsgrößen werden weitere Kennzahlen betrachtet. Dabei werden diese Daten erneut segmentspezifisch ausgewertet. Der Ansatz ist in Abb. 5.1 dargestellt. Analog zu den Bewegungsgrößen wird die Fahrzeugzugehörigkeit zu den Clustern, bzw. die repräsentativen Profile, verwendet, um die Metadaten für das jeweilige Segment zu bestimmen. Zuerst wird auf den Energieverbrauch näher eingegangen. Anschließend werden weitere Kennzahlen vorgestellt.

Energieverbrauch

Zur Ermittlung des erforderlichen Energieverbrauchs für ein Straßensegment wird ein Längsdynamikmodell eingesetzt. Die Fahrleistung wird über Gl. 5.28 aus der Geschwindigkeit und den

Fahrwiderständen bestimmt.

$$P = vF_A = v(F_L + F_R + F_{St} + F_B) \quad (5.28)$$

Die Antriebskraft F_A setzt sich nach [268, S. 50] aus den Fahrwiderständen zusammen. Dabei gilt es den Luftwiderstand F_L , den Rollwiderstand F_R , den Steigungswiderstand F_{St} , sowie den Beschleunigungswiderstand F_B zu überwinden. Für den Steigungswiderstand werden die Höhendaten aus Kapitel 3.1.2 herangezogen. Für eine detaillierte Beschreibung der Modellierung wird auf [272] verwiesen.

Für die Ermittlung des Energieverbrauchs über einen längeren Horizont, erfolgt die Berechnung mit Hilfe des repräsentativen Profils schrittweise innerhalb der Segmente. Somit werden mögliche Unstetigkeiten zwischen den Streckensegmenten vernachlässigt und der berechnete Wert kann fahrzeugspezifisch für das Profil gespeichert werden. Die Kartendaten können somit um spezifische, zu erwartende Energieverbräuche, angereichert werden. Neben der Prädiktion des Fahrprofils kann somit gleichzeitig der zu erwartende Energieverbrauch ermittelt werden. Dieser ist jedoch stark von den Fahrzeugeigenschaften abhängig.

Weitere segmentspezifische Kennzahlen

Neben dem Energieverbrauch können weitere segmentspezifische Kennzahlen erhoben werden. Als Beispiel wird in Abb. 5.14 die Fahrzeit für die beiden vorgestellten Straßensegmente gezeigt. Neben der Verteilung der Fahrdauer über die jeweiligen Cluster kann auch eine Fahrzeit über die repräsentativen Geschwindigkeitsprofile bestimmt werden. Des Weiteren kann eine mittlere Stillstandszeit pro Cluster für mögliche Ampelphasen bzw. häufige Stillstandsorte als Haltelinien extrahiert werden.

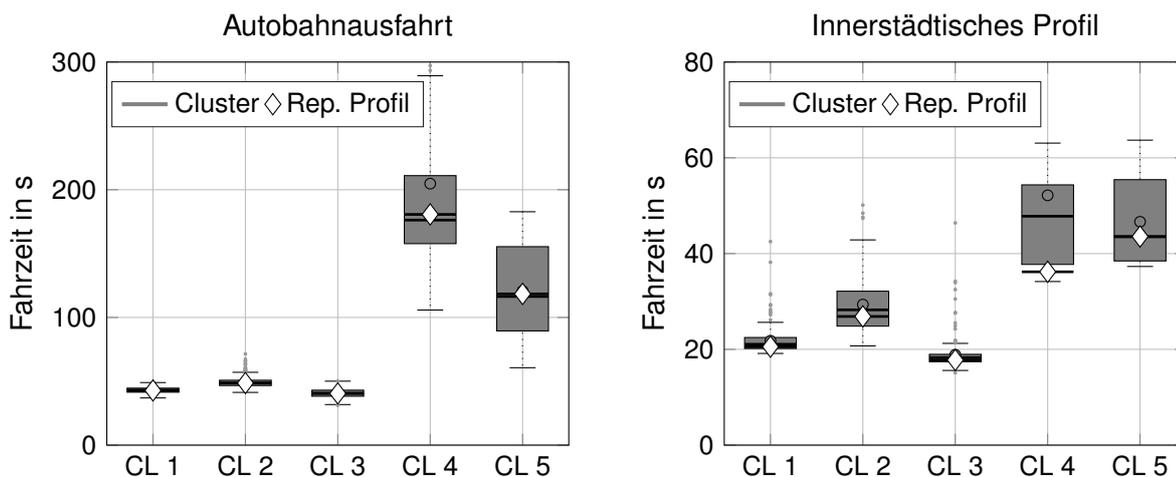


Abbildung 5.14 : Fahrzeit für ein innerstädtisches Segment

Darüber hinaus kann die Häufigkeit der Cluster über die Tageszeit ermittelt werden. Beispielhaft ist dies in Abb. 5.15 für eine Auffahrt auf den mittleren Ring dargestellt. Es zeigt sich eine Übereinstimmung der Fahrten im stockenden Verkehr mit den Verkehrsfluss [97].

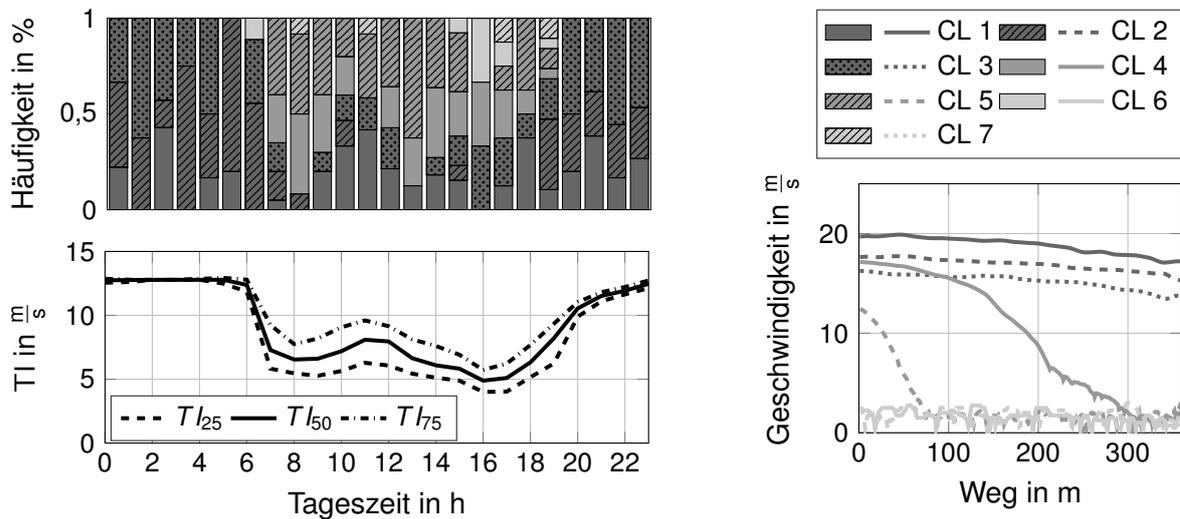


Abbildung 5.15 : Häufigkeit der Cluster für ein Straßensegment nach [97]

5.3 Ergebnisse

Die Ergebnisse zur Fahrprofilprädiktion werden wie folgt untergliedert. Es wird eine Bewertung der repräsentativen Profile vorgenommen. Anschließend wird beispielhaft ein Prädiktionsmodell für ein Straßensegment analysiert. Darauf folgend wird die Güte der Prädiktion für mehrere Segmente betrachtet und weitere Einflussfaktoren herausgehoben. Es folgt die Bewertung der Prädiktion über die Länge des Horizonts. Abschließend werden die prädizierten Fahrprofile zusammen mit den weiteren Bewegungsgrößen und Metadaten an Beispielfahrten veranschaulicht.

5.3.1 Bewertung der repräsentativen Profile

In Kapitel 5.2.2 wurde die Erzeugung der charakteristischen Fahrprofile erläutert. Die Profile spiegeln mögliche Zustände wider, welche die Fahrten für das jeweilige Segment annehmen können. Die repräsentativen Profile stellen dabei eine Näherung für alle Fahrprofile in dem zugehörigen Cluster dar. Zur Bewertung dieser Annäherung werden verschiedene Kenngrößen herangezogen. Nachfolgend wird auf die durchschnittliche Geschwindigkeit, die Fahrzeit, mittlere Beschleunigungen und den Energieverbrauch eingegangen. Abschließend erfolgt eine Bewertung der Zahl der Cluster pro Streckensegment.

Durchschnittliche Geschwindigkeit und Zeitdifferenz

Zum einen wird die durchschnittliche Geschwindigkeit auf dem Streckensegment betrachtet. Dabei wird die durchschnittliche Geschwindigkeit aller Fahrten in einem Cluster mit der des repräsentativen Profils verglichen. Zum anderen wird die Differenz der durchschnittlichen Fahrzeit mit der des charakteristischen Profils untersucht. In Abb. 5.16 sind diese Ergebnisse dargestellt.

Der quadratische Mittelwert der Zeitdifferenz liegt bei 1,24 s und damit die mittlere Abweichung bei 8,22 %. Die durchschnittliche Fahrzeit pro Segment beträgt 16,6 s. Der quadratische Mittelwert der Geschwindigkeitsdifferenz liegt bei $0,41 \frac{\text{m}}{\text{s}}$ und die mittlere Abweichung beträgt 3,55 %. Die Profile eignen sich daher auch, um damit Kennwerte zu der Fahrgeschwindigkeit und der Reisezeit abzuleiten. In Abb. 5.16 ist die relative Häufigkeitsdichte der durchschnittlichen Geschwindigkeit über alle ermittelten Cluster und Profile dargestellt. Es zeigt sich eine nahezu vollständige Überdeckung. Die Verteilung spiegelt mit dem hohen Anteil der durchschnittlichen Geschwindigkeit bei $8-12 \frac{\text{m}}{\text{s}}$ zudem den großen Anteil der innerstädtischen Straßensegmente in der Datenbasis wider.

Durchschnittliche Beschleunigungen

Des Weiteren wird die Übereinstimmung der Beschleunigungen mit dem Cluster untersucht. Abb. 5.17 zeigt die Differenz der mittleren Beschleunigungen \bar{a} , und deren positive \bar{a}_p und negative

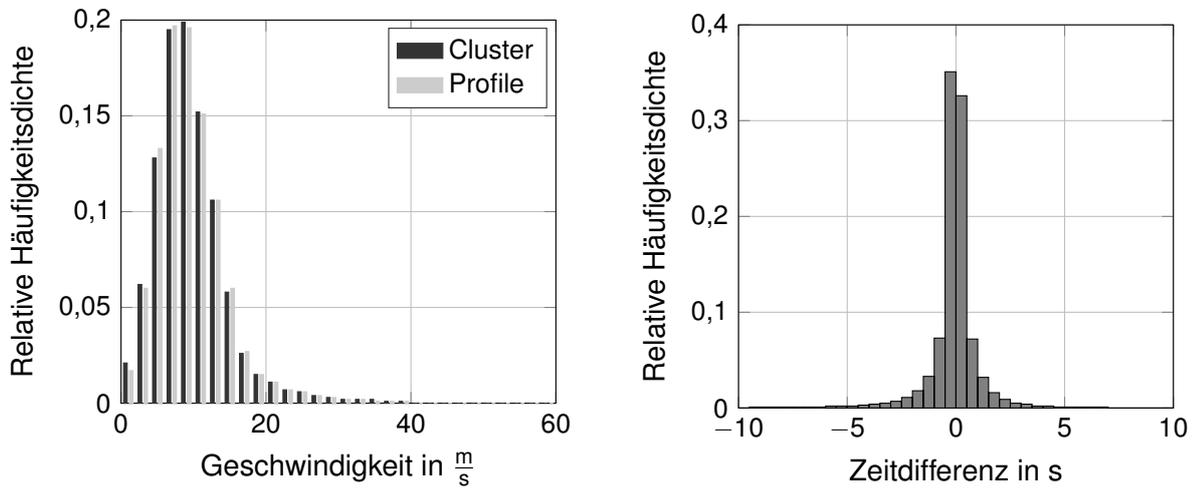


Abbildung 5.16 : Übereinstimmung der Profile mit den Cluster

\bar{a}_n Anteile. Der quadratische Mittelwert liegt für die mittlere Beschleunigung bei $0,31 \frac{m}{s^2}$. Der positive Anteil liegt bei $0,17 \frac{m}{s^2}$ und der negative Anteil bei $0,31 \frac{m}{s^2}$.

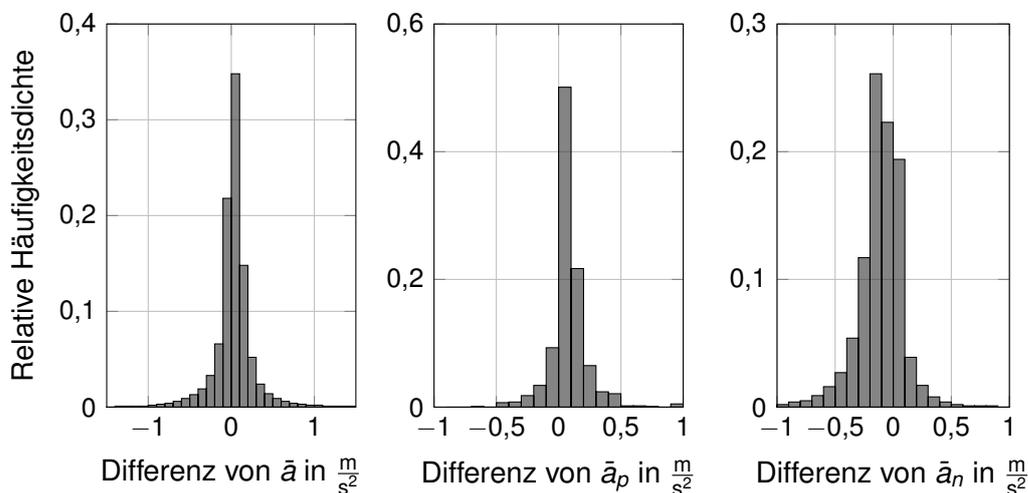


Abbildung 5.17 : Übereinstimmung der Beschleunigungen mit dem Cluster

Energieverbrauch

Zum Vergleich des Energieverbrauchs unterschiedlicher Fahrzeuge für ein Streckensegment werden die Fahrzeuge BMW i3, Nissan Leaf, Tesla Model S und VW eGolf herangezogen. Abb. 5.18 zeigt den unterschiedlichen Energieverbrauch der Fahrzeuge für eine Autobahnausfahrt. Dabei werden alle Fahrten mit den entsprechenden Fahrzeugen simuliert und der ermittelte Verbrauch für die jeweiligen Fahrprofilcluster in Boxplots dargestellt. Zudem wird der Verbrauch, ermittelt aus dem NEFZ-Fahrzyklus, für das Segment dargestellt. Es wird deutlich, dass der Verbrauch auf dem Segment von dem Fahrprofil abhängt. So ist es für manche Cluster (CL 3 und CL 5) möglich, auf dem Straßensegment zu rekuperieren, während bspw. im zähfließenden Verkehr ein erhöhter Energieverbrauch erforderlich ist.

Zudem wird der Energieverbrauch dargestellt, der für eine simulierte Fahrt mit dem repräsentativen Fahrprofil für das jeweilige Cluster ermittelt wird. In Abb. 5.18 zeigt sich eine gute Übereinstimmung für das repräsentative Profil und den mittleren Energieverbrauch für das Straßensegment.

Als ein weiteres Beispiel ist in Abb. 5.19 der Energieverbrauch für ein innerstädtisches Straßen-segment dargestellt. Wiederum zeigt sich eine Abhängigkeit des Verbrauchs von der Clusterzuordnung. Der Energieverbrauch, der über die repräsentativen Profile ermittelt wird, liegt in zwei Fällen unterhalb des unteren Quartils. Dies ist dem späten Haltevorgang, bzw. der Verzögerung

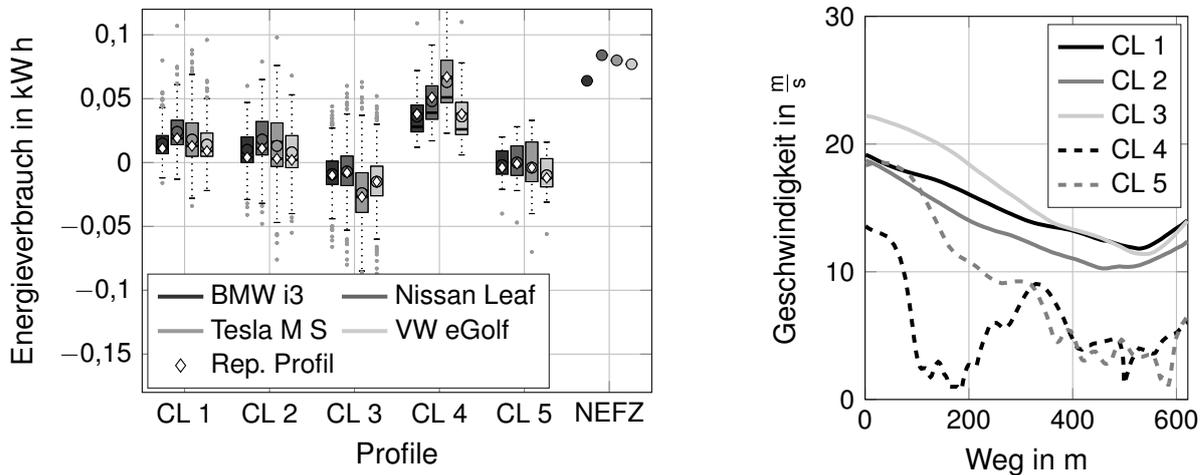


Abbildung 5.18 : Energieverbrauch bei einer Autobahnausfahrt

rungsphase geschuldet. Fahrten, die ebenfalls in diesem Cluster liegen, haben nach dem Stopp einen längeren Anfahr- bzw. Beschleunigungsvorgang und deshalb einen höheren Energieverbrauch. CL 2 zeigt zudem eine hohe Streuung des Energieverbrauchs und wenig vergleichbare repräsentative Profile, weil dieser keine direkte Einflussgröße bei dem Clusterverfahren ist.

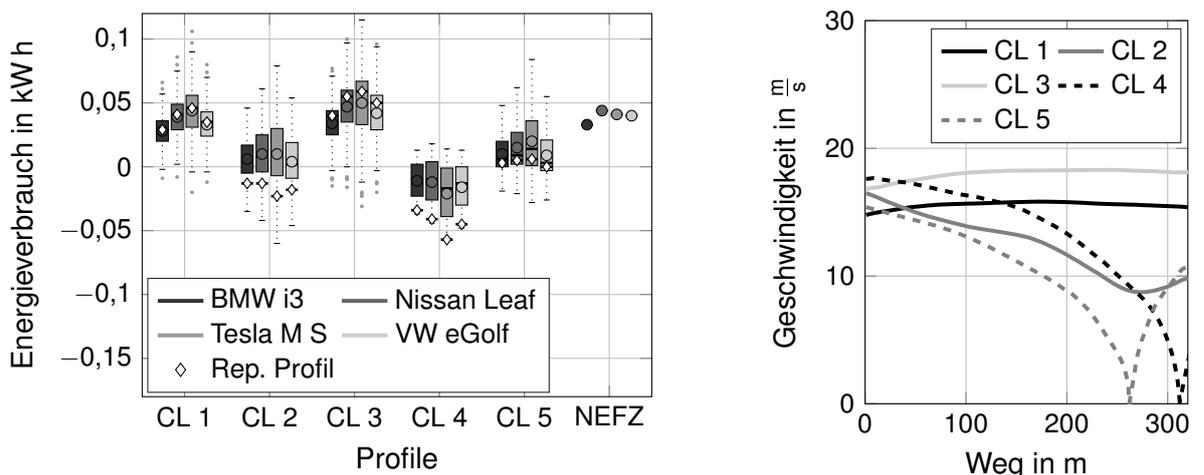


Abbildung 5.19 : Energieverbrauch für ein innerstädtisches Segment

Anzahl der Cluster pro Streckensegment

Für die Prädiktion des Fahrzustands dienen die nominalskalierten Cluster der vorangegangenen Straßensegmente über eine One-Hot-Kodierung als unabhängige Variablen. Die Zahl der Cluster pro Segment wirkt sich daher auf die Zahl der Regressionskoeffizienten aus. Die maximale Zahl der Cluster für ein Segment bezogen auf die routenabhängige Betrachtung liegt durch die haltestellenbedingte Vorsortierung und Parametrierung (max. fünf Cluster, Kapitel 5.2.2) bei 20 Cluster. In Abb. 5.20 wird die Anzahl der Fahrprofile pro Segment und Richtung in einer kumulierten Häufigkeitsverteilung dargestellt. Demnach werden bereits 50 % der Segmente mit nur zwei Profilen beschrieben. Bei 90 % der Segmente genügen sieben Profile für die Beschreibung.

Die Anzahl der Profile pro Straßensegment werden in Abhängigkeit von der Straßenklasse untersucht. In Abb. 5.21 ist sowohl der Anteil der Profile von der jeweiligen Straßenklasse, als auch die Aufteilung gemäß der beschreibenden Profile dargestellt. Es zeigt sich, dass Autobahnen in den häufigsten Fällen mit weniger als fünf Profilen beschrieben werden. Dies wird durch die Vorsortierung mit Hilfe der Haltevorgänge, die auf Autobahnen nur in Stausituationen eintreten, bekräftigt. Dementsprechend zeichnen sich die Segmente auf dem vergleichswei-

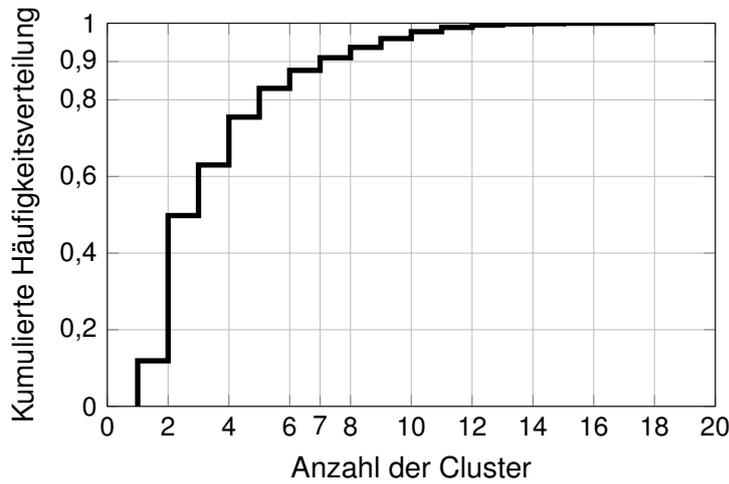


Abbildung 5.20 : Anzahl der Fahrprofile pro Segment und Richtung

se vielbefahrenen Sekundärstraßennetz (Klasse 21/22), das auch die Hauptverkehrswege in München darstellt, durch eine durchschnittlich höhere Zahl an charakteristischen Profilen aus.

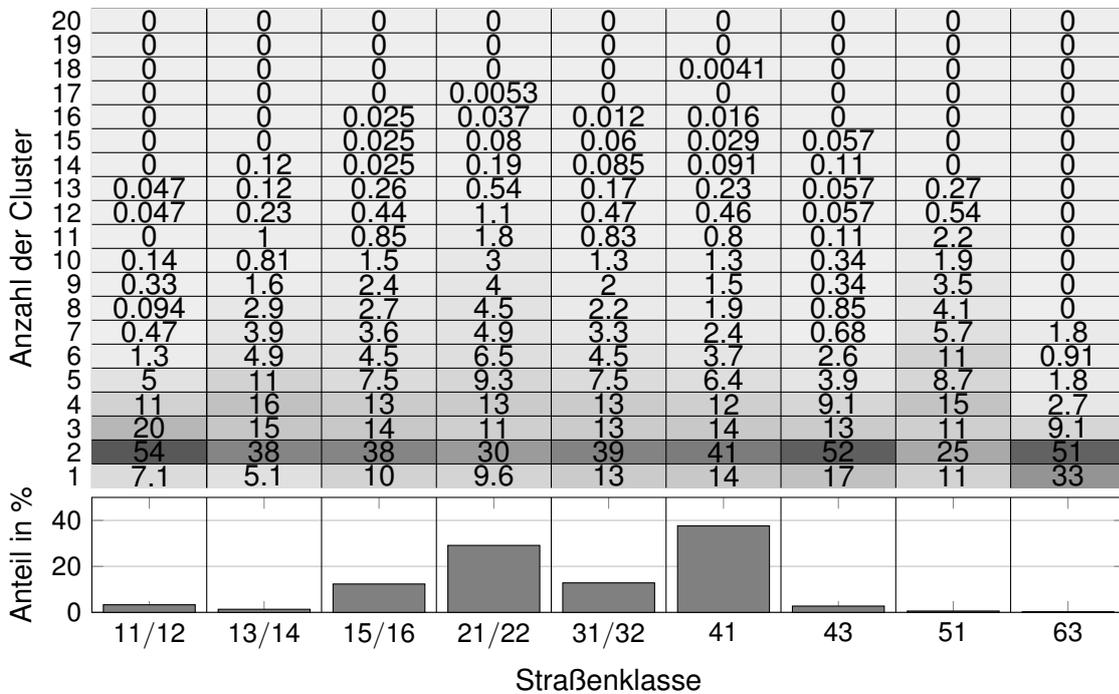


Abbildung 5.21 : Anzahl der Profile pro Straßenklasse (Tab. 3.2)

5.3.2 Analyse eines Prädiktionsmodells

Im folgenden wird ein Prädiktionsmodell für ein Straßensegment genauer erläutert. Hierfür wird zur Wahl der unabhängigen Variablen auf die Einflussfaktoren für die Prädiktion eingegangen. Anschließend wird exemplarisch der Signifikanztest der Regressionskoeffizienten durchgeführt. Abschließend wird das Klassifikationsergebnis für das Segment vorgestellt.

Übersicht über die Einflussfaktoren

Als Modellerweiterungen werden folgende Parameter für die Regression verwendet.

- A: Kennzahlen zum Beschleunigungs- und Verzögerungsverhalten des Fahrers, basierend auf den bereits befahrenen Streckensegmenten

- TI:
Verkehrsflussinformationen für das vorausliegende Streckensegment
- A+TI:
Eine Kombination beider Einflussparameter

Folgend wird der Einfluss der Parameter für die Regression anhand eines Segments erläutert. Es erfolgt die Prädiktion der Zuordnung der Fahrten zu einem Cluster auf dem nachfolgenden Streckensegment. Als Vergleichsbasis dient ein Modell mit ausschließlich der Clusterzuordnung des aktuellen Straßensegments, ohne weiteren Einflussfaktoren. Zur Überprüfung werden die in Kapitel 5.2.3 vorgestellten Informationsmaßzahlen DEV, AIC und BIC verwendet. Abb. 5.22 stellt die Informationsmaße für das Vergleichsmodell und die Modellerweiterungen.

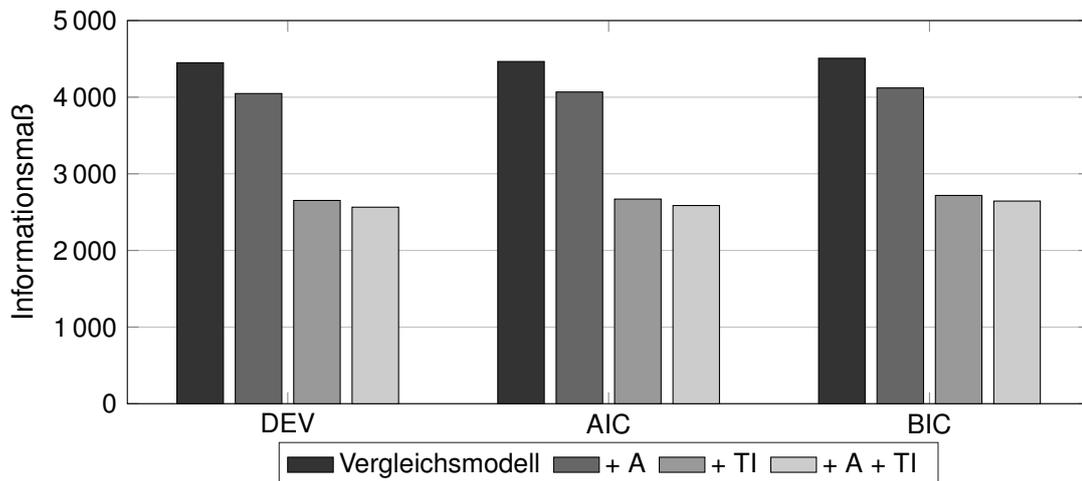


Abbildung 5.22 : Informationsmaßzahlen (Devianz, Akaike- und Bayessesches Informationskriterium) für ein Segment abhängig von den Eingangsgrößen

Für dieses Straßensegment zeigen alle Informationsmaßzahlen ein ähnliches Bild. In allen Fällen findet eine Verbesserung des Modells statt, dass sich durch eine niedrige Informationsmaßzahl kennzeichnet. Der größte Effekt zeigt sich in der Erweiterung um den Verkehrsfluss. Die Maßzahlen reduzieren sich auf ca. 60 % des ursprünglichen Werts. Aus der Kombination des Verkehrsflusses mit den Beschleunigungskennzahlen resultiert nur eine geringfügige Verbesserung.

Die Zahl der zurückliegenden Streckensegmente, die für die Prädiktion mit herangezogen werden, wird als weiterer Einflussfaktor untersucht. In Abb. 5.23 wird der BIC als Informationsmaßzahl dargestellt. Neben dem bereits vorgestellten Vergleichsmodell ist der Einfluss der zurückliegenden Streckensegmente, sowie die Zusatzinformationen wie in Abb. 5.22, dargestellt. Als Referenz dient der BIC des Modells ohne Erweiterung. In diesem Beispiel ist für die Verwendung eines weiteren Segments (Seg 1) nahezu keine Verbesserung ersichtlich. Bei zwei zurückliegenden Segmenten wird eine leichte Modellverbesserung erzielt. Eine deutliche Verbesserung des Modells wird durch die Hinzunahme des drittletzten Segments erreicht.

Demnach kann ein Einfluss der zurückliegenden Streckensegmente festgestellt werden. Allerdings ist dieser nicht zwingend auf die Reihenfolge der Segmente zu beziehen. Abschnitte die weiter zurückliegen können einen merklichen Einfluss auf das Prädiktionsmodell haben. Für den Verkehrsfluss und die weiteren Kennzahlen zeigt sich unabhängig von der Anzahl an Streckensegmenten ein ähnliches Bild. Die Kombination der zurückliegenden Segmente mit dem Verkehrsfluss bzw. der Beschleunigungskennzahl führt zu einer weiteren Verbesserung des Modells.

Signifikanztest

Anschließend wird ein Signifikanztest für das Modell mit dem niedrigsten Informationsmaß für diesen Streckenabschnitt durchgeführt. Insgesamt werden vier zurückliegende Segmente, die Kennzahlen zum Beschleunigungsverhalten und der Verkehrsindex betrachtet. Für das zu prädi-

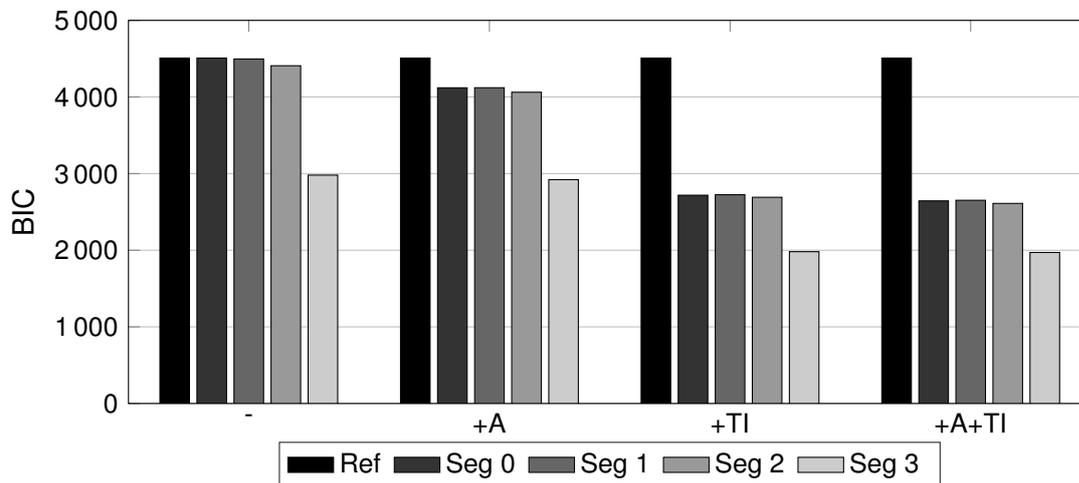


Abbildung 5.23 : Informationsmaßzahlen für ein Segment abhängig von der Anzahl der zurückliegenden Segmenten

zierende nächste Streckensegment stehen fünf repräsentative Profile zur Auswahl. In Tab. 5.6 sind die p-Werte für die entsprechenden Regressionsparameter hinterlegt. Die fünfte Kategorie wird als Referenzkategorie nicht angegeben, weil die Wahrscheinlichkeit für das Eintreten über Gl. 5.10 (bzw. für MLR Gl. 5.16 und 5.17) über alle anderen Kategorien berechnet werden kann. Neben der Regressionskonstante ist die Zahl der Regressionskoeffizienten abhängig von den möglichen Clustern der zurückliegenden Segmenten. Demnach sind im ausgewählten Beispiel für Seg_0 neun, für Seg_2 drei Cluster möglich. Ein Cluster dient ebenso als Referenz und wird für die Berechnung entfernt. Zwei Koeffizienten werden für die Kennzahlen zur Beschleunigung und ein Koeffizient für den Verkehrsfluss verwendet. Die Parameter werden nach dem Grad der Signifikanz ([266, S. 740]) unterschieden.

In allen Fällen sind die Parameter für den Verkehrsindex mindestens signifikant. Es zeigen manche Cluster in den Segmenten für die ein oder andere Kategorie einen signifikanten Beitrag. In diesem Fall hat das erste zurückliegende Segment keinen signifikanten Einfluss auf das Prädiktionsmodell. Dagegen hat das am weitesten zurückliegende Element in vielen Fällen einen hoch signifikanten Einfluss. Nicht signifikante Eingangsgrößen werden nicht entfernt und bleiben Bestandteil des Modells.

Klassifikation

Im nächsten Schritt wird an diesem Beispiel das Ergebnis der Klassifikation für das nächste Segment vorgestellt. Auf dem tertiären Straßensegment werden insgesamt 1 439 Fahrten den fünf möglichen Kategorien zugewiesen. Tab. 5.7 zeigt das Ergebnis. Die Trefferquote TQ für das Element beträgt 76 %. Dabei liegt die individuelle Trefferquote für die Klassen zwischen 65 % und 95 %. Diese ist unabhängig von der Zahl der Fahrten in dem Cluster.

5.3.3 Bewertung der Prädiktionsmodelle

Nach der Untersuchung der Einflussfaktoren und der Prädiktionsgüte für ein Segment wird dies nun auf mehrere Segmente ausgeweitet. Da für jedes Straßensegment ein eigenes Modell verwendet wird, ist für eine automatisierte Erstellung ein gleicher Satz an Eingangsgrößen vorteilhaft. Um die Wahl der Eingangsparameter global festzusetzen, werden mehrere Segmente gleichzeitig untersucht. Die Eingangsparameter werden wie in Kapitel 5.3.2 betrachtet. Anschließend wird der Einfluss der Datenbasis auf die Modellerstellung evaluiert. Als Basis hierfür dienen 375 Segmente, die jeweils mindestens 1 000 Fahrten aufweisen. Dabei wird für alle Modellvarianten die gleiche Basis an Fahrten auf den Segmenten verwendet, damit die Informationsmaßzahlen der einzelnen Varianten verglichen werden können.

Einflussfaktoren

Um die Einflussfaktoren zu bewerten, werden die Modelle mit den unterschiedlichen Ausprägung

Tabelle 5.6 : Beispiel p-Werte der Einflussfaktoren

Einfluss	$K = 1$	$K = 2$	$K = 3$	$K = 4$
β_0	0.050*	0.000***	0.000***	0.000***
	0.622	0.743	0.094	0.201
	0.239	0.697	0.087	0.092
	0.613	0.478	0.436	0.082
Seg_0	0.002**	0.047*	0.234	0.081
	0.001***	0.146	0.019*	0.165
	0.104	0.257	0.399	0.045*
	0.008**	0.202	0.142	0.390
	0.075	0.267	0.060	0.027*
		0.210	0.069	0.387
Seg_1	0.070	0.439	0.321	0.122
		0.010**	0.029*	0.001***
Seg_2	0.000***	0.002**	0.001***	0.002**
	0.000***	0.000***	0.405	0.001***
	0.000***	0.443	0.774	0.504
	0.000***	0.109	0.001***	0.504
	0.000***	0.127	0.002**	0.535
	0.900	0.417	0.109	0.558
A	0.004**	0.004**	0.978	0.543
	0.000***	0.431	0.004**	0.912
TI	0.017*	0.000***	0.000***	0.000***

*** = hoch signifikant, ** = sehr signifikant, * = signifikant
nach [266, S. 740]

Tabelle 5.7 : Beispiel Ergebnis Klassifikationstabelle tertiäres Straßensegment

		Prädiziert					Summe
		$K = 1$	$K = 2$	$K = 3$	$K = 4$	$K = 5$	
Beobachtet	$K = 1$	207	21	24	20	28	300
	$K = 2$	6	275	7	0	0	288
	$K = 3$	20	2	166	5	62	255
	$K = 4$	17	0	0	201	14	232
	$K = 5$	58	0	42	21	243	364
	Summe	308	298	239	247	347	1 439
TQ_k	69 %	95 %	65 %	87 %	67 %	76 %	

an Regressionsparametern untersucht. Dabei wird der Anteil der Segmente, bei denen eine Erweiterung des Modells zu einer Verbesserung, also einem kleineren Informationsmaß, führt, betrachtet.

In Abb. 5.24 ist diese Modellverbesserung für alle Informationsmaße dargestellt. Als Referenz dient wie im vorherigen Abschnitt das Modell, bei dem lediglich das aktuelle Straßensegment ohne zurückliegende Abschnitte, bzw. weiteren Kennzahlen verwendet wird. Bei allen Informationsmaßen ist ein vergleichbarer Effekt zu erkennen. Im Vergleich zur Referenz stellt die Erweiterung der Modelle mit den Kennzahlen zur Beschleunigung in über 94 % der Fälle eine Verbesserung dar. Im Vergleich zu dieser Konfiguration wird eine Erweiterung mit dem Verkehrsindex und dem aktuellen Segment betrachtet. 99 % der Segmente zeigen eine Verbesserung des Modells. Eine Kombination aus den Kennzahlen und dem Verkehrsindex bringt im Vergleich zur Variante mit lediglich dem Verkehrsindex seltener eine Verbesserung mit sich. In diesem Fall wird auch der Unterschied in der Betrachtung der Informationsmaßzahlen deutlich. Der BIC zeigt klarer, dass durch die Hinzunahme der Kennzahlen in nur ca. 71 % der Fälle eine Modellverbesserung erzielt wird, der DEV liegt bei 95 %.

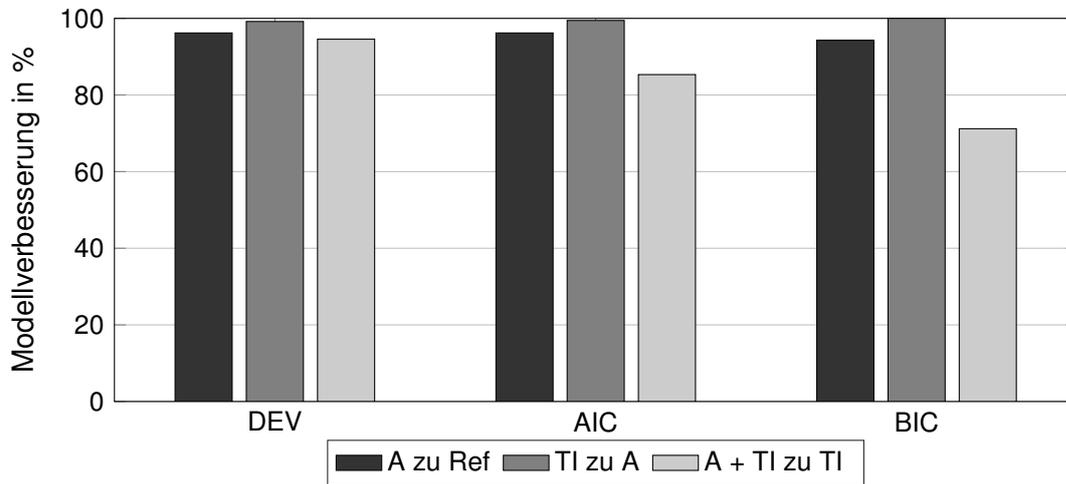


Abbildung 5.24 : Schrittweise Verbesserung der Modelle abhängig von den Erweiterungen

Um den Einfluss der zurückliegenden Streckensegmente auf die Modelle zu bewerten, wird erst der BIC qualitativ über alle untersuchten Segmente betrachtet. Dies ist in Abb. 5.25 dargestellt. Es ist zu erkennen, dass die zurückliegenden Segmente einen Einfluss auf die Modelle haben, besonders dann, wenn diese ausschließlich oder nur mit den Kennzahlen für das Beschleunigungsverhalten aufgebaut werden. Die Kennzahlen haben nur eine geringe Relevanz für die Modelle. Das Beschleunigungs- und Verzögerungsverhalten wird durch die Fahrten, die außerhalb der Standardabweichung liegen, beschrieben. Dieses Fahrverhalten auf den zurückliegenden Segmenten ist nicht zwingend auf den zu prädizierenden Segmenten identisch. Einen wesentlicheren Einfluss zeigt der Verkehrsindex. Allerdings ist auch eine weitere Verbesserung durch eine Kombination aus den zurückliegenden Segmenten mit dem Verkehrsfluss ersichtlich. Der Verkehrsindex bezieht sich direkt auf das zu prädizierende Segment. Sowohl freifließender Verkehr, als auch Stau haben einen signifikanten Effekt auf das Geschwindigkeitsprofil.

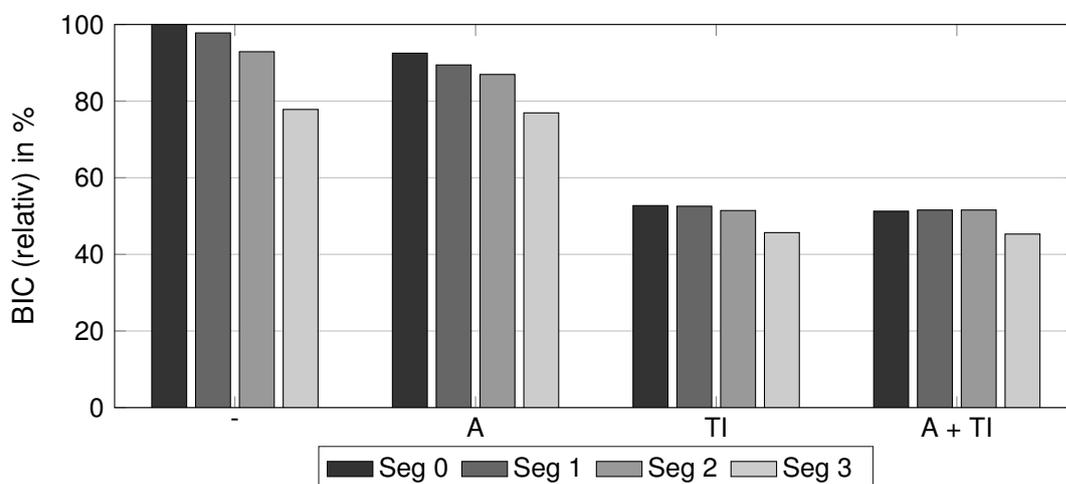


Abbildung 5.25 : Einfluss der zurückliegenden Segmente über alle Modelle

Anschließend wird erneut eine Untersuchung der Modellverbesserung anhand des BIC durchgeführt. Abb. 5.26 zeigt hierzu die Ergebnisse. Die Referenz entspricht dem Modell ohne weitere zurückliegende Segmente. Ebenso werden keine zusätzlichen Modellerweiterungen verwendet. Durch die Hinzunahme ausschließlich zurückliegender Segmente verbessert sich das Modell von Segment zu Segment in ca. 80 % der Fälle. Ein ähnliches Resultat ist bei der Erweiterung um die Kennzahlen für die Beschleunigung zu erkennen. In 99 % der Fälle steigert sich die Modellqualität bei Hinzunahme des Verkehrsindex. Hier wird deutlich, dass das erste zurückliegende Segment in nur ca. 50 % der Fälle eine Modellverbesserung bewirkt. Jedoch steigert das dritte zurückliegende Segment dagegen in über 75 % der Fälle die Modellqualität.

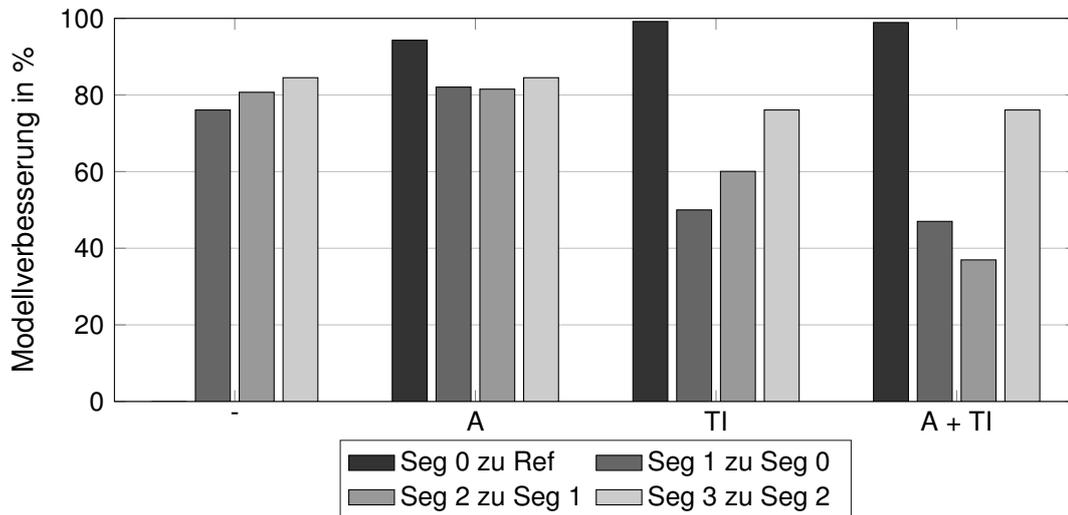


Abbildung 5.26 : Schrittweise Verbesserung der Modelle abhängig von den zurückliegenden Segmenten

Einfluss der Datenbasis

Bei der Erweiterung der Modelle mit Parametern aus den zurückliegenden Segmenten wird die verwendbare Datenbasis eingeschränkt. Aus allen aufgezeichneten Fahrten, die über das Segment gingen, kann nur das Subset an Fahrten verwendet werden, die alle aufeinander folgenden Segmente beinhalten. Durch Abbiegevorgänge, Ausreißer in dem Clusterverfahren oder Signalunterbrechungen reduziert sich die Zahl der Fahrten für jedes hinzugefügte Segment. Dies wird in Abb. 5.27 veranschaulicht. Die relative Anzahl der Fahrten für die betrachteten Modelle in Abhängigkeit der zurückliegenden Streckensegmente werden dargestellt. Im Falle der Berücksichtigung von drei weiteren Segmenten reduziert sich die Datenbasis auf 52 % der ursprünglichen Fahrten des betrachteten Segments.

Gleichzeitig steigt die Zahl der Regressionskoeffizienten mit zunehmenden Streckensegmenten deutlich an. Hier verdreifachen sich die Koeffizienten. Bei der Berücksichtigung der Modellverbesserung wird der BIC als Informationsmaß verwendet. Dieser berücksichtigt bereits die Zahl der Koeffizienten, sowie die Größe der Datenbasis. In den oben genannten Fällen wird die Datenbasis konstant gehalten. Ein niedrigerer BIC zeigt daher trotz einer größeren Zahl an Regressionskoeffizienten eine Verbesserung des Modells.

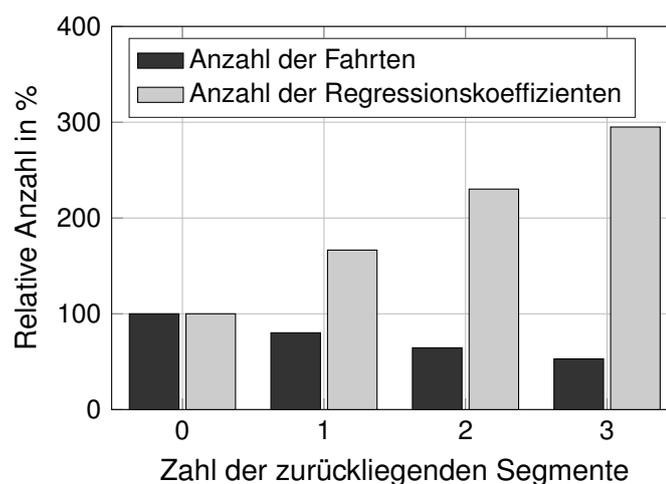


Abbildung 5.27 : Einfluss der Anzahl der zurückliegenden Segmente auf die Datengrundlage

Klassifikation

Abschließend wird für alle untersuchten Segmente das Ergebnis der Klassifikation vorgestellt. Hierfür wird für die betrachteten Elemente eine 10-fache Kreuz-Validierung mit je zwei Versuchsreihen durchgeführt. Die Modelle werden mit je vier zurückliegenden Segmenten und

dem Verkehrsindex aufgebaut. Da der Einfluss der Beschleunigungskennzahl in dem vorherigen Abschnitt vergleichsweise gering ausgefallen ist, wird dieser im Folgenden vernachlässigt.

Abb. 5.28 zeigt den Einflussfaktor der Straßenkategorie auf die resultierende Trefferquote. Dabei entspricht die Kategorie von Autobahnen den Straßenklassen 11-14, die der Landstraßen bzw. Sekundärstraßen den Klassen 15-22 und der Innenstadtbereich den Straßenklassen größer als 22. Der Einfluss auf die Trefferquote ist ersichtlich. Während die Trefferrate für Autobahnen im Mittel bei über 96 % liegt, sinkt der Wert für Landstraßen bereits auf 83 %. Besonders anspruchsvoll ist die Prädiktion im Innenstadtbereich. Hier liegt der Mittelwert bei 78 %, die Spannweite der zu erwartenden Ergebnisse ist darüber hinaus vergleichsweise groß. Besonders im Innenstadtbereich ist durch Ampelschaltungen, Fußgängerüberwege oder unvorhersehbare Verkehrssituationen die Prädiktion schwierig. Dies spiegelt sich auch in der Trefferquote wider.

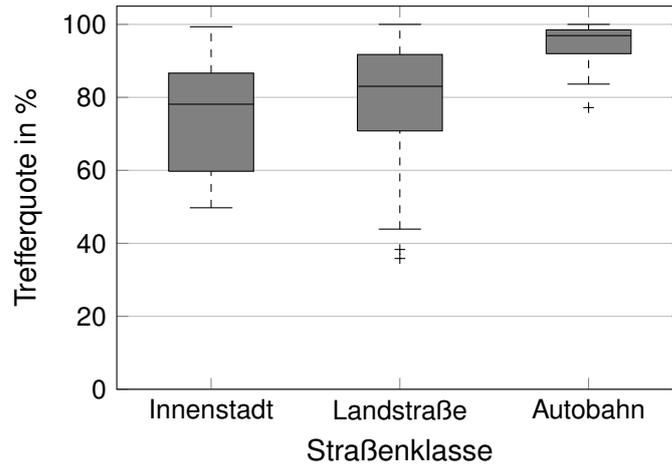


Abbildung 5.28 : Einfluss der Straßenkategorie auf die Trefferrate

Des Weiteren wird der Einfluss der Anzahl der Kategorien für das nächste Streckensegment betrachtet. In Abschnitt 5.3.1 und Abb. 5.21 werden die repräsentativen Profile und deren Zahl pro Segment bereits bewertet. Der Effekt auf die Trefferrate wird in Abb. 5.29 dargestellt. Je mehr Möglichkeiten für das nächste Segment bestehen, desto geringer fällt die Trefferquote aus. Eine korrekte Zuweisung wird schwieriger. Hier zeigt sich auch der Bezug zur Straßenklasse. Da für Autobahnen auch weniger Kategorien benötigt werden, steigt die Quote an.

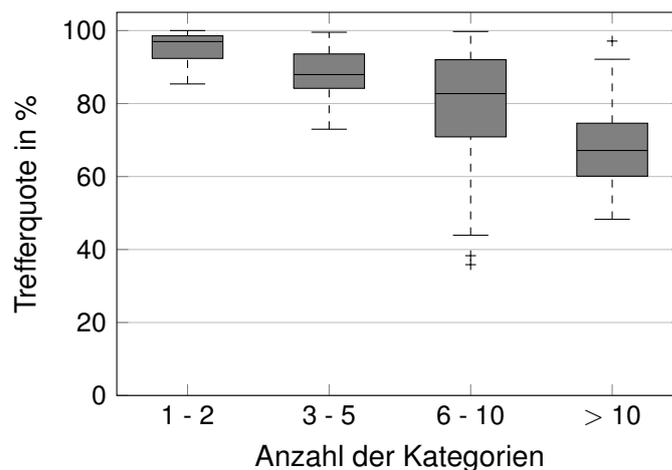


Abbildung 5.29 : Einfluss der Anzahl der Kategorien auf die Trefferrate

5.3.4 Prädiktionshorizont

Um den Einfluss des Horizonts auf das Prädiktionsergebnis zu bewerten, werden exemplarisch drei Routen näher untersucht. Eine Übersicht zu den ausgewählten Strecken ist in Tab. 5.8

gegeben. Die Strecken liegen jeweils auf den Straßenkategorien von Autobahnen, Landstraßen bzw. Sekundärstraßen und im Innenstadtbereich.

Tabelle 5.8 : Übersicht zu den ausgewählten Strecken

	Innerorts	Landstraße	Autobahn
ID Startknoten	71444	234673	539169
ID Zielknoten	4217	211047	24122
Anzahl der Segmente	30	30	30
Strecke in km	3,6	4,11	48,2
Durchschnittsgeschwindigkeit in $\frac{m}{s}$	11,75	12,3	29,1
Durchschnittlicher Energieverbrauch (BMW i3) in kW h	0,27	0,32	7,47
Mittlere Trefferrate in %	81,0	82,0	92,2
25tes Perzentil der Trefferrate in %	69,0	80,0	88,5
75tes Perzentil der Trefferrate in %	93,1	93,4	99,9
Durchschnittliche Anzahl an Cluster	7,16	4,65	2,72

Ähnlich zu dem Einfluss der Anzahl der zurückliegenden Streckensegmente auf die Datengrundlage in Abb. 5.27, ist die Zahl der durchgängigen Fahrten hinweg über alle Segmente auf einer Route begrenzt. Dies wird in Abb. 5.30 dargestellt. Für die Evaluation werden die Modelle wie im vorherigen Abschnitt mit je vier zurückliegenden Segmenten und dem Verkehrsindex erstellt. Die Modelle werden dabei mit allen verfügbaren Fahrten trainiert. Zur Bewertung werden nur die vollständig durchgängigen Fahrten herangezogen.

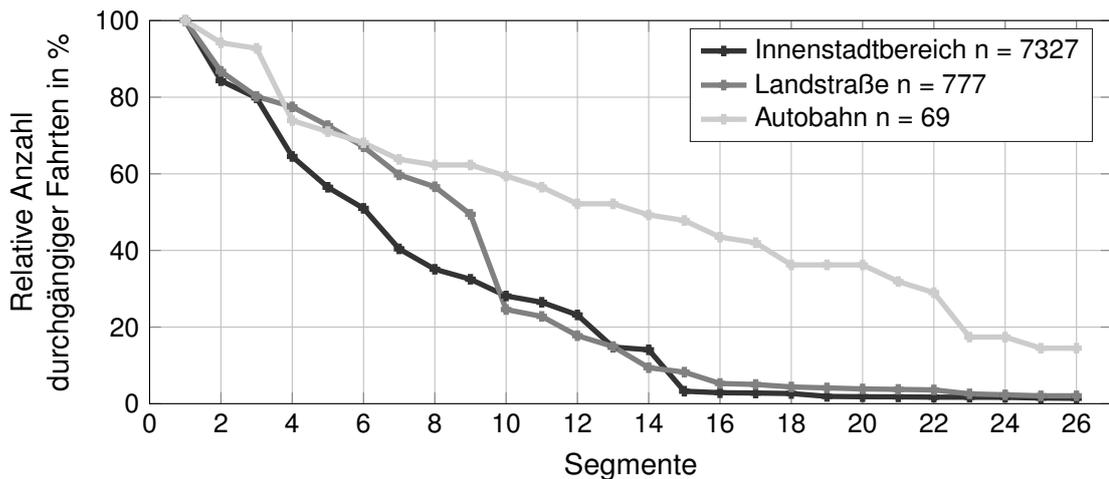


Abbildung 5.30 : Einfluss des Horizonts auf die Datengrundlage

Der Prädiktionshorizont wird als Anzahl der vorausliegenden Streckensegmente definiert. Die Güte der Prädiktion ist nicht direkt mit der Güte der Trip-Prädiktion in Abb. 4.12 vergleichbar. Bei der Trip-Prädiktion führt ein falsch vorhergesagtes Segment dazu, dass alle folgenden Segmente zwingenderweise nicht korrekt sind. Dies ist bei der Fahrprofilprädiktion nicht der Fall, weil eines der repräsentativen Profile bestimmt wird. Bei einem fortlaufenden Horizont kann daher auch nach einer falschen Prädiktion im vorherigen Segment das nächste korrekt bestimmt werden. Die Prädiktion ist jedoch abhängig von den vorherigen Segmenten.

Um die Treffquote über den Horizont zu bestimmen, werden die vorgestellten Strecken in Abschnitte aufgeteilt und beginnend von dem Startpunkt der Fahrt eine Prädiktion über die Horizontlänge durchgeführt. Dies ist exemplarisch für eine Autobahnfahrt in Abb. 5.31 dargestellt. Das Ergebnis für den Einfluss des Horizonts auf die Trefferrate wird in Abb. 5.32 gezeigt. Hierbei wird wiederum die Straßenkategorie unterschieden. Vor allem für die Fahrten innerorts zeigt sich, dass mit zunehmendem Prädiktionshorizont eine größere Spreizung der Trefferrate über die Segmente auftritt. Der mittlere Wert der Quote verändert sich unwesentlich.

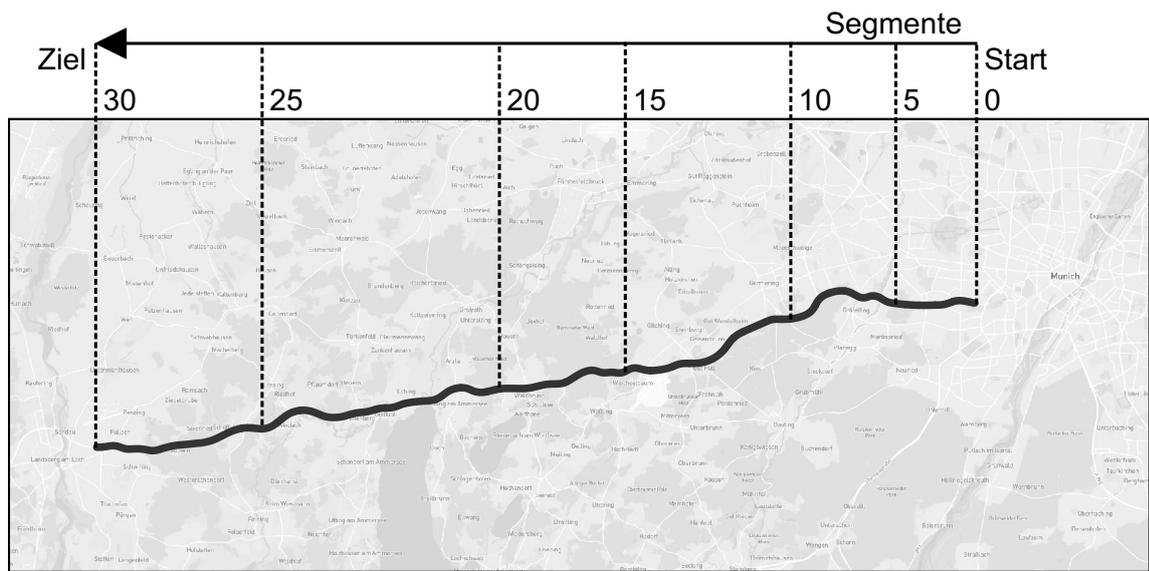


Abbildung 5.31 : Streckenlänge und Segmentabschnitte für die Autofahrt (Karte: OSM ©)

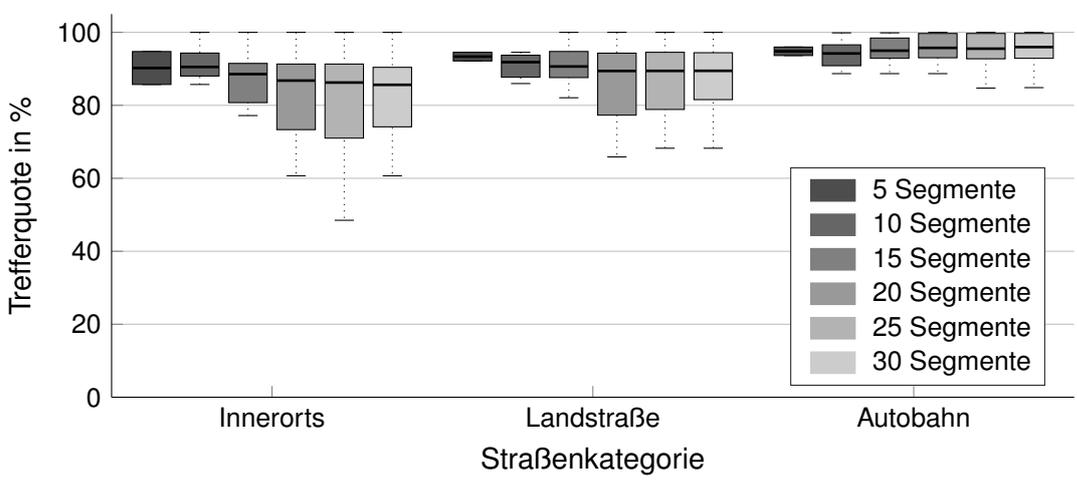


Abbildung 5.32 : Einfluss des Horizonts auf die Trefferquote

5.3.5 Beispielfahrt

Die resultierenden Ergebnisse für zusammenhängende Fahrprofile werden wie folgt untersucht: Anhand der in Tab. 5.8 vorgestellten Strecken wird der Einfluss des Verkehrs auf das Fahrprofil veranschaulicht. Anschließend wird für eine Fahrt mit kürzerem Prädiktionshorizont das Aktualisierungsverfahren der Prädiktion vorgestellt. Darüber hinaus wird auf die Prädiktion von weiteren Bewegungsgrößen und abschließend auf die Ermittlung des Energieverbrauchs für eine längere Fahrt eingegangen.

Verkehrseinfluss

Es wird der Einfluss des Verkehrs auf das Beispiel der Fahrt auf einer Landstraße untersucht. Abb. 5.33 zeigt drei unterschiedliche Situationen der identischen Route. Neben den drei realen Fahrten wird jeweils ausgehend von den ersten vier Segmenten die vollständige Route unter Verwendung des Verkehrsflusses prädiziert. Dabei wird die Bedeutung der Verkehrssituation deutlich. Im ersten Fall ist das resultierende Fahrprofil dem aufgezeichneten Profil ähnlich. Die Geschwindigkeit liegt bis auf zwei Verzögerungsvorgänge an der erlaubten Höchstgeschwindigkeit.

Im zweiten Fall ist die durchschnittliche Geschwindigkeit niedriger. Die Prädiktion kann die Tendenz mit den Beschleunigungs- und Verzögerungsvorgängen abbilden. Eine kurzzeitige,

starke Verzögerung gegen Ende der Fahrt (Position: 2850 m) wird von dem Modell nicht vorhergesehen.

Im letzten Fall handelt es sich um einen vergleichsweise stockenden Verkehrsfluss mit mehreren Haltevorgängen. Die Tendenz wird auch in dieser Situation widerspiegelt. Manche Verzögerungen (Positionen: 1600 m, 3200 m) werden nicht im gleichen Maße detektiert.

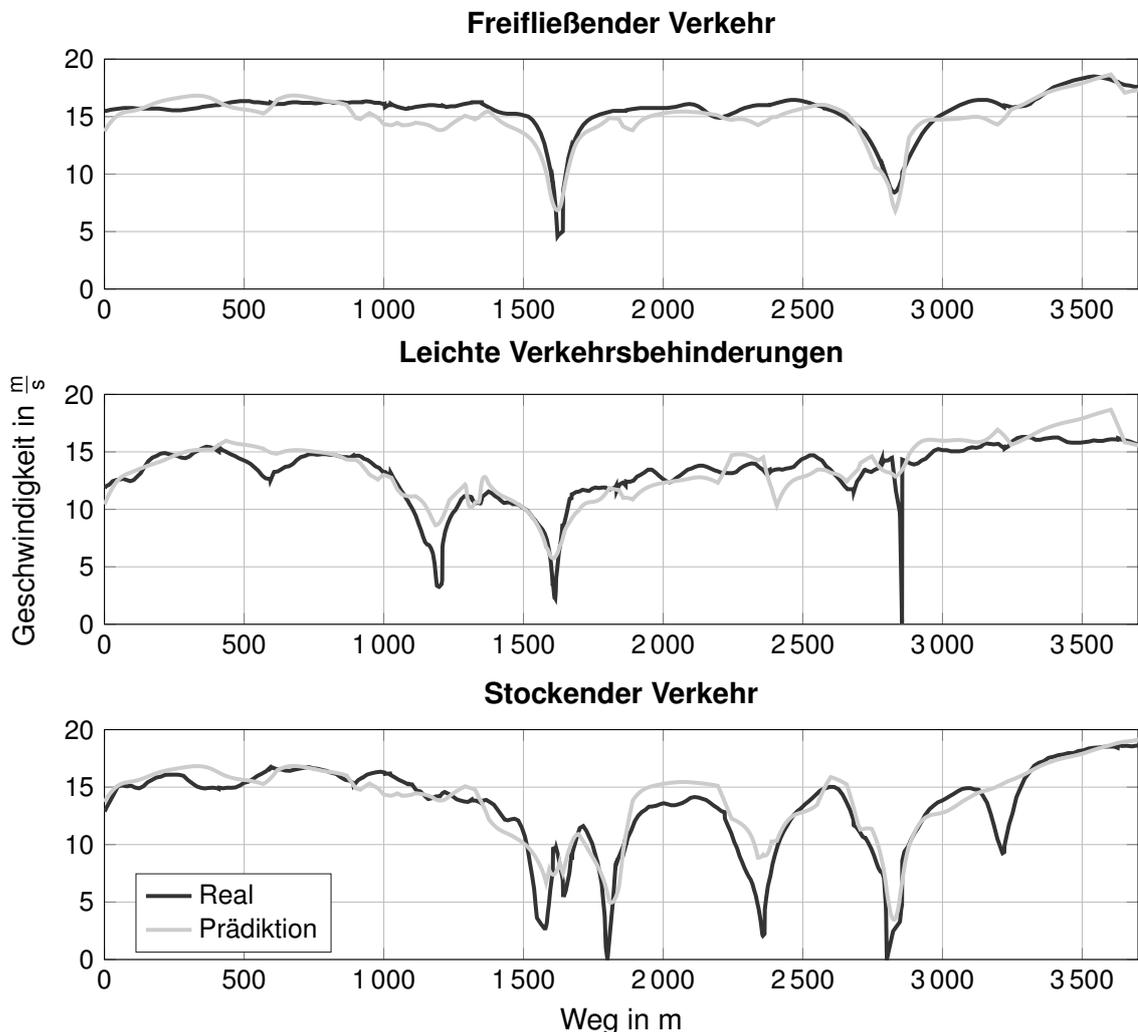


Abbildung 5.33 : Einfluss der Verkehrsbedingung auf das resultierende Fahrprofil

Aktualisierung

Für eine innerstädtische Fahrt wird das Verfahren mit einem kürzeren Prädiktionshorizont veranschaulicht. Abb. 5.34 zeigt die Fahrt mit einem Horizont von fünf Segmenten, die als Abschnitte definiert werden. Zur besseren Veranschaulichung wird jeweils nur die erste Prädiktion in dem Abschnitt dargestellt. Neben der Realfahrt werden zusätzlich die für diese Fahrt zugewiesenen Cluster illustriert. Bei der Prädiktion dienen die jeweils letzten vier zurückliegenden Segmente und der Verkehrsfluss als Basis. Die prädizierten Cluster zeigen größtenteils gute Übereinstimmungen mit der realen Fahrt. Der fortlaufende Horizont verdeutlicht, dass nach einer falschen Prädiktion das reale Fahrprofil dazu verwendet werden kann, um erneut eine Anpassung der Vorausschau zu ermöglichen. Dies ist zwischen Abschnitt zwei und drei zu erkennen. Im zweiten Abschnitt wird gegen Ende ein Haltevorgang prädiziert, der in der Realität nicht auftritt. Im dritten Abschnitt erfolgt erneut eine initiale Prädiktion mit der Information aus der Realfahrt. Dadurch können die folgenden Segmente besser vorhergesagt werden.

Gierrate

Anhand zweier Straßensegmente wird die Prädiktion des bevorstehenden Gierratenverlaufs veranschaulicht. Neben dem zugehörigen Geschwindigkeitsprofil werden jeweils drei Beispielfahrten, die dem Cluster zugehörig sind, dargestellt. Wie in Kapitel 5.2.4 werden zudem die

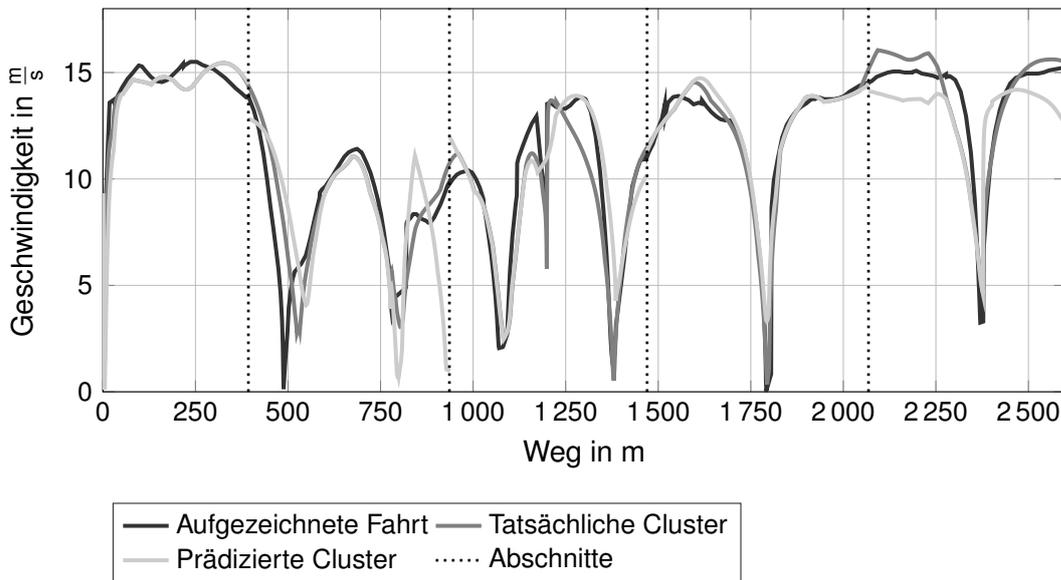


Abbildung 5.34 : Fortschreitender Horizont bei einer innerstädtischen Fahrt

entsprechenden Konfidenzintervalle angefügt. Die Fahrt links zeigt eine Auffahrt auf den mittleren Ring. Der gemessene Gierratenverlauf der Beispielfahrten liegt in allen Fällen innerhalb des 68 % Intervalls und wird damit ausreichend abgebildet. Die Fahrt rechts zeigt eine vergleichsweise schnell gefahrene S-Kurve am Flughafen. Die Beispiele zeigen alle einen ähnlichen Geschwindigkeitsverlauf. Das dargestellte Konfidenzintervall von 68 % wird in wenigen Fällen überschritten. In der ersten Fahrt ist dies einem Offset des vollständigen Verlaufs geschuldet, der allerdings innerhalb der Standardabweichung der Gierrate für die Geradeausfahrt liegt.

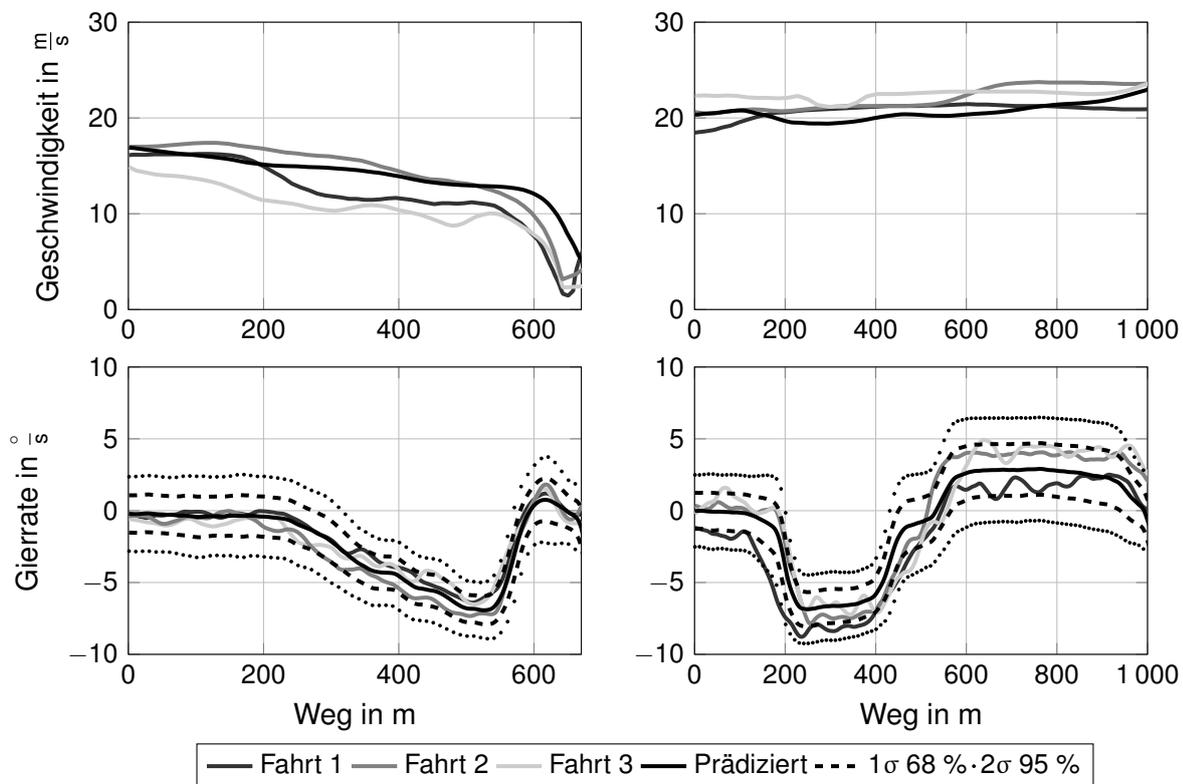


Abbildung 5.35 : Prädizierter Gierratenverlauf für zwei Streckenabschnitte

Energieverbrauch

Zuletzt wird der Energieverbrauch für eine Autobahnfahrt mit zwei unterschiedlichen Verkehrssituationen analysiert. Abb. 5.36 stellt die beiden Fahrten dar. Die Prädiktion erfolgt dabei zu Beginn der Fahrt über die vollständige Route über 37 km für alle Segmente. Die Eingangsparamete-

ter sind dabei der Verkehrsfluss und vier zurückliegende Streckensegmente. Als Vergleich dient eine Energieverbrauchssimulation des aufgezeichneten Fahrprofils. Der Geschwindigkeitsverlauf wird über die Fahrt in beiden Fällen gut angenähert. Die Berechnung des Energieverbrauchs erfolgt segmentweise und liefert eine gute Näherung des Verbrauchs gegenüber dem aufgezeichneten Fahrprofil. Bei der Analyse von 16 Fahrten auf dieser Route liegt der mittlere absolute prozentuale Fehler, der durch die Fahrprofilprädiktion verursacht wird, bei 3,3 %.

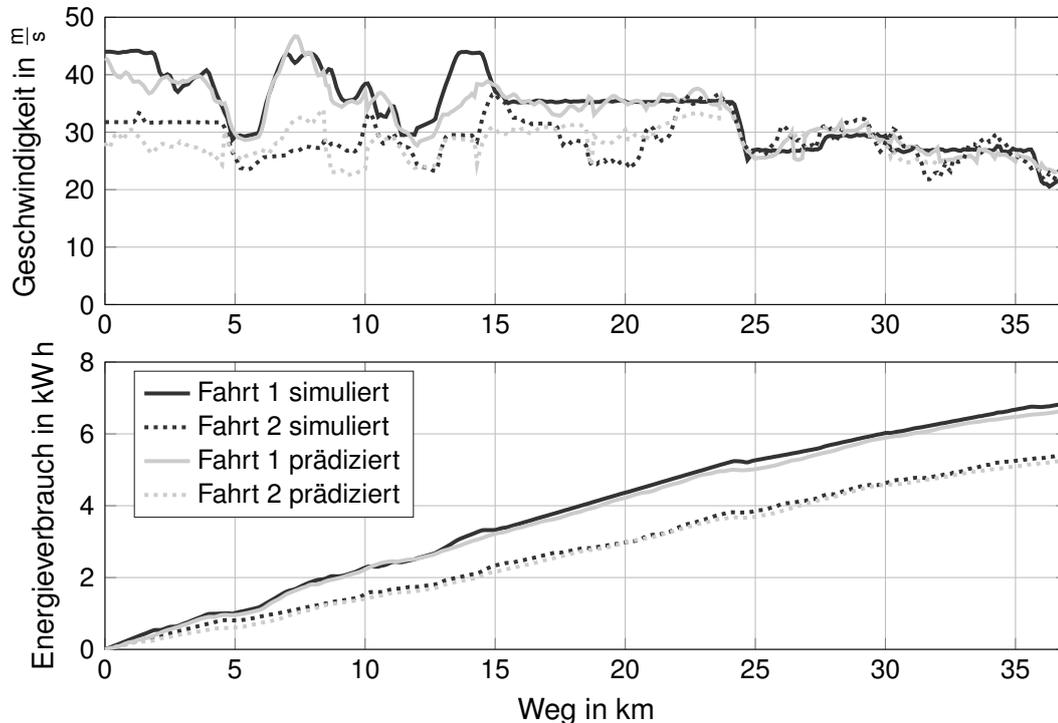


Abbildung 5.36 : Energieverbrauch prädiziert für zwei Autobahnfahrten

5.4 Bewertung

Es folgt eine Bewertung und Diskussion zur Fahrprofilprädiktion und den Beiträgen dieser Arbeit. Dabei wird zu Beginn auf den untersuchten Stand der Forschung und die Datenbasis eingegangen. Anschließend wird die angewandte Methodik diskutiert. Abschließend findet eine Bewertung der Ergebnisse statt.

Stand der Forschung

Die vorgestellten Quellen geben einen ausführlichen Überblick über die bisherige und aktuelle Forschung auf dem Gebiet. Eine inhaltliche Bewertung und Abgrenzung dieser Arbeit wurde in Kapitel 5.1.4 durchgeführt. Das breite Forschungsfeld unterteilt sich in mathematische und datengetriebene Ansätze. Ein wichtiger Ansatz für diese Arbeit ist der Beitrag von Kerper [226], der aufgegriffen und weiterentwickelt wird. Die Abgrenzung zum Stand der Forschung hat außerdem ergeben, dass im Bereich der datengetriebenen Ansätze fortführende Untersuchungen notwendig sind. Zudem werden weitere Bewegungsgrößen in dem Bereich kaum behandelt.

Datenbasis

Entscheidend für die Untersuchung ist der Umfang der Datenbasis. Im Vergleich zum Stand der Forschung handelt es sich um einen der größten Datensätze. Allerdings ist neben der Zahl der gefahrenen Kilometer die Verteilung auf den Untersuchungsraum ausschlaggebend. Die Häufigkeit, mit der zusammenhängende Streckenabschnitte befahren werden, hat einen direkten Einfluss auf den Trainingsdatensatz für ein Modell. Abb. 3.4 verdeutlicht dies. Besonders im Hauptverkehrsnetz sind eine Vielzahl an Fahrten vorhanden. In vielen weiteren Fällen allerdings ist die Ausgangslage begrenzt. Zudem reduziert sich der Datensatz für ein Segment

deutlich, wenn für die Modelle länger zusammenhängende Routenabschnitte benötigt werden. Signalabbrüche bspw. bei Fahrten durch Tunnel, Messausreißer oder längere Haltevorgänge von Taxis am Straßenrand reduzieren darüber hinaus die Datenbasis.

Neben der Größe des Datensatzes ist die Datenqualität von Bedeutung. Bei den aufgezeichneten Daten handelt es sich um Messungen mit Smartphones. Die Güte der Signale wurde in Kapitel 3.1.1 untersucht. Für die GPS-Position, die Geschwindigkeit, die Querbeschleunigung und die Gierrate werden die Fehlerabweichungen ermittelt.

Methodik

Ein Baustein der Methodik ist die Erstellung von charakteristischen Fahrprofilen. Hier wird das Verfahren aus dem Stand der Forschung von [226] adaptiert und für den Einsatz mit größeren Datenmengen und einer kartenbasierten Betrachtung erweitert. Über das FastDTW kann zudem ein Nachteil in der Datenqualität berücksichtigt werden. Auf Grund von unterschiedlichen Haltepositionen der Fahrzeuge und da die Genauigkeit des GPS der Smartphones nicht exakt ist, kommt es bei mehreren Fahrten zu einer Abweichung von der tatsächlichen Position. Das DTW-Verfahren erlaubt diese Verschiebung bei der Bewertung der Fahrten. Das spektrale Clusterverfahren kann darauf aufbauend die Unterteilung der Fahrten in verschiedene Klassen übernehmen. Die flexible Anzahl der Profile erlaubt eine vollständig maschinelle Bewertung, ohne vorgegebene Grenzwerte. Das Verfahren kann dadurch auf unterschiedliche Straßenklassen und Segmente angewendet werden.

Der Kern der Methode ist ein MLR zur Prädiktion des erwarteten Fahrprofilclusters. Als Eingangsdaten werden Daten unterschiedlicher Ausprägung verwendet. So können gleichzeitig die Cluster der vorherigen Segmente als auch der Verkehrsfluss oder fahrerspezifische Elemente angesetzt werden. Das Verfahren kann mit vergleichsweise wenig Datensätzen trainiert werden, was für manche Routenkombinationen erforderlich ist. Die Verwendung der Cluster des vorherigen Segments ermöglicht einen sequentiellen Ansatz bei der Prädiktion, bei dem die Ausgangsgrößen als Eingänge für die nächsten Modelle verwendet werden.

Das Verfahren wird ergänzt um die Prädiktion weiterer Bewegungsgrößen. Über die Einschränkung auf Straßensegmente kann mit den prädizierten Clustern die Vorhersage ausgeweitet werden. Die Methode ist durch die Datenqualität beschränkt. Zum einen ist die Aufzeichnung der Gierrate und der Beschleunigung Ungenauigkeiten unterworfen und durch die Halterung der Smartphones im Fahrzeug entsprechen die Werte nicht zwingend den Bewegungsgrößen des Fahrzeugs. Zum anderen ist die Positionsgenauigkeit unzureichend, um eine fahrspurspezifische Aussage zu treffen. Die Bewertung erfolgt daher für ein Segment mit mehreren Spuren in gleicher Weise. Um dennoch eine Aussage treffen zu können, werden Konfidenzintervalle im Bereich der Standardabweichung der Messungen für die Größen eingeführt. Um die Aussagekraft weiter zu steigern, ist eine intensivere Vorverarbeitung der vollständigen Datenbasis inklusive einer Sensordatenfusion erforderlich.

Neben den Bewegungsgrößen werden weitere Metadaten betrachtet. Insbesondere der Energieverbrauch kann für eine Vielzahl an Applikationen eingesetzt werden. Der Ansatz erlaubt eine segmentweise Vorabberechnung des Verbrauchs für Fahrzeuge. Dadurch kann der Energieverbrauch ohne eine vollständige Längsdynamiksimulation über den Prädiktionshorizont zusammengesetzt werden. Die Methode setzt sich nicht zum Ziel, den Verbrauch möglichst exakt zu bestimmen. Es zeigt allerdings die Anwendbarkeit des Verfahrens für weitere abhängige Größen.

Ergebnisse

Die repräsentativen Profile zeigen eine gute Übereinstimmung mit den Fahrten innerhalb des Clusters in Bezug auf die mittlere Geschwindigkeit, sowie die Fahrzeit. Das verwendete Ähnlichkeitsmaß unterteilt die Fahrten in sinnvolle Bereiche. Die Zahl der Cluster ist nach oben begrenzt, allerdings wird in den meisten Fällen nur eine geringere Anzahl von dem Algorithmus generiert. Dies resultiert in einem Fahrprofil, das eine Fülle an Fahrten versucht widerzuspiegeln. Dies reduziert zum einen Ausreißer, zum anderen beschränkt es auch das aufgezeichnete Fahrverhalten. Da die Zahl der Cluster begrenzt ist, bzw. in manchen Fällen auch keine klare Trennung

der Datensätze für das maschinelle Verfahren möglich ist, werden stärkere Beschleunigungs- oder Verzögerungsvorgänge mit den durchschnittlichen zusammengefasst. Ein aggressiveres oder defensiveres Fahrerverhalten ist dann zum Teil nicht mehr in den repräsentativen Profilen ersichtlich.

Dies hat auch einen Effekt auf die Prädiktionsmodelle. Die individuelle Beschleunigung im Vergleich zu der Gesamtheit zeigt kaum einen Einfluss auf die Güte des Modells. Eine Personalisierung auf Grund des Fahrerverhaltens ist nur begrenzt möglich. Den größten Effekt auf die Modellqualität hat der Verkehrsfluss. Dies ist nachvollziehbar, weil sich die durchschnittliche Geschwindigkeit direkt auf die Gestalt der Fahrprofile auswirkt. Die zurückliegenden Segmente wirken sich auf das Profil aus. Es sticht hervor, dass der entscheidende Einfluss nicht zwingenderweise in dem letzten zurückliegenden Segment generiert wird.

Bei der Klassifikation zeigt sich ein Einfluss der Straßenklasse. Die Trefferquote ist im innerstädtischen Bereich deutlich niedriger als auf Autobahnen. Dies ist insofern logisch, weil sich die Fahrt auf Autobahnen häufig allein anhand des Verkehrsflusses oder der erlaubten Höchstgeschwindigkeit einstellt. Es existiert zudem eine höhere Anzahl an möglichen Clustern im innerstädtischen Gebiet. Weiterhin ist die Chance auf ein unvorhergesehenes Verkehrseignis in der Innenstadt wesentlich größer. Es kann unter Umständen über die Betrachtung der zurückliegenden Segmente der Einfluss von Lichtsignalanlagen berücksichtigt werden, allerdings nur, wenn eine ausreichend große Datenmenge vorliegt. Die Klassifikation zeigt sich auch von der Anzahl der Cluster abhängig. Die Trefferquote steht im Gegensatz zu einer hohen Anzahl an Fahrten. Eine Einschränkung der Anzahl der Cluster ist jedoch nicht sinnvoll, weil somit die automatisierte Erstellung eingeschränkt wird und dies der Vereinbarkeit von charakteristischen Profil und Fahrprofil widerspricht.

Die betrachteten Größen zeigen eine gute Übereinstimmung mit den realen Fahrten. Das Verhalten von Beschleunigungs- und Verzögerungsvorgängen wird bei einer korrekten Prädiktion zielführend wiedergegeben. Der Verlauf der Gierrate und der Beschleunigungen ist von der Datenqualität abhängig und diese können daher nur über entsprechende Vertrauensintervalle ausgegeben werden. Der Energieverbrauch zeigt für längere Fahrten eine gute Übereinstimmung.

6 Anwendung in einer Betriebsstrategie

In diesem Kapitel wird untersucht, welcher quantifizierbare Nutzen sich aus der Prädiktion des Trips und des Fahrprofils für ein Fahrzeug ergibt, um die offene Forschungsfrage aus Kapitel 2.3.1 zu beantworten. Hierfür wird eine Betriebsstrategie auf Basis des Prädiktionsmodells für ein PHEV entwickelt. Ziel ist es, den Kraftstoffverbrauch zu reduzieren, ohne auf die Angabe eines Fahrtziels zurückgreifen zu müssen. Dafür wird ein datengetriebener Ansatz verwendet, der auf zusätzliche statistische Daten zurückgreift. Die Betriebsstrategie muss sich robust gegenüber Unsicherheiten aus der Prädiktion erweisen und fortlaufend aktualisiert werden. Es wird der gewählte Systemansatz in Anlehnung an Kapitel 2.3.2 aufgezeigt. Anschließend wird die Regelstrategie kurz erläutert. Eine detaillierte Beschreibung liefert [273]. Abschließend werden die gewonnenen Ergebnisse vorgestellt.

6.1 Systemansatz der Betriebsstrategie

Die grundsätzliche Architektur für LBS wird in Kapitel 2.2.2 bereits dargelegt. Aufbauend auf die vorgeschlagene Systemarchitektur nach Kapitel 2.2.2 wird diese für die Betriebsstrategie eines vernetzten PHEV erweitert (Abb. 6.1).

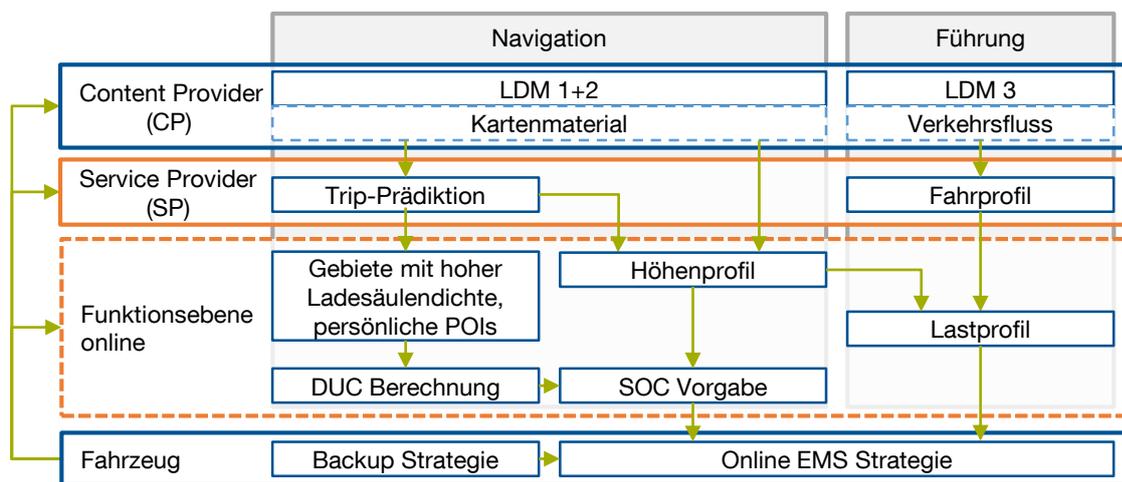


Abbildung 6.1 : Ansatz der Systemarchitektur für eine online basierte Betriebsstrategie

Die Herausforderung liegt in diesem Falle darin, dass das Fahrtziel nicht vorab bekannt ist und somit keine Aussage darüber getroffen werden kann, wie groß die Distanz ist, die das Fahrzeug zurücklegen wird. Um die Kraftstoffersparnis zu optimieren, ist es von Vorteil, den zur Verfügung stehenden State of Charge (SOC) möglichst vollständig auszunutzen. Jedoch ist die *Charge Depleting Methode*, zuerst den elektrischen Energiespeicher gänzlich zu leeren und anschließend hauptsächlich verbrennungsmotorisch zu fahren, für längere Fahrten nicht ideal [274], weil das Potential, den Verbrennungsmotor effizient einzusetzen, verloren geht. Es ist ein gemischter Betrieb aus *Charge Depleting* und *Charge Sustaining* optimal abzustimmen.

Der Service Provider liefert mit aktuellen Daten des Verkehrsflusses und den benötigten Echtzeitinformation aus dem Fahrzeug die Prädiktion für die vorausliegenden Streckenabschnitte und das Fahrprofil. Es wird daher ein Verfahren über einen gleitenden Horizont gewählt, wobei der Zustand des Fahrzeugs und insbesondere der SOC am Ende des Horizonts unbekannt ist. Die notwendige Information der Distance Until Charge (DUC) kann vereinfacht über statistische Daten bestimmt werden, wenn diese zu Beginn der Fahrt nicht über das Fahrtziel bekannt ist. Zudem kann die Wahrscheinlichkeit eines Ladevorgangs mit berücksichtigt werden. Dies kann bspw. über persönlich bevorzugte Ladestandorte, oder die Nähe zur öffentlichen Ladeinfrastruktur

tur und deren Verfügbarkeit, erfolgen. Über den Prädiktionshorizont kann das Wissen dadurch erweitert werden und somit bei längeren Fahrten der SOC besser eingeteilt, bzw. auch gezielter aufgebraucht werden.

Die spezifische Funktionsebene nach der Trip- und Fahrprofilprädiktion teilt sich in einen online- und einen fahrzeugbasierten Teil auf. Im ersteren wird aufbauend auf der Trip-Prädiktion die Zahl der Ladesäulen in der Region am Ende des Prädiktionshorizont bestimmt. Die Regionen sind nach den Sekundärgebieten in Kapitel 3.2.2 definiert. Mit einer statistischen DUC-Verteilung aus vergangenen Fahrten und dem Höhenprofil wird der SOC für die Fahrt kontinuierlich angepasst. Mit dem prädizierten Fahrprofil und der Steigung aus dem Höhenprofil kann ein Lastprofil für die vorausliegenden Abschnitte berechnet werden.

Die SOC-Vorgabe und das Lastprofil werden an das Fahrzeug übertragen. Dort erfolgt die Berechnung der *Energie Management Strategie*. Diese muss mit Unsicherheiten in der Prädiktion umgehen können. Für Fehler in der Übertragung oder Ausfällen wird darüber hinaus eine Rückfallebene im Fahrzeug benötigt. Der Ansatz ist nicht auf PHEV beschränkt, sondern kann auch für HEV eingesetzt werden. Ebenso könnten Optimierungsmaßnahmen für Elektrofahrzeuge basierend auf dem Lastprofil zur Reichweitenverlängerung entwickelt werden.

6.2 Regelstrategie

Das Energiemanagement von Hybridfahrzeugen kann grundsätzlich in regelbasierte Strategien oder Optimierungsmethoden unterteilt werden [275]. Erstere optimieren dabei Teilsysteme, benötigen wenig Rechenleistung und sind leicht zu implementieren. Daher eignen sich diese als Rahmenbetriebsstrategien, in diesem Falle als Back-up Strategie im Falle eines Systemausfalls. Bei den Optimierungsmethoden wird das Fahrzeug mathematisch modelliert. Darauf aufbauend werden Kostenfunktionen aufgestellt und die Kosten für die Fahrt minimiert. Eine Übersicht über mögliche Methoden wird in [275–277] gegeben.

6.2.1 Modell

Das System wird folgendermaßen definiert: Die Zustandsgröße x (6.1) wird durch den SOC des Energiespeichers beschrieben. Die Stellgröße u (6.2) entspricht dem Moment des Elektromotors. Die Störgröße z besteht aus dem Drehmoment $T_{\text{Tran,In}}$ und der Drehzahl $\omega_{\text{Tran,In}}$ am Getriebeeingang (6.3).

$$x(t) = \text{SOC}(t), \quad (6.1)$$

$$u(t) = T_{\text{EM}}(t) \quad \text{und} \quad (6.2)$$

$$z(t) = \begin{bmatrix} T_{\text{Tran,In}}(t) \\ \omega_{\text{Tran,In}}(t) \end{bmatrix}. \quad (6.3)$$

Die SOC-Rate bestimmt sich nach (6.4) aus der Stellgröße und den Störgrößen. Das zu Grunde liegende Batteriemodell wird ähnlich zu [278] basierend auf einem Ersatzschaltbild mit den Parametern zu Innenwiderstand, Leerlaufspannung und Kapazität, gehalten. Die Kostenfunktion (6.5) berücksichtigt ausschließlich den Kraftstoffverbrauch $g(u(t), z(t))$. Bei der Berechnung müssen die Beschränkungen für den Zustand (6.6) und die Stellgröße (6.7) eingehalten werden. Die Störgrößenbeschränkung (6.8) ergibt sich aus den Fahrzeugeigenschaften \mathcal{Z} .

$$\dot{x}(t) = f(u(t), z(t)), \quad (6.4)$$

$$J(u(t), z(t)) = \int_{t_0}^{t_e} g(u(t), z(t)) dt, \quad (6.5)$$

$$x(t) \in [\text{SOC}_{\min}, \text{SOC}_{\max}], \quad (6.6)$$

$$u(t) \in [u_{\min}(z(t)), u_{\max}(z(t))] \text{ und} \quad (6.7)$$

$$z(t) \in \mathcal{Z}. \quad (6.8)$$

Mit der notwendigen Definition der Initialbedingung $x(t_0) = x_0$ und der Endbedingung $x(t_e) = x_e$ handelt es sich dadurch um ein Zwei-Punkt-Randwert-Problem [245, S. 540–593]. Bei dem Kraftstoffverbrauch, sowie der SOC-Rate handelt es sich um nichtlineare Funktionen. Für eine detaillierte Übersicht zur Modellierung wird auf [279] verwiesen.

6.2.2 Optimierung

Die dynamische Programmierung DP wird in vielen Fällen als Referenz über eine bekannte Fahrt herangezogen, weil diese eine optimale Lösung bestimmen kann [280]. In der ursprünglichen Form ist DP rechenintensiv und muss für den Einsatz mit einer ungenauen Prädiktion der Fahrt erweitert werden. Eine weit verbreitete Variante ist zudem das Pontryagin's Minimum Prinzip (PMP) [143] basierend auf der Variationsrechnung [281]. Hierbei wird die Hamilton-Gleichung H (6.9) minimiert. Dafür wird der Lagrange-Multiplikator λ eingeführt [281].

$$\min_u H(u(t), z(t)) = g(u(t), z(t)) + \lambda(t)f(u(t), z(t)) \quad (6.9)$$

Zur Minimierung der Gleichung erfolgt die Ermittlung von λ mit einem Schießverfahren nach [273]. Durch diese Methode ist keine Rückwärtsberechnung wie bei DP erforderlich, was die Rechenzeit deutlich reduziert. Zudem kann λ adaptiv für jeden Zeitschritt für den Prädiktionshorizont und dem wahrscheinlichen Lastprofil bestimmt werden. In [273] wird der Einfluss des Prädiktionshorizonts untersucht.

Die Variante ist daher abhängig von dem SOC im betrachteten Horizont. Dieser kann, wie in Kapitel 6.1 beschrieben, für die Fahrt angepasst werden. Die Daten für die initiale DUC-Verteilung stammen aus dem sun2car-Datensatz (Kapitel 3.1.1). Dieser beinhaltet für die Fahrten der Probandengruppe mit zehn Fahrzeugen, die genaue Auflistung der Ladevorgänge, dem Rest-SOC und der dazwischen gefahrenen Distanz. Der Algorithmus zur adaptiven Vorgabe des SOC nach der Zahl der Ladesäulen ist in [273] detailliert beschrieben. Der SOC orientiert sich dabei an der DUC-Verteilung. Dabei wird die Steigung der Verteilung über die Distanz angepasst. Der Prädiktionshorizont wird mit einer Anzahl an Straßensegmenten festgelegt. Liegt das Ende des Prädiktionshorizonts in einem Gebiet von einer hoher Ladesäulendichte, wird die SOC-Vorgabe stärker als die DUC-Verteilung abgesenkt. Dabei wird die Länge des Horizonts mitberücksichtigt. Liegt das Ende des Horizonts in einem Gebiet von geringer Ladesäulendichte ist ein Ladevorgang weniger wahrscheinlich. Die Steigung der SOC-Vorgabe wird vermindert, um eine größere Energiemenge vorzuhalten. Dies wird exemplarisch in Kapitel 6.3.2 veranschaulicht.

Zum Vergleich wird neben der Referenz mit DP eine Strategie ohne Vorausschau gewählt. Die Equivalent Consumption Minimization Strategy (ECMS) minimiert einen äquivalenten Kraftstoffverbrauch, der sich aus der Kraftstoffverbrauchsrate und der SOC-Rate ergibt. Anstelle des Schießverfahrens wird ein PI-Regler nach [282, 283] eingesetzt.

6.3 Ergebnisse

Folgend wird eine Einzelfahrt genauer beleuchtet. Hierfür wird der durch den Algorithmus resultierende SOC näher betrachtet und die Kraftstoffeinsparungen für diese Fahrt berechnet. Anschließend werden die zu erwartenden Kraftstoffeinsparungen für mehrere Fahrten bewertet.

6.3.1 Beispielfahrt

Die Fahrt ist in Abb. 6.2 dargestellt. Sie startet in München und endet in Garmisch-Partenkirchen. Für die Fahrt werden die nächsten zehn Streckensegmente prädiziert. Es werden keine personalisierten Informationen verwendet. Die Prädiktion wird nach der Methodik in Kapitel 4 durchgeführt. Zu Beginn trifft der Algorithmus innerorts an wenigen Stellen nicht den korrekten Weg. Außerdem liegt die Prädiktion kurz vor der Ausfahrt nach Starnberg falsch. Im übrigen Verlauf wird die Fahrt korrekt prädiziert. In der Karte sind die Gebiete mit der Zahl der Ladestationen farblich markiert. Die Fahrt führt durch Gebiete mit einer erhöhten Zahl an Ladestationen. Bei der anschließenden Überlandfahrt ist kaum eine öffentliche Lademöglichkeit vorhanden. In Garmisch-Partenkirchen ist auf Grund der hohen Zahl an Ladestationen davon auszugehen, dass der Fahrer das Fahrzeug nachladen kann.

Des Weiteren ist der Geschwindigkeitsverlauf für die reale Fahrt, sowie die prädizierte Geschwindigkeit nach der Methodik in Kapitel 5 dargestellt. Das reale Verhalten der Beschleunigungs- und Verzögerungsphasen wird mit der Prädiktion angenähert. An manchen Stellen (Positionen: 16 km, 28 km) weicht der Absolutwert der Prädiktion von der Realfahrt ab. Zudem wird die Steigung über das Höhenprofil nach Kapitel 3.1.2 bestimmt. Es zeigt sich ein moderater Verlauf der Steigung für die gewählte Route.

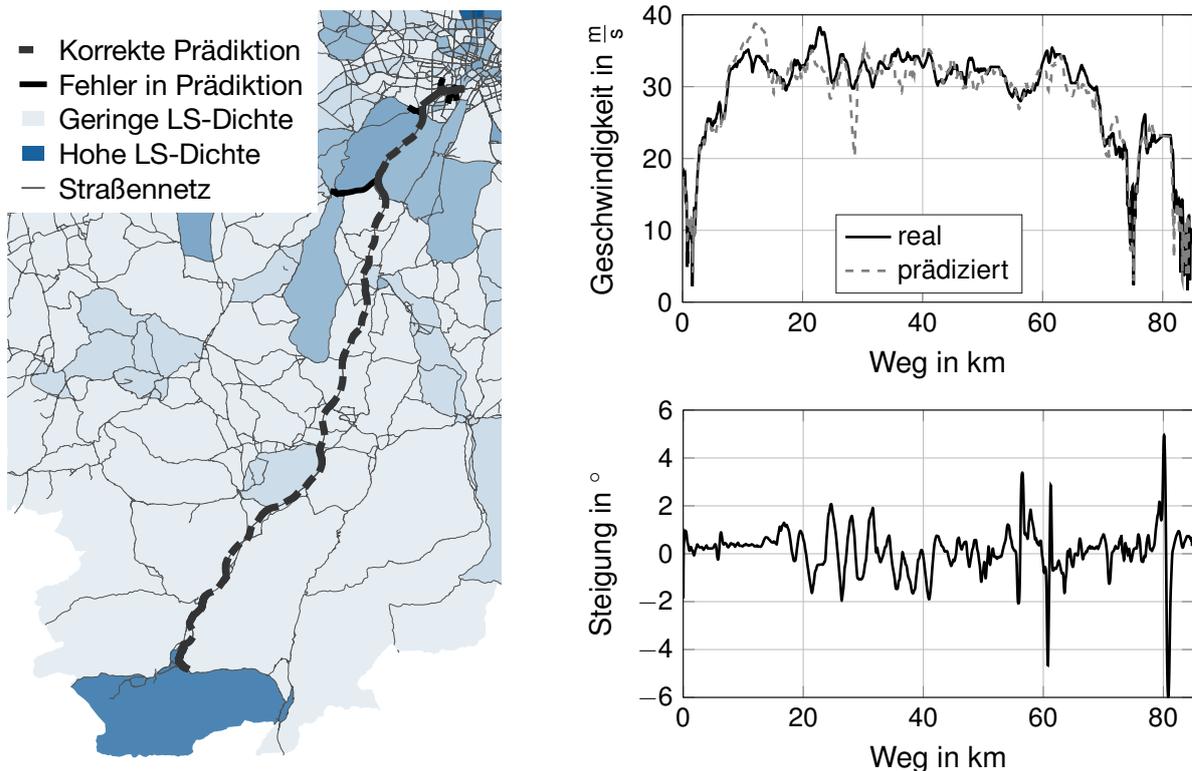


Abbildung 6.2 : Beispielfahrt über eine längere Distanz nach [273] (Karte: OSM © ⓘ ⓘ)

6.3.2 SOC-Verlauf

Das Ergebnis der Regelstrategie für das Fahrzeug wird anhand des SOC-Verlaufs vorgestellt. In Abb. 6.3 werden drei verschiedene SOC-Vorgaben dargestellt. Alle Varianten enden oberhalb der vorgegebenen Grenze von 80 %. Zum einen wird ein linearer Verlauf des SOC über der

Distanz dargestellt. Diese Vorgabe ist nach [284, 285] bevorzugt anzusetzen, wenn die Distanz der Fahrt bekannt ist. Des weiteren ist der SOC-Verlauf über die DUC-Verteilung dargestellt. Dieser kann wie in 6.1 vorgestellt als Alternative angenommen werden, wenn das Ziel der Fahrt nicht bekannt ist. Zuletzt ist der SOC-Verlauf für den adaptiven Ansatz basierend auf der Trip-Prädiktion nach [273] dargestellt. Wie in Kapitel 6.1 beschrieben, wird die SOC-Vorgabe über die Verfügbarkeit an Ladesäulen über den Prädiktionshorizont angepasst. In Abb. 6.3 wird zudem die Zahl der Ladesäulen in den angrenzenden Gebieten, die in der Trip-Prädiktion auf der Strecke voraus liegen, dargestellt. Über die gestrichelten Linien wird der Prädiktionshorizont in Fällen veranschaulicht, bei denen eine Änderung der Zahl der Ladesäulen vorliegt. Die Fehlerlinien verdeutlichen die Neukalkulation in der Trip-Prädiktion, für den Fall, dass der Fahrer von der prädizierten Route abweicht. Dies ist zu Beginn und an der Ausfahrt zu Starnberg der Fall.

Es werden Gebiete mit vergleichsweise wenig Ladesäulen prädiziert. Ab ca. 10 km setzt die Trip-Prädiktion die nächsten Streckensegmente in Richtung Starnberg. Dort existiert eine hohe Zahl an Ladestationen und der Algorithmus reduziert die SOC-Vorgabe. Bei ca. 18 km wählt der Fahrer eine andere Route. Die Prädiktion wird für die Fahrt neu ausgeführt. Die nun wahrscheinlichere Route führt über ländliches Gebiet mit geringer Ladesäulendichte. Daraufhin wird die SOC-Vorgabe wieder angehoben, um an die DUC-Verteilung aufzuschließen. Bis Garmisch-Partenkirchen ist keine Lademöglichkeit in Sicht. Die SOC-Vorgabe wird über die DUC-Verteilung gesetzt, um über der SOC-Grenze zu liegen und flexibler mit den restlichen Energieinhalt umgehen zu können. Die Fahrt endet in Garmisch-Partenkirchen und der SOC wird auf das Minimum reduziert.

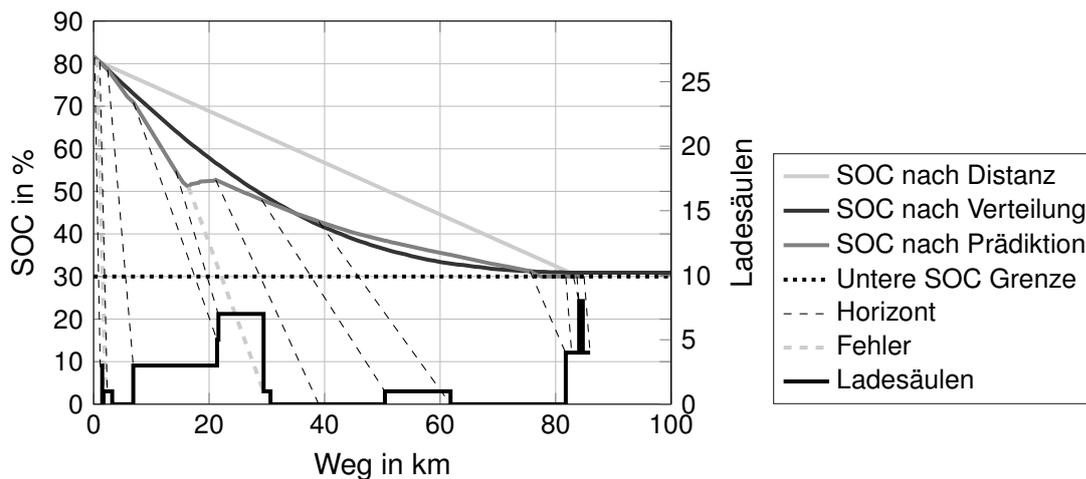


Abbildung 6.3 : Einfluss der Ladesäulen über Prädiktionshorizont auf die SOC Vorgabe

Der resultierende SOC-Verlauf wird in Abb. 6.4 dargestellt. Es werden vier Varianten miteinander verglichen. In der Legende wird die relative Kraftstoffersparnis prozentual angegeben. Die Rahmenbetriebsstrategie dient als Referenz. Bei dieser wird mit *Charge Depleting* der SOC bis zur unteren Grenze reduziert und anschließend dieser Wert gehalten. Vergleichend ist der Verlauf für DP dargestellt. Dieser weicht deutlich von der Rahmenbetriebsstrategie ab. Die Strategie ist unabhängig von einer SOC-Vorgabe und zeigt, dass eine gemischte Strategie von *Charge Depleting* und *Charge Sustaining* deutliche Vorteile bietet. Der resultierende Verlauf ist verglichen mit Abb. 6.3 der linearen Vorgabe über der bekannten Distanz am nächsten. DP zeigt die maximale Kraftstoffersparnis. Zudem ist die ECMS-Strategie dargestellt. Diese setzt den SOC nach der DUC-Verteilung ohne eine Vorausschau um. Der resultierende Verlauf weicht kaum von der Vorgabe ab.

Die vorgestellte aPMP Variante agiert nach der SOC-Vorgabe, die durch die Trip-Prädiktion angepasst wird und die Informationen aus der Fahrprofilprädiktion enthält. Verglichen mit ECMS zeigt sich die unterschiedliche Vorgabe in den resultierenden Verläufen. Weiterhin setzt aPMP den Energieinhalt flexibler ein. ECMS und aPMP liegen nahe an der maximalen Kraftstoffer-

sparsnis von DP und zeigen in diesem Falle kaum Unterschiede. Die Kraftstoffersparnis der vorgestellten Strategie liegt trotz einer teilweise falsch prädizierten Route leicht über der Variante ohne Prädiktion. Das Potential der Vorausschau zeigt sich vor allem auf kürzeren Fahrten, weil der SOC besser ausgenutzt werden kann. Die aPMP-Strategie ist nach [273] grundsätzlich in der Lage auch nahe an der oberen oder unteren SOC-Grenze zu agieren. Durch die Vorausschau wird das voraus liegende Höhenprofil berücksichtigt, wodurch Rekuperationsphasen strategisch eingeplant werden können. Bei Fahrtbeginn mit einer darauffolgenden Rekuperationsphase wird aPMP den Elektromotor bevorzugt verwenden. ECMS wird sich in diesem Falle strikter an die SOC-Vorgabe halten.

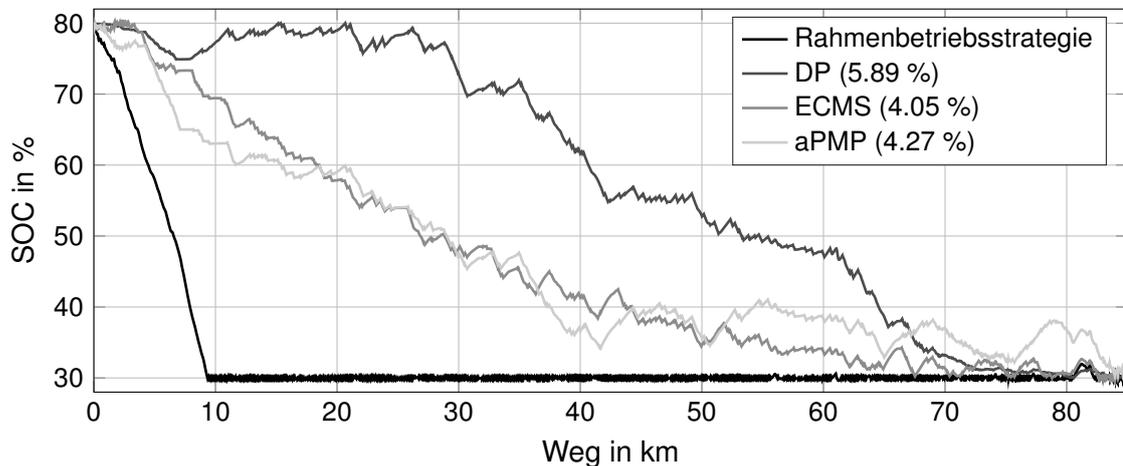


Abbildung 6.4 : Resultierender SOC Verlauf für verschiedene Strategien nach [273]

6.3.3 Erwartete Kraftstoffeinsparungen

Um den Einfluss der Prädiktion für mehrere Fahrten mit unterschiedlicher Länge zu bewerten, werden 214 Fahrten untersucht. Dabei werden die Varianten ECMS und aPMP mit Trip- und Fahrprofilprädiktion miteinander verglichen. In Abb. 6.5 ist die relative Kraftstoffersparnis von aPMP gegenüber ECMS über die Fahrdistanz in Boxplots dargestellt. Für längere Fahrten von mehr als 30 km zeigt sich ein genereller, leichter Vorteil der Strategie mit dem Prädiktionshorizont. Das große Potential wird bei kürzeren Fahrten deutlich, weil der SOC in diesen Fällen besser ausgenutzt wird. Allerdings kann in seltenen Fällen auch ECMS bessere Ergebnisse liefern. Um eine repräsentative Kraftstoffersparnis zu bestimmen, wird die Häufigkeit der Fahrten mit der jeweiligen Distanz mit der DUC-Verteilung berücksichtigt. Es resultiert ein Vorteil von ca. 8,0 % gegenüber der Prädiktion ohne Vorausschau mit ECMS.

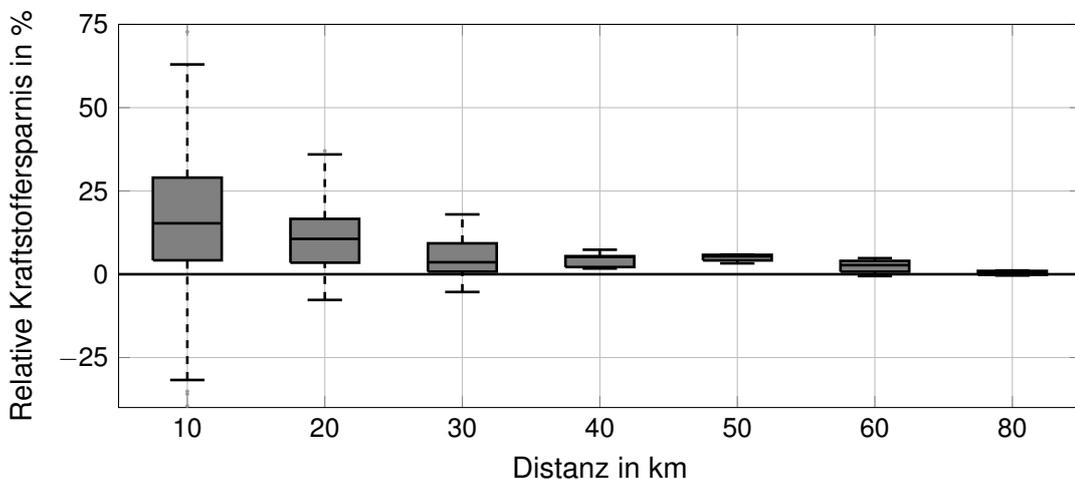


Abbildung 6.5 : Ergebnisse für mehrere Fahrten unterschiedlicher Länge nach [273]

7 Diskussion

In diesem Abschnitt werden die Beiträge dieser Arbeit diskutiert. Es werden die gesetzten Ziele hinterfragt. Anschließend wird untersucht, ob die Forschungsfragen im Rahmen dieser Arbeit ausreichend beantwortet wurden. Zuletzt werden die Einschränkungen und die Übertragbarkeit der Ergebnisse beleuchtet.

7.1 Zielsetzung

Das übergeordnete Ziel dieser Arbeit ist es, einen Mehrwert aus aufgezeichneten Fahrdaten zu erzeugen. Als Funktion soll dieser weiteren Fahrzeugen wiederum zur Verfügung gestellt werden. Zahlreiche Studien, die in Kapitel 1 aufgeführt werden, bestätigen diesen Bedarf für fahrzeugtechnische Innovationen in Bezug auf datengetriebene Dienste, weil hier ein enormer Wachstumsmarkt entsteht. Diese Art der Fragestellung ist für die Automobilindustrie derzeit von hoher Relevanz.

Diese Arbeit widmet sich der Entwicklung von konkreten Funktionen. Dies ist zum aktuellen Zeitpunkt möglich, weil die Grundlage für standortbezogene Dienste bereits geschaffen wurde. In Kapitel 2.1 werden die dafür erforderlichen fahrzeuginternen Technologien beleuchtet. Zudem existiert eine Vielzahl an Standards und Protokollen für die Übertragung der Daten. Geeignete Back-End-Architekturen wurden in mehreren Konsortien bereits untersucht. Die Forschung kann sich nun der Entwicklung von geeigneten datengetriebenen Funktionen widmen.

Die Funktionen, die in dieser Arbeit untersucht werden, haben zum Ziel, die vorausliegenden Fahrtabschnitte (Trip) und das Fahrprofil über einen begrenzten Horizont zu präzisieren. Die anknüpfenden Kundenfunktionen sind vielseitig. Im Bereich der Energieeffizienz ist das Geschwindigkeitsprofil die Haupteinflussgröße. Aufbauend auf der Prädiktion des Bewegungszustandes können Verbrauchseinsparungen erzielt werden, was in Kapitel 6 gezeigt wird. Im Bereich der Fahrsicherheit und des Komforts sind neben der Geschwindigkeit weitere Bewegungsgrößen von Bedeutung. Diese werden in der Funktion integriert. Die Route, die ausschlaggebend für das Profil aller Bewegungsgrößen ist, muss daher jederzeit bekannt sein. In der Motivation in Kapitel 1.1 wird herausgearbeitet, dass dies nicht garantiert werden kann. Somit ist dies ein wichtiger Bestandteil für die Erfüllung der Funktionalität. Die Fragestellung ist daher relevant für Hersteller und Kunden und kann einen Mehrwert aus den Mobilitätsdaten generieren.

7.2 Forschungsfragen

Im folgenden Abschnitt werden die gesetzten Forschungsfragen näher untersucht und diskutiert, ob diese geeignet beantwortet wurden.

Wie kann die geplante Fahrtroute des Fahrzeugs ohne Angabe eines Fahrtziels präzisiert werden?

Diese Frage wurde im Rahmen der Trip-Prädiktion beantwortet. In Kapitel 4 wurde die Methodik erläutert und in Abschnitt 4.4 bereits entsprechend kritisch bewertet. Die umgesetzte Methode ist in der Lage, die nächsten Straßensegmente vorherzusagen. Dabei sind für die Grundfunktionalität weder das Fahrtziel, noch persönliche Informationen notwendig. Ausschlaggebend für die Bewertung ist der Prädiktionshorizont. Dieser kann zwar von variabler Länge sein, doch sinkt die Zuverlässigkeit der Vorhersage über die zusammenhängende Fahrstrecke. Dies muss bei der Gestaltung darauf aufbauender Kundenfunktionen ohne persönliche Informationen berücksichtigt werden. Eine über die komplette Fahrt korrekte Prädiktion ist mit der Grundfunktionalität unwahrscheinlich. Da die Daten für das Trainingsverfahren hauptsächlich von Taxis stammen,

werden zwar öfter kombinierte Routenabschnitte erfolgreich erkannt, aber an entscheidenden Kreuzungspunkten kann keine zuverlässige Aussage getroffen werden. Die Kundenfunktionen müssen mit derartigen Unsicherheiten umgehen können.

Wie kann das Fahrprofil des Fahrzeugs möglichst detailliert für einen mittelfristigen Horizont präzisiert werden?

In Kapitel 5 wurde die Prädiktion des Fahrprofils untersucht und die Ergebnisse anschließend in Abschnitt 5.4 diskutiert. Das Geschwindigkeitsprofil ist die entscheidende Größe für die Bewegung des Fahrzeugs. Der Stand der Forschung hat gezeigt, dass hierbei zwischen einer kurz- und einer längerfristigen Prädiktion unterschieden werden muss. Mit der Methode der Prädiktion von repräsentativen Profilen wird auf einen längeren Horizont gesetzt. Durch das Abstrahieren der Zeitreihen auf standortbezogene wiederkehrende Muster werden Verallgemeinerungen getroffen, welche die Genauigkeit gegenüber kurzfristigen Vorhersagen mindern. Die genaue Lokalisierung von bspw. Haltepunkten oder die Wahl der Geschwindigkeit auf Grund des Fahrzeugfolgeverhaltens kann mit diesem Ansatz nicht abgedeckt werden. Unter diesen Einschränkungen kann allerdings eine Prädiktion für einen mittelfristigen Horizont erzeugt werden. Da die Untersuchungen gezeigt haben, dass diese Muster existieren und abhängig von dem Verkehrsfluss und der Wahl der Route auftreten, wurde versucht dies vorherzusagen. Diese repräsentativen Profile enthalten die erforderlichen Details, um bspw. den Energieverbrauch oder die Fahrzeit aus der Prädiktion abzuleiten. Neben der Geschwindigkeit wurde die Querbeschleunigung und die Gierrate mit untersucht. Eine Vorhersage kann hier getroffen werden, auch wenn die Qualität auf Grund der verwendeten Smartphone-Sensordaten deutlich eingeschränkt ist. Diese liefern aber eine gültige Tendenz für den Bewegungszustand des Fahrzeugs.

Wie kann diese Funktion für unterschiedliche Ziele zur Optimierung von Fahrzeugfunktionen eingesetzt werden?

Die entwickelten Funktionen dienen als Basis für weitere Kundenfunktionen, die darauf aufsetzen. Da die ermittelten Größen für viele Punkte benötigt werden, ist ein breites Spektrum denkbar. Mögliche Kategorien wurden in Kapitel 1.1 bereits erläutert. In dieser Arbeit wurde eine Anwendung zur Steigerung der Effizienz untersucht. In Kapitel 6 wurde dabei die Betriebsstrategie für ein Hybridfahrzeug betrachtet. Dies ist für die Prädiktion der Geschwindigkeit ein häufiges Anwendungsbeispiel und wurde daher gezielt ausgewählt. Es wurde gezeigt, dass mit dem vorgestellten Ansatz eine Kraftstoffeinsparung möglich ist.

Welcher quantifizierbare Nutzen ergibt sich aus einer derartigen Anwendung?

Der Nutzen wurde speziell für den Einsatz eines PHEV untersucht. Der Ansatz wurde um die Berücksichtigung der Ladeinfrastruktur und des Ladeverhaltens des Kunden erweitert; eine Kundenfunktion, die auf der entwickelten Trip- und Fahrprofilprädiktion aufsetzt. Die Strategie wurde an einer Vielzahl realer Fahrprofile untersucht. In Abschnitt 6.3.3 wurden die berechneten Kraftstoffeinsparungen vorgestellt. Da es sich weiterhin um Simulationen handelt und auch keine expliziten Aufzeichnungen über das tatsächliche Ladeverhalten der verwendeten Beispiele vorliegen, sind dies Erwartungswerte, die in Realversuchen bestätigt werden müssen. Da der Vorteil gegenüber einem System ohne Vorausschau bei ca. 8 % liegt, wird trotz Vereinfachungen in der Modellierung ein Effizienzgewinn erwartet.

Wie können diese Ziele unter der Verwendung einer Vielzahl gesammelter Fahrprofile erreicht werden?

Die gewählten Methoden sind rein datenbasiert. In Kapitel 3 wurden die verwendete Datenbasis und die notwendigen Vorverarbeitungsschritte aufgezeigt. Der Datensatz ist im Vergleich zu den meisten anderen Arbeiten aus dem Stand der Forschung deutlich größer. Da die gefahrenen Kilometer allerdings über den Untersuchungsraum Oberbayern verteilt sind, ist die Datenlage nicht immer ausreichend. Bei der Prädiktion muss stets bewertet werden, ob für das Trainingsverfahren ausreichend Datensätze vorhanden sind, um eine aussagekräftiges Ergebnis zu erhalten. Neben der Größe der Datenbasis ist auch die Qualität entscheidend. Die Vielzahl der aufgezeichneten Fahrten bietet die Möglichkeit, reale Muster in den Datensätzen zu erkennen. Allerdings ist durch die Erfassung der Daten mit Smartphones die Genauigkeit im Vergleich zur Sensorik im Fahrzeug eingeschränkt. Die Ausrichtung der Smartphones muss überprüft und die

erfolgreiche Übertragung und Aufzeichnung der Daten kann nicht gewährleistet werden. Eine umfangreiche Aufbereitung der Sensordaten über die komplette Datenbasis ist für eine höhere Genauigkeit erforderlich. Die gewählten datenbasierten Verfahren können iterativ angepasst werden. Die Erweiterbarkeit für andere Regionen ist durch die kartenbasierte Betrachtung gegeben. Die Funktionen nutzen ausschließlich ortsspezifisch aufgezeichnete Daten. Dadurch werden die Prädiktionsmodelle an die Umgebungsbedingungen angepasst.

Inwiefern können diese Ziele anonymisiert bzw. personalisiert erreicht werden?

In dieser Arbeit wurden zum Großteil Fahrten von Taxis verwendet. Sie liefern die Basis für die Trip-Prädiktion und die Fahrprofilprädiktion. Für das Training der Modelle ist nur ein Ausschnitt der Fahrt ohne Angabe des Fahrtziels oder des Startpunkts erforderlich. Für die Vorhersage benötigen die Modelle die Information über die aktuelle Position und einen Ausschnitt aus den Fahrdaten für eine begrenzte Anzahl zurückliegender Segmente. Werden diese unabhängig von dem Fahrzeug im Back-End bearbeitet, ist eine anonyme Bearbeitung möglich. Für die Prädiktion der nächsten Streckensegmente wurde die Funktion in Kapitel 4.2.4 um eine personalisierte Variante erweitert. Hier wurde ein deutlicher Vorteil gegenüber der Grundfunktionalität festgestellt. Die Zahl der Fehlprädiktionen wurde um über 17 % gesenkt. Nicht nur Einzelpersonen, sondern auch Nutzergruppen können von der Verwendung von personalisierten Daten profitieren. Die Entscheidung, ob er davon Gebrauch machen will, bleibt dem Kunden überlassen. Bei der Fahrprofilprädiktion wurde in Abschnitt 5.4 erörtert, dass eine fahrerspezifische Vorhersage kaum möglich ist. Eine Anpassung des Verfahrens auf nutzerspezifische repräsentative Profile ist auf Grund der unzureichenden Datenmenge nicht umsetzbar.

7.3 Einschränkungen und Übertragbarkeit

Die erstellten Modelle sind nicht uneingeschränkt für den Untersuchungsraum anwendbar. Dies liegt wie bereits beschrieben an der Datenbasis. Besonders in Wohngebieten oder Vororten sind Teile des Straßennetzes nur selten befahren worden, was die Aussagekraft des Modells stark einschränkt. Es ist notwendig, eine flächendeckend ausreichend große Menge an Fahrten zu analysieren, um für alle Gebiete die Modelle zu erzeugen. Die Methode ist auf Grund der straßensegment- und gebietweisen Betrachtung einfach auf weitere Regionen übertragbar. Das verwendete Kartenmaterial, bzw. die Quelle für den Verkehrsfluss können ausgetauscht werden. Die Modelle müssen entsprechend neu angelern werden. Die Aussagekraft hat, bezogen auf die Datenerfassung mit den Smartphones, die beschriebenen Einschränkungen hinsichtlich der Genauigkeit. Eine vergleichbare Datensammlung mit der Sensorik eines Fahrzeugs kann eine deutliche Verbesserung schaffen. Die Methodik muss diesbezüglich nicht verändert werden.

In dieser Arbeit wurde die Integration einer solchen Funktion nicht genauer untersucht. Die Einbindung in eine Back-End Architektur ähnlich des NGTP und die Integration der übertragenen Daten in das ADASIS Protokoll ist allerdings möglich. Die Algorithmen zur Prädiktion sind auf Grund einfacher mathematischer Operationen und der segmentweisen Betrachtung auch auf größere Regionen skalierbar.

8 Zusammenfassung und Ausblick

Die Vernetzung der Fahrzeuge ist maßgeblich für die zukünftige Entwicklung in der Automobilindustrie. Das Fahrzeug als Sensor sammelt eine Fülle an Daten. Als Herausforderung gilt, einen Mehrwert aus diesen Daten zu erzeugen. So ist es das Ziel, fahrzeugtechnische Innovationen zu generieren, bei denen der Kunde den Vorteil in der Verwendung seiner Daten erkennt. Derartige Funktionen liegen in den Bereichen Energieeffizienz, Sicherheit und Komfort. Die Grundlage für viele dieser neuartigen Dienste bilden Mobilitätsdaten. Je nach Position werden unterschiedliche Informationen erhoben bzw. dem Fahrzeug mitgeteilt. Dabei ist die Prädiktion der Umgebungsbedingungen oder des Fahrzustands ein Mehrwert, der über historische Daten geschaffen werden kann. Als entscheidende Information für viele Funktionen gilt das Geschwindigkeitsprofil. Es ist die Haupteinflussgröße für den Verbrauch und wichtig für Sicherheitsfunktionen während der Fahrt. Es ist überwiegend abhängig von der Position, der gewählten Route und dem Verkehr. Um eine Vorhersage dieser Größe zu jedem Zeitpunkt zu ermöglichen, ist darüber hinaus eine Prädiktion des nächsten Routenabschnitts notwendig. Daher wurden in dieser Arbeit zwei Funktionen entwickelt, die Prädiktion des Trips sowie die Prädiktion des Fahrprofils entlang der erwarteten Route, auf deren Basis weitere prädiktive, standortbezogene Kundenfunktionen aufbauen können.

Um die erforderlichen Schritte einordnen zu können, wurden die Grundlagen der Vernetzung von Fahrzeugen untersucht. Dabei wurden die notwendigen fahrzeuginternen Technologien, die Übertragungsmechanismen und mögliche Back-End Architekturen evaluiert. Es wurde festgehalten, dass für den Bereich der standortbezogenen Dienste bereits die erforderliche Basis existiert und die funktionale Seite betrachtet werden kann.

Im Rahmen der Entwicklung wurden knapp 5 Mio. km realer Fahrprofile von über 200 verschiedenen Fahrern verwendet. Diese wurden in mehreren Flottenversuchen in München und Oberbayern mit Hilfe von Smartphones aufgezeichnet. Neben der Position und der Geschwindigkeit wurde auch die Gierrate, sowie die Querschleunigung des Fahrzeugs evaluiert. Über mehrere Vorverarbeitungsschritte wurden die Daten um zusätzliche Kennwerte und vor allem Karteninformationen ergänzt.

Zur Prädiktion der nächsten Routenabschnitte wurde ein neuronales Netz, im speziellen ein *End-to-End Memory Network Model*, eingesetzt. Die Fahrprofile wurden dabei auf die befahrenen Straßensegmente reduziert und die Abfolge der Segmente mit dem Verfahren des maschinellen Lernens trainiert. Neben dem allgemeinen Ansatz, der hauptsächlich auf den Fahrten mehrerer Taxis basiert, wurde die Methode um eine nutzerindividuelle Variante erweitert. Hierbei wird das Netz bei fehlerhafter Prädiktion weiter trainiert und somit für den Einsatzzweck angepasst. Die Methode wurde mit über 1 200 Fahrten einer Nutzergruppe getestet. Es zeigt sich ein Vorteil des *End-to-End Memory Network Model* gegenüber einem *Variable Order Markov Model* aus dem Stand der Forschung. Die Zahl der fehlerhaften Prädiktionen über die Fahrten reduziert sich um ca. 5 %, mit der nutzerspezifischen Erweiterung um über 17 %.

Bei der Fahrprofilprädiktion wurden Muster in dem Bewegungsverhalten der Fahrzeuge auf einem Streckensegment erkannt. Über das *Fast Dynamic Time Warping*-Verfahren wurde die Ähnlichkeit in den Geschwindigkeitsprofilen berechnet und über ein spektrales Clusterverfahren wurden Klassen gebildet. Diesen wurde jeweils ein repräsentatives Profil zugewiesen. Als Prädiktionsmodell wurde eine *Multinomiale Logistische Regression* verwendet, das neben Klassen der bereits zurückgelegten Straßensegmente vor allem den Verkehrsfluss berücksichtigt. Über einen sequentiellen Ansatz kann der Prädiktionshorizont beliebig lang gestaltet werden. Neben der Geschwindigkeit wurden Konfidenzintervalle für die Gierrate und die Querschleunigung für die Straßensegmente und den zugehörigen Klassen ermittelt. Es wurde gezeigt, dass ebenso weitere Kenngrößen wie bspw. der Energieverbrauch ermittelt werden können. Der Verkehrsfluss wurde als Haupteinflussgröße für die Modellgüte bestimmt. Die zurückliegenden Streckensegmente zeigen einen relevanten Beitrag für die Prädiktion. Die Trefferquote beträgt im Durchschnitt

ca. 84 % und hat sich als abhängig von der Straßenklasse erwiesen. In mehreren Experimenten wurde das Prädiktionsmodell für Fahrten auf verschiedenen Routen eingesetzt und dadurch gezeigt, dass eine gute Übereinstimmung der prädizierten Größen der Geschwindigkeit und der Gierrate mit den aufgezeichneten Messgrößen besteht.

Um einen quantifizierbaren Nutzen der beiden Funktionen zu zeigen, wurde eine Betriebsstrategie für ein vernetztes PHEV entwickelt. Ein gleitender Horizont prädiziert dabei die Routenabschnitte und das Geschwindigkeitsprofil des Fahrzeugs im Back-End. Die Vorgabe für den SOC-Verlauf wurde zudem abhängig von der prädizierten Route und der Ladeinfrastruktur gestaltet. Die Optimierung erfolgte nach *Pontryagin's Minimum Prinzip*. Für über 200 Fahrten wurden Kraftstoffeinsparungen von ca. 8,0 % gegenüber einem nicht vorausschauenden Verfahren ermittelt.

In der Diskussion wurde überprüft, ob die Forschungsfragen geeignet beantwortet wurden. In dieser Arbeit wurde gezeigt, dass rein datengetriebene Funktionen zur Prädiktion des Trips und der Bewegungszustände im Fahrzeug möglich sind. Zudem können diese in Kombination zur Effizienzsteigerung von Fahrzeugen eingesetzt werden.

Das Fundament dieser Arbeit stellten die mit Hilfe von Smartphones aufgezeichneten Flotendaten. In mehreren Fällen wurden die erforderlichen Einschränkungen an die Genauigkeit aufgezeigt. Als Ausblick ist in erster Linie die Verwendung von interner Fahrzeugsensorik zu nennen, um die Qualität der Ergebnisse zu steigern und auch weitere Zeitreiheninformationen bspw. von Lenkradwinkel oder Fahrpedalstellungen zu integrieren. Ein weiterer wichtiger Punkt ist dabei die Größe der Datenbasis. Um die Aussagekraft möglichst hoch zu halten, ist eine weitaus größere Datenmenge erforderlich, die nur durch einen Automobilhersteller oder einen Echtzeitkartendienst bereitgestellt werden kann.

In dieser Arbeit wurden explizit rein historische Daten untersucht. Darüber hinaus ist es sinnvoll, weitere Echtzeitinformationen einzubinden. Verkehrsbehinderungen, wie die Kenntnis von Baustellen kann die Trip-Prädiktion verbessern. Der kurzfristige Horizont der Fahrprofilprädiktion kann mit der zuverlässigen Information von Ampelphasen oder einer Modellierung des Fahrzeugfolgeverhaltens auf Basis der Signale von Radarsensoren verbessert werden.

Als weiterer Schritt muss eine prototypische Implementierung auf Fahrzeug und Back-End vorgenommen werden, um mit Versuchen im Realverkehr und in Echtzeit den Mehrwert der entwickelten Funktionen zu zeigen.

Abbildungsverzeichnis

Abbildung 1.1	Infrastrukturtechnologien für den Umgang mit Fahrzeugdaten [12]	2
Abbildung 1.2	Fahreraufgaben im Zeitbereich nach Donges [35].	4
Abbildung 1.3	Überblick über den Aufbau dieser Arbeit	6
Abbildung 2.1	Aufbau Next Generation Telematics Pattern nach [69]	9
Abbildung 2.2	Übersicht über Service und Content Provider	10
Abbildung 2.3	Beispielhafter Verlauf von zwei Geschwindigkeitsprofilen auf einer Route (Karte: OSM © ⓘ ⓘ)	14
Abbildung 2.4	Vorgehensmodell mit Wissensgenerierung in Datenbanken aufbauend auf Fayyad [98]	15
Abbildung 2.5	Entwicklungsumgebung	16
Abbildung 3.1	Datenbasis der relevanten Flottenversuche nach [100] (Karte: OSM © ⓘ ⓘ)	17
Abbildung 3.2	Vergleich der Zeitreihen und Fehlerhistogramme	19
Abbildung 3.3	Vergleich von Datenquellen für Höheninformationen	21
Abbildung 3.4	Befahrene Straßenabschnitte im Stadtgebiet München (Ausschnitt zwi- schen Altstadt und Ostbahnhof)(Karte: OSM © ⓘ ⓘ)	22
Abbildung 3.5	Anteile an den Straßenklassen für unterschiedliche Nutzertypen nach [100]	22
Abbildung 3.6	Tagesverlauf des Verkehrsflusses auf Rosenheimer Straße	24
Abbildung 3.7	Tagesverlauf des Verkehrsflusses auf einem Abschnitt der A9	24
Abbildung 4.1	Semantische Beschreibung von Trajektorien	26
Abbildung 4.2	Ansatz der Trip-Prädiktion	35
Abbildung 4.3	Links: Einlagiges Netzwerk, Rechts: Dreilagiges Netzwerk nach Sukh- baatar [187]	36
Abbildung 4.4	Prädiktion des nächsten Streckensegments mit VOMM	39
Abbildung 4.5	Unterteilung in Gebiete nach [97] (Karte: OSM © ⓘ ⓘ)	39
Abbildung 4.6	Start- und Endpunkte: Links ein Taxi, Rechts ein Handwerksbetrieb (Karte: OSM © ⓘ ⓘ)	40
Abbildung 4.7	Vorgehen bei der Prädiktion (Karte: OSM © ⓘ ⓘ)	41
Abbildung 4.8	Isodistanz mit mehreren wahrscheinlichen Pfaden nach [97] (Karte: OSM © ⓘ ⓘ)	42
Abbildung 4.9	Relative Zahl der Fehlprädiktionen im Vergleich der Modelle	43
Abbildung 4.10	Reduktion der Fehlprädiktionen im zeitlichen Verlauf	43
Abbildung 4.11	Häufigkeit und Position der nutzerindividuellen Updates (Karte: OSM © ⓘ ⓘ)	44
Abbildung 4.12	Abhängigkeit der Prädiktionsgüte vom Horizont in Segmenten	44
Abbildung 4.13	Abhängigkeit der Prädiktionsgüte vom Horizont in km	45
Abbildung 4.14	Prädiktionsgüte bezogen auf die betrachtete Datenbasis	45
Abbildung 4.15	Vergleich von VOMM und MEM an einer Beispielfahrt (Karte: OSM © ⓘ ⓘ)	46
Abbildung 5.1	Ansatz der Fahrprofilprädiktion	56
Abbildung 5.2	Geschwindigkeitsprofil und Geschwindigkeitsganglinie nach Leutzbach [246, S. 32]	58
Abbildung 5.3	Projektion des optimalen Pfades aus niedriger Auflösung (links) in höhere Auflösungen (rechts) nach Salvador [251]	60
Abbildung 5.4	Skalierung der Distanzmatrix in den zweidimensionalen Raum	61
Abbildung 5.5	Geschwindigkeitsprofile und Clusterzugehörigkeit nach [244]	61
Abbildung 5.6	Einfluss der Anzahl der Fahrten auf die Clustergenerierung	62
Abbildung 5.7	Einfluss der Anzahl der Fahrten auf die Clustergenerierung	63
Abbildung 5.8	Kenngößen und Beziehungen der logistischen Regression	65
Abbildung 5.9	Sequentieller Ansatz der Prädiktion	68
Abbildung 5.10	Messwerte der Gierrate und der Querschleunigung bei Geradeausfahrt	69
Abbildung 5.11	Kurvenradien einer Autobahnausfahrt (Karte: OSM © ⓘ ⓘ)	70
Abbildung 5.12	Gierrate bei einer Autobahnausfahrt	71
Abbildung 5.13	Querschleunigung bei einer Autobahnausfahrt	71

Abbildung 5.14	Fahrzeit für ein innerstädtisches Segment	72
Abbildung 5.15	Häufigkeit der Cluster für ein Straßensegment nach [97]	73
Abbildung 5.16	Übereinstimmung der Profile mit den Cluster	74
Abbildung 5.17	Übereinstimmung der Beschleunigungen mit dem Cluster	74
Abbildung 5.18	Energieverbrauch bei einer Autobahnausfahrt	75
Abbildung 5.19	Energieverbrauch für ein innerstädtisches Segment.....	75
Abbildung 5.20	Anzahl der Fahrprofile pro Segment und Richtung.....	76
Abbildung 5.21	Anzahl der Profile pro Straßenklasse (Tab. 3.2)	76
Abbildung 5.22	Informationsmaßzahlen (Devianz, Akaikes- und Bayessesches Informati- onskriterium) für ein Segment abhängig von den Eingangsgrößen	77
Abbildung 5.23	Informationsmaßzahlen für ein Segment abhängig von der Anzahl der zurückliegenden Segmenten.....	78
Abbildung 5.24	Schrittweise Verbesserung der Modelle abhängig von den Erweiterungen.	80
Abbildung 5.25	Einfluss der zurückliegenden Segmente über alle Modelle.....	80
Abbildung 5.26	Schrittweise Verbesserung der Modelle abhängig von den zurückliegen- den Segmenten	81
Abbildung 5.27	Einfluss der Anzahl der zurückliegenden Segmente auf die Datengrundlage	81
Abbildung 5.28	Einfluss der Straßenkategorie auf die Trefferrate	82
Abbildung 5.29	Einfluss der Anzahl der Kategorien auf die Trefferrate	82
Abbildung 5.30	Einfluss des Horizonts auf die Datengrundlage	83
Abbildung 5.31	Streckenlänge und Segmentabschnitte für die Autobahnfahrt (Karte: OSM ©(i)©)	84
Abbildung 5.32	Einfluss des Horizonts auf die Trefferquote	84
Abbildung 5.33	Einfluss der Verkehrsbedingung auf das resultierende Fahrprofil	85
Abbildung 5.34	Fortschreitender Horizont bei einer innerstädtischen Fahrt	86
Abbildung 5.35	Prädizierter Gierratenverlauf für zwei Streckenabschnitte	86
Abbildung 5.36	Energieverbrauch prädiziert für zwei Autobahnfahrten.....	87
Abbildung 6.1	Ansatz der Systemarchitektur für eine online basierte Betriebsstrategie	90
Abbildung 6.2	Beispielfahrt über eine längere Distanz nach [273] (Karte: OSM ©(i)©) ...	93
Abbildung 6.3	Einfluss der Ladesäulen über Prädiktionshorizont auf die SOC Vorgabe....	94
Abbildung 6.4	Resultierender SOC Verlauf für verschiedene Strategien nach [273]	95
Abbildung 6.5	Ergebnisse für mehrere Fahrten unterschiedlicher Länge nach [273]	95
Abbildung C.1	Zeit für die Erstellung eines Trip-Prädiktion Modells	133
Abbildung C.2	Zeit für die Erstellung eines Fahrprofilprädiktion Modells	133

©(i)©Das in dieser Arbeit verwendete Kartenmaterial entstammt ©OpenstreetMap [112] unter der Open Data Commons Open Database Lizenz ist es frei verfügbar. Es ist freigestellt die Daten zu kopieren, weiterzugeben, zu übermitteln sowie anzupassen. Die Darstellung der Kartographie in den Kartenkacheln erfolgte mit Hilfe von ©Mapbox [286], das ebenfalls Daten von ©OpenStreetMap bezieht, die über die Creative-Commons Lizenz CC-BY-SA 2.0 frei zu Verfügung stehen.

Tabellenverzeichnis

Tabelle 1.1	Haupteinflussgrößen aus Fahrprofilen für Emissionen und Verbrauch nach Ericsson [28]	4
Tabelle 2.1	Beispiele für die Einordnung von Location Based Services.....	12
Tabelle 3.1	Übersicht zu den Kartenattributen	20
Tabelle 3.2	Übersicht zu den Straßensegmenten im Zieldatensatz	22
Tabelle 4.1	Literaturübersicht Fahrtzielprediktion	28
Tabelle 4.2	Literaturübersicht Routenprediktion	29
Tabelle 4.3	Literaturübersicht Trip-Prädiktion	33
Tabelle 5.1	Literaturübersicht mathematische Modelle	51
Tabelle 5.2	Literaturübersicht datenbasierte Ansätze	53
Tabelle 5.3	Literaturübersicht Bewertung von Fahrprofilen	55
Tabelle 5.4	Übersicht über die Modellausprägungen	64
Tabelle 5.5	Klassifikationstabelle	67
Tabelle 5.6	Beispiel p-Werte der Einflussfaktoren	79
Tabelle 5.7	Beispiel Ergebnis Klassifikationstabelle tertiäres Straßensegment	79
Tabelle 5.8	Übersicht zu den ausgewählten Strecken.....	83
Tabelle A.1	Literaturübersicht Trip-Prädiktion	131
Tabelle B.1	Literaturübersicht Fahrprofilprediktion	132

Formelzeichen

Formelzeichen

Formelzeichen	Einheit	Beschreibung
\bar{a}	$\frac{m}{s^2}$	Durchschnittliche Längsbeschleunigung
\bar{a}_n	$\frac{m}{s^2}$	Durchschnittlicher negativer Anteil der Längsbeschleunigung
\bar{a}_p	$\frac{m}{s^2}$	Durchschnittlicher positiver Anteil der Längsbeschleunigung
a_{tn}	%	Zeitanteil der negativen Längsbeschleunigung
a_{tp}	%	Zeitanteil der positiven Längsbeschleunigung
a_y	$\frac{m}{s^2}$	Querschleunigung
A	-	Kennzahl aus dem Beschleunigungs- und Verzögerungsverhalten
c	-	Eintrag im Ausgangsvektor in MEM
AIC	-	Akaikes Informationskriterium
BIC	-	Bayessches Informationskriterium
DEV	-	Devianz
F	N	Fahrwiderstände
H	-	Hamilton-Gleichung
i	-	Allgemeiner Index
I	-	Anzahl
j	-	Allgemeiner Index
J	-	Kosten zur Lösung eines Optimierungsproblems
K	-	Anzahl der Kategorien in MLR
l	-	Zahl der Eigenvektoren von L
l_b	-	Zahl der Cluster beim spektralen Clusterverfahren
$logit$	-	Logit-Wert in MLR
L	-	Kombinierte Wahrscheinlichkeit in MLR
LL	-	Log-Likelihood in MLR
m	-	Eintrag im Speichervektor in MEM
n	-	Allgemeiner Index
n_{stop}	-	Anzahl der Stopps
n_{ap}	-	Anzahl der Beschleunigungsvorgänge
n_{an}	-	Anzahl der Verzögerungsvorgänge
n_p	-	Anzahl der zu prädzierenden Streckensegmente
n_z	-	Anzahl der zurückliegenden Streckensegmente
N	-	Anzahl/Anzahl der unabhängigen Variablen in MLR
O	-	Komplexitätsgrad in DTW
$odds$	-	Chancenverhältnis in MLR
P	-	Wahrscheinlichkeit in MLR
P	W	Fahrleistung

p	-	Eintrag im Wahrscheinlichkeitsvektor in MEM
r	-	Radius in DTW
R	-	Kurvenradius
R^2	-	Bestimmtheitsmaß eines statistischen Modells
s	m	Weg
S	-	Segmentlänge
t	s	Zeit
T	N m	Drehmoment
TQ	%	Trefferquote
TI_S	$\frac{m}{s}$	Segment-spezifischer Verkehrsfluss
t_{stop}	s	Haltedauer
u	-	Stellgröße
v_s	$\frac{m}{s}$	Räumliche Geschwindigkeit
\bar{v}_{s25}^S	$\frac{m}{s}$	Segment-spezifisches 25-Perzentil der mittleren räumlichen Geschwindigkeit
v	$\frac{m}{s}$	Fahrgeschwindigkeit
v_t	$\frac{m}{s}$	Zeitliche Geschwindigkeit
w	-	Wort in MEM
W	-	Wald-Test
x	-	Zustandsgröße
Y	-	Abhängige Variable in MLR
y_{qc}	-	Element der Ähnlichkeitsmatrix in DTW
z	-	Aggregierte Einflussstärke in MLR
α	%	Signifikanzniveau in MLR
β	-	Regressionskoeffizient in MLR
$\dot{\beta}$	$\frac{^\circ}{s}$	Schwimmwinkelrate
γ	-	Kummulative Distanz in DTW
η	-	Lernrate in MEM
λ	-	Lagrange-Multiplikator
$\dot{\psi}$	$\frac{^\circ}{s}$	Gierrate
σ	-	Standardabweichung
ω	$\frac{1}{s}$	Drehgeschwindigkeit

Vektoren und Matrizen

Formelzeichen	Einheit	Beschreibung
\hat{A}	-	Ähnlichkeitsmatrix beim spektralen Clusterverfahren
c	-	Zeitreihenvektor in DTW
d	-	Differenz der Zeitreihenvektoren in DTW
D	-	Diagonalmatrix in DTW
DTW	-	DTW-Pfad
E	-	Einbettung in MEM
L	-	Normierte Ähnlichkeitsmatrix beim spektralen Clusterverfahren
q	-	Zeitreihenvektor in DTW
T	-	Merkmalsraum beim spektralen Clusterverfahren
T_A	-	Zeitliche Encodierung des Eingangs in MEM
T_C	-	Zeitliche Encodierung des Ausgangs in MEM
V	-	Wörterverzeichnis in MEM
W	-	Gewichtungsmatrix in MEM
w	-	Warping Pfad in DTW
x	-	Unabhängige Variablen in MLR
Y	-	Distanzmatrix in DTW
z	-	Störgröße

Indizes

Formelzeichen	Beschreibung
\bar{x}	Arithmetischer Mittelwert von x
\hat{x}	Prognose von x
x_i	i -tes Element von x
x_j	j -tes Element von x
x_n	n -tes Element von x

Abkürzungen

A

ACP	Application Communication Protocol	8
ADASIS	Advanced Driver Assistance Systems Interface Specifications	11
AIC	Akaikes Informationskriterium	67

B

BIC	Bayessches Informationskriterium	67
BRNN	Bidirektionales RNN	

C

C2C	Car-to-Car	7
C2I	Car-to-Infrastructure	7
CDP	Customer Data Provider	8
CNN	Convolutional Neural Network	
CP	Content Provider	8

D

DBSCAN	Dichtebasierte räumliche Clusteranalyse mit Rauschen	27
DEV	Devianz	66
DGM	Digitales Geländemodell	20
DP	Dynamische Programmierung	50
DSPT	Dispatcher	8
DTW	Dynamic Time Warping	52
DUC	Distance Until Charge	90

E

eCall	automatisches Notrufsystem	8
ECMS	Equivalent Consumption Minimization Strategy	92
ESP	Elektronischen Stabilitätsprogramm	2
ETSI	Europäisches Institut für Telekommunikationsnormen	11

F

FCD	Floating Car Data	8
------------	-------------------	---

G

GATS	Global Automotive Telematics Standard	8
GIS	Geoinformationssystem	9
GML	Geography Markup Language	11
GPS	Global Positioning System	9

H

HEV	Hybrid Electric Vehicle	48
HMM	Hidden Markov Model	27
HPC	Hochleistungsrechner	16

I

IDM	Intelligent Driver Model	48
------------	--------------------------	----

IKT	Informations- und Kommunikationstechnologie.....	7
IOT	Internet of Things	1
ISO	Internationale Organisation für Normung.....	7
ITS	Intelligent Transportation Systems.....	11

L

LBS	Location-based Services	3
LDM	Local Dynamic Map	11
LSA	Lichtsignalanlage	11
LTE	Long Term Evolution	7

M

MDS	Multidimensionale Skalierung	61
MEM	End-to-End Memory Network.....	34
MLR	Multinomiale Logistische Regression	64
MPP	Most Probable Path	30

N

NGTP	Next Generation Telematics Patterns	8
NTM	Neural Turing Machine	35

O

OBD	On-Board-Diagnose	1
OD	Origin-Destination.....	26
OGC	Open Geospatial Consortium	10
OSI	Open Systems Interconnection Model	7
OSM	OpenStreetMap.....	20

P

PCA	Principal Component Analysis	53
PDP	Provisioning Data Provider	8
PHEV	Plug-In Hybrid Electric Vehicle	48
PMP	Pontryagin's Minimum Prinzip	92
POI	Point of Interest	2
PSAP	Public Safety Answering Point	8

R

RNN	Rekurrentes Neuronales Netz	
------------	-----------------------------	--

S

SH	Service Handler	8
SI	Service Integrator	8
SOC	State of Charge	90
SP	Service Provider	8
SRTM	Shuttle Radar Topography Mission	21
SVM	Support Vector Machine.....	53
SVD	Singular Value Decomposition.....	51

T

TCP	Transmission Control Protocol	8
TU	Telematic Unit	8

U

UMTS	Universal Mobile Telecommunications System	8
USNG	United States National Grid	27

V

VEM	Virtuelle Elektromobilität für den Taxi- und Handwerksverkehr München	17
VOMM	Variable Order Markov Model	27

X

xFCD	extended FCD	8
-------------	--------------------	---

Literatur

- [1] KPMG International Cooperative, Hrsg., *KPMG's 17th Global Automotive Executive Survey 2016: From a product-centric world to a service-driven digital universe*, 2016. Adresse: <https://assets.kpmg.com/content/dam/kpmg/pdf/2016/01/gaes-2016-sec-de.pdf> (besucht am 01. 12. 2016).
- [2] H.-j. Gunther, O. Trauer und L. Wolf, "The potential of collective perception in vehicular ad-hoc networks", in *14th International Conference on ITS Telecommunications (ITST)*, IEEE, 2015, S. 1–5, ISBN: 978-1-4673-9382-9. DOI: 10.1109/ITST.2015.7377190.
- [3] A. Picot, S. Hopf und J. Sedlmeir, "Digitalisierung als Herausforderung für die Industrie – Das Beispiel der Automotive Branche", in *Technologie, Strategie und Organisation*, W. Burr und M. Stephan, Hrsg., Wiesbaden: Gabler, 2017, S. 87–112, ISBN: 978-3-658-16041-8. DOI: 10.1007/978-3-658-16042-5_5.
- [4] KPMG LLP, Hrsg., *Connected and Autonomous Vehicles - The UK Economic Opportunity*, 2015. Adresse: https://www.kpmg.com/BR/en/Estudos_Analises/artigosepublicacoes/Documents/Industrias/Connected-Autonomous-Vehicles-Study.pdf (besucht am 01. 12. 2016).
- [5] H.-P. Huber, J. Lohrer, J. Hutzschenreuther und M. Lienkamp, "Potenzialanalyse von innovativen Diensten für vernetzte Fahrzeuge", *ZfAW, Zeitschrift für die gesamte Wertschöpfungskette der Automobilwirtschaft*, Nr. 3, S. 12–22, 2015.
- [6] Commerz Finanz GmbH, Hrsg., *Vernetztes und autonomes Fahren. Was erwarten die Verbraucher? Automobilbarometer 2016 – International*, 2016.
- [7] F. Klanner, C. Ruhhammer, A. Bartsch, R. Rasshofer, W. Huber und S. Rauch, "Mehr Komfort und Sicherheit durch zunehmende Vernetzung", *Elektronik automotive*, Jg. 6, Nr. 7, S. 26–31, 2013.
- [8] A. Talebpour und H. S. Mahmassani, "Influence of connected and autonomous vehicles on traffic flow stability and throughput", *Transportation Research Part C: Emerging Technologies*, Jg. 71, S. 143–163, 2016, ISSN: 0968090X. DOI: 10.1016/j.trc.2016.07.007.
- [9] McKinsey & Company, Hrsg., *Ten ways autonomous driving could redefine the automotive world*, 2015. Adresse: <http://www.mckinsey.com/industries/automotive-and-assembly/our-insights/ten-ways-autonomous-driving-could-redefine-the-automotive-world> (besucht am 01. 12. 2016).
- [10] SAE International Surface Vehicle Recommended Practice, *Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles*. Adresse: http://standards.sae.org/j3016_201609/ (besucht am 07. 12. 2016).
- [11] A. Cacilo, S. Schmidt, P. Wittlinger, F. Herrmann, W. Bauer, O. Sawade, H. Doderer, M. Hartwig und V. Scholz, Hrsg., *Hochautomatisiertes Fahren auf Autobahnen - industriepolitische Schlussfolgerungen: Dienstleistungsprojekt 15/14, Studie im Auftrag des Bundesministeriums für Wirtschaft und Energie (BMWi)*, Stuttgart, 1. Jan. 2015. Adresse: <http://publica.fraunhofer.de/dokumente/N-372308.html> (besucht am 09. 01. 2017).
- [12] McKinsey & Company, Hrsg., *Monetizing car data: New service business opportunities to create new customer benefits*, 2016. Adresse: https://www.mckinsey.de/files/mckinsey_car_data.pdf (besucht am 01. 12. 2016).
- [13] Intel Corporation, Hrsg., *Technology and Computing Requirements for Self-Driving Cars: Moving beyond horsepower to deliver a new era of driving safety and convenience*, 2014. Adresse: <http://www.intel.com/content/dam/www/public/us/en/documents/white-papers/automotive-autonomous-driving-vision-paper.pdf> (besucht am 01. 12. 2016).

- [14] S. Bratzel, *Connected Car Innovation 2016 (CCI)*, Center of Automotive Management GmbH Co. KG, Hrsg., 2016.
- [15] Capgemini, Hrsg., *CARS ONLINE 2015 The Selfie Experience: The evolving behavior of the connected customer*, 2015. Adresse: https://www.de.capgemini.com/resource-file-access/resource/pdf/cars_online_2015_1.pdf (besucht am 01. 12. 2016).
- [16] BITKOM, Hrsg., *Leitlinien für den Big-Data-Einsatz: Chancen und Verantwortung*, 2015. Adresse: <https://www.bitkom.org/noindex/Publikationen/2015/Leitfaden/LF-Leitlinien-fuer-den-Big-Data-Einsatz/150901-Bitkom-Positionspapier-Big-Data-Leitlinien.pdf> (besucht am 01. 12. 2016).
- [17] K. Boriboonsomsin, M. J. Barth, W. Zhu und A. Vu, "Eco-Routing Navigation System Based on Multisource Historical and Real-Time Traffic Information", *IEEE Transactions on Intelligent Transportation Systems*, Jg. 13, Nr. 4, S. 1694–1704, 2012, ISSN: 1524-9050. DOI: 10.1109/TITS.2012.2204051.
- [18] M. Heimrath, P. Reinisch und M. Werling, "Fahrstabilisierung für ein Fahrzeug", DE102015 201383 (A1), 2016.
- [19] N. Casott und N. Mäthner, "Verfahren und Vorrichtung zur adaptiven Zielermittlung bei Navigationssystemen", DE102007024061 (A1), 2008.
- [20] D. S. Touney, "Navigation system and travel coordinator with dynamic traffic data", US2003065442 (A1), 2003.
- [21] C. Demir und B. Kerner, "Dynamic vehicle routing procedure uses off board processing in central unit with routing and updates for traffic conditions sent by wireless link", DE10324269 (A1), 2004.
- [22] S. Grubwinkler und M. Lienkamp, "A modular and dynamic approach to predict the energy consumption of electric vehicles", in *Conference on Future Automotive Technology*, München, 2013.
- [23] M. Barth und K. Boriboonsomsin, "Real-World Carbon Dioxide Impacts of Traffic Congestion", *Transportation Research Record: Journal of the Transportation Research Board*, Jg. 2058, Nr. 1, S. 163–171, 2008, ISSN: 0361-1981. DOI: 10.3141/2058-20.
- [24] P. Hemmerle, M. Koller, H. Rehborn, G. Hermanns, B. S. Kerner und M. Schreckenberger, "Increased Consumption in Oversaturated City Traffic Based on Empirical Vehicle Data", in *Advanced Microsystems for Automotive Applications 2014*, Ser. Lecture Notes in Mobility, J. Fischer-Wolfarth und G. Meyer, Hrsg., Cham: Springer International Publishing, 2014, S. 71–79, ISBN: 978-3-319-08086-4. DOI: 10.1007/978-3-319-08087-1_7.
- [25] G. Nilsson, "Traffic Safety Dimensions and the Power Model to Describe the Effect of Speed on Safety", Dissertation, Universität Lund, Lund, 2004.
- [26] E. I. Vlahogianni, M. G. Karlaftis und J. C. Golias, "Short-term traffic forecasting: Where we are and where we're going", *Transportation Research Part C: Emerging Technologies*, Jg. 43, S. 3–19, 2014, ISSN: 0968090X. DOI: 10.1016/j.trc.2014.01.005.
- [27] A. M. Rao und K. R. Rao, "Measuring Urban Traffic Congestion-A Review", *International Journal for Traffic and Transport Engineering*, Jg. 2, Nr. 4, S. 286–305, 2012. DOI: 10.7708/ijtte.2012.2(4).01.
- [28] E. Ericsson, "Independent driving pattern factors and their influence on fuel-use and exhaust emission factors", *Transportation Research Part D: Transport and Environment*, Jg. 6, Nr. 5, S. 325–345, 2001, ISSN: 13619209. DOI: 10.1016/S1361-9209(01)00003-7.
- [29] M. Kerper, C. Wewetzer, H. Trompeter, W. Kiess und M. Mauve, "Driving More Efficiently - The Use of Inter-Vehicle Communication to Predict a Future Velocity Profile", in *Vehicular Technology Conference (VTC Spring), 2011 IEEE 73rd*, 2011, S. 1–5, ISBN: 978-1-4244-8332-7. DOI: 10.1109/VETECS.2011.5956199.

- [30] C. Sun, X. Hu, S. J. Moura und F. Sun, "Velocity Predictors for Predictive Energy Management in Hybrid Electric Vehicles", *IEEE Transactions on Control Systems Technology*, Jg. 23, Nr. 3, S. 1197–1204, 2015, ISSN: 1063-6536. DOI: 10.1109/TCST.2014.2359176.
- [31] B. Bader, "An energy management strategy for plug-in hybrid electric vehicles", Dissertation, Universitat Politècnica de Catalunya, Barcelona, 2013.
- [32] H. Schüller, "Modelle zur Beschreibung des Geschwindigkeitsverhaltens auf Stadtstraßen und dessen Auswirkungen auf die Verkehrssicherheit auf Grundlage der Straßengestaltung", Dissertation, Technische Universität Dresden, Dresden, 2009.
- [33] R. Denaro, "Data mining in a digital map database to identify speed changes on upcoming curves along roads and enabling precautionary actions in a vehicle", US20090299624A1.
- [34] J. Cao und M. Menendez, "System dynamics of urban traffic based on its parking-related-states", *Transportation Research Part B: Methodological*, Jg. 81, S. 718–736, 2015, ISSN: 01912615. DOI: 10.1016/j.trb.2015.07.018.
- [35] E. Donges, "Fahrerverhaltensmodelle", in *Handbuch Fahrerassistenzsysteme: Grundlagen, Komponenten und Systeme für aktive Sicherheit und Komfort*, H. Winner, S. Hakuli, F. Lotz und C. Singer, Hrsg., Wiesbaden: Springer Fachmedien Wiesbaden, 2015, S. 17–26, ISBN: 978-3-658-05734-3. DOI: 10.1007/978-3-658-05734-3_2.
- [36] Generali Versicherung AG, Hrsg., *Generali Autostudie 2014*, 2014.
- [37] TEMA-Q Marktforschung, Hrsg., *Navigationssysteme: Teil 1: Nutzungsverhalten und Trends*, 2011. Adresse: http://www.tema-q.de/uploads/media/TEMA-Q_TdM_02-11_Navigationssysteme_Teil_I.pdf (besucht am 01.09.2017).
- [38] TOMORROW FOCUS Media GmbH, Hrsg., *Social Trends: Automobil und Mobilität der Zukunft*, 2015.
- [39] T. Skoglund, P. Wallgren, M. Karlsson und S. Franzén, "Users' Perception And Reported Effects Of Long-term Access To In-Vehicle Traffic Information Services Mediated Through Nomadic Devices", *Transport and Telecommunication Journal*, Jg. 16, Nr. 3, 2015, ISSN: 1407-6179. DOI: 10.1515/ttj-2015-0018.
- [40] M. Karlsson, T. Skoglund, P. Wallgren, M. Alonso, L. Guidotti, O. Martin und A. May, "Patterns of use, perceived benefits and reported effects of access to navigation support systems: An inter-European field operational test", *IET Intelligent Transport Systems*, Jg. 9, Nr. 8, S. 802–809, 2015, ISSN: 1751-956X. DOI: 10.1049/iet-its.2014.0233.
- [41] N. L. Forbes, "Behavioural adaptation to in-vehicle navigation systems", Thesis (PhD), University of Nottingham, Nottingham, 2009.
- [42] M. Göktürk und A. Pakkan, "Effects of In-Car Navigation Systems on User Perception of the Spatial Environment", in *Design, User Experience, and Usability. User Experience in Novel Technological Environments*, Ser. Lecture Notes in Computer Science/Information Systems and Applications, Incl. Internet/Web, and HCI, D. Hutchison, T. Kanade, J. Kittler u. a., Hrsg., Bd. 8014, Berlin: Springer, 2013, S. 57–64, ISBN: 978-3-642-39237-5. DOI: 10.1007/978-3-642-39238-2_7.
- [43] J. Speake und S. Axon, "I Never Use 'Maps' Anymore": Engaging with Sat Nav Technologies and the Implications for Cartographic Literacy and Spatial Awareness", *The Cartographic Journal*, Jg. 49, Nr. 4, S. 326–336, 2013, ISSN: 0008-7041. DOI: 10.1179/1743277412Y.0000000021.
- [44] Landesamt für Bürger- und Ordnungsangelegenheiten Berlin, Hrsg., *Ortskundekatalog*, 2015. Adresse: https://www.berlin.de/lab0/_assets/kraftfahrzeugwesen/ortskundekatalog_berlin_2015.pdf (besucht am 01.09.2017).
- [45] C. Buckl, A. Camek, L. Mercep, H. Stähle und A. Knoll, "The Software Car: Building ICT Architectures for Future Electric Vehicles", in *2012 IEEE International Electric Vehicle Conference*. DOI: 10.1109/IEVC.2012.6183198.

- [46] M. Broy, G. Reichart und L. Rothhardt, "Architekturen softwarebasierter Funktionen im Fahrzeug: von den Anforderungen zur Umsetzung", *Informatik-Spektrum*, Jg. 34, Nr. 1, S. 42–59, 2011, ISSN: 0170-6012. DOI: 10.1007/s00287-010-0507-6.
- [47] ForTISS GmbH, Hrsg., *Mehr Software (im) Wagen: Abschlussbericht des vom Bundesministerium für Wirtschaft und Technologie geförderten Verbundvorhabens "eCar-IKT-Systemarchitektur für Elektromobilität"*, 2012. Adresse: <http://download.fortiss.org/public/ikt2030/ikt2030de.pdf> (besucht am 01.09.2017).
- [48] V. Scheuch, "E/E-Architektur für Batterie-Elektrische Fahrzeuge", *ATZ Elektronik*, Nr. 06, S. 28–33, 2011, ISSN: 2192-8878.
- [49] M. Zeller und C. Prehofer, "Self-adaptation in Automotive Embedded Systems using a Multi-layered Control Approach", *Proceedings of the 2nd International Conference on Pervasive Embedded Computing and Communication Systems*, S. 459–468, 2012. DOI: 10.5220/0003942304590468.
- [50] T. Streichert und M. Traub, *Elektrik/Elektronik-Architekturen im Kraftfahrzeug*. Berlin und Heidelberg: Springer Berlin Heidelberg, 2012, ISBN: 978-3-642-25477-2. DOI: 10.1007/978-3-642-25478-9.
- [51] T. Weidener, D. Stüker, S. Wender und R. Katzwinkel, "Skalierbare E/E-Architekturen für Fahrerassistenzfunktionen", *ATZ Elektronik*, Nr. 06, S. 428–433, 2013, ISSN: 2192-8878. DOI: 10.1365/s35658-013-0361-9.
- [52] T. Tang, D. Soto-Setzke, C. Kohl, T. Köhn, J. Lohrer und J. Betz, "EE-Architektur für mobile Dienste", *ATZextra*, Jg. 19, Nr. 14, S. 40–45, 2014, ISSN: 2195-1454. DOI: 10.1365/s35778-014-1356-8.
- [53] R. Von Vignau, "Kommunikationsplattform für Autos in intelligenten Verkehrssystemen", *ATZ Elektronik*, Nr. 06, S. 20–25, 2011, ISSN: 2192-8878. DOI: 10.1365/s35658-011-0095-5.
- [54] B. Glas, O. Sander, K. Müller-Glaser und J. Becker, "Echtzeitfähige CAR-TO-X Kommunikationsabsicherung und E/E-Architekturintegration", *ATZ Elektronik*, Nr. 01, S. 68–76, 2012, ISSN: 2192-8878. DOI: 10.1007/978-3-658-04019-2.
- [55] H. Zimmermann, "OSI Reference Model–The ISO Model of Architecture for Open Systems Interconnection", *IEEE Transactions on Communications*, Jg. 28, Nr. 4, S. 425–432, 1980, ISSN: 0096-2244. DOI: 10.1109/TCOM.1980.1094702.
- [56] R. Schmidt, T. Leinmüller und B. Bötdecker, "V2X-Kommunikation", *17. Aachener Kolloquium Fahrzeug- und Motorentechnik*, 2008.
- [57] G. Behrens, V. Kuz und R. Behrens, *Softwareentwicklung von Telematikdiensten: Konzepte, Entwicklung und zukünftige Trends*, Ser. VDI-Buch. Hedielberg: Springer, 2010, ISBN: 978-3-540-88969-4.
- [58] K. Janschek, M. Scharfe, M. Buchroithner, N. Prechtel und E. Michler, "Synergetic telecommunication and Earth observation data fusion for road traffic monitoring", in *IAF, International Astronautical Congress, 50 th, Amsterdam, Netherlands*, 1999.
- [59] M. Linauer, "Generierung streckenbezogener Verkehrsdaten als Basis für den Einsatz in Verkehrstelematiksystemen", Dissertation, Universität für Bodenkultur, Wien, 2005.
- [60] European Commission, Hrsg., *DATEX II v2.3 MODELLING METHODOLOGY*, 2014. Adresse: http://www.datex2.eu/sites/www.datex2.eu/files/DATEX_II_Methodology_4.pdf (besucht am 01.09.2017).
- [61] R. Kelpin, A. Sohr, M. Umlauf, P. Brosch, G. Schimon und D. Pfoser, *Project Track&Trade: Deliverable 1.1 Data Model and Interfaces*, 2007. Adresse: <http://www.trackandtrade.org/wp-content/deliverables/D1.1.pdf> (besucht am 01.09.2017).

- [62] T. Stottan, "xFCD als Basistechnologie für die Mobilität 3.0", in *Schritte in die künftige Mobilität: Technische und betriebswirtschaftliche Aspekte*, H. Proff, W. Pascha, J. Schönharting und D. Schramm, Hrsg., Wiesbaden: Springer Fachmedien Wiesbaden, 2013, S. 47–59, ISBN: 978-3-8349-4308-8. DOI: 10.1007/978-3-8349-4308-8_4.
- [63] S. Wedeniwski, *The mobility revolution in the automotive industry: How not to miss the digital turnpike*. Heidelberg: Springer, 2015, ISBN: 978-3-662-47787-8.
- [64] A. Solberg, A. Schmid, M. Baggen u. a., *D.CVIS.3.4 Final Architecture and System Specifications: Deliverable*, CVIS, Hrsg., 2010.
- [65] U. Noyer, T. Schlauch, P. Cercato und L. Mikkelsen, "MOBiNET – architecture overview of an innovative platform for European mobility services", *22nd ITS World Congress, Bordeaux, France*, Nr. ITS-1980, 2015.
- [66] U. Noyer, T. Schlauch, B. Wissingh und L. Mikkelsen, "MOBiNET: Architecture and experience from a marketplace for mobility services", *11th ITS European Congress, Glasgow, Scotland*, Nr. EU-TP0005, 2016.
- [67] I. Passchier, M. van Sambeek und F. Ophelders, *DITCM Architecture Version 1.0*, DITCM Innovations, Hrsg., 2014.
- [68] M. van Sambeek, F. Ophelders, T. Bijlsma, B. van der Kluit, O. Turetken, R. Eshuis u. a., *Towards an Architecture for Cooperative ITS Applications in the Netherlands*, DITCM Innovations, Hrsg., 2015.
- [69] NGTP Group members, Hrsg., *NGTP in a nutshell*, 2010. Adresse: http://ngtp.org/wp-content/uploads/2013/12/NGTP20_nutshell.pdf (besucht am 01.09.2017).
- [70] M. Nyström und A. Bolinder, "Volvo IT - WirelessCar", *RedHat Forum*, 2014.
- [71] T. Haberle, L. Charissis, C. Fehling, J. Nahm und F. Leymann, "The Connected Car in the Cloud: A Platform for Prototyping Telematics Services", *IEEE Software*, Jg. 32, Nr. 6, S. 11–17, 2015, ISSN: 0740-7459. DOI: 10.1109/MS.2015.137.
- [72] A. J. Brimicombe, "GIS-Where are the frontiers now", in *Proceedings GIS 2002*, S. 33–45.
- [73] S. Steiniger, M. Neun, A. Edwardes und B. Lenz, "Foundations of LBS", *CartouCHE-Cartography for Swiss Higher Education. Obtido em*, Jg. 20, 2008.
- [74] R. Abdalla, *Introduction to geospatial information and communication technology (GeoICT)*. Switzerland: Springer, 2016, ISBN: 978-3-319-33602-2.
- [75] Open Geospatial Consortium Inc., Hrsg., *OpenGIS Location Services (OpenLS): Core Services*, 2008. Adresse: http://portal.opengeospatial.org/files/?artifact_id=22122 (besucht am 08.02.2017).
- [76] K. Virrantaus, J. Markkula, A. Garmash, V. Terziyan, J. Veijalainen, A. Katanosov und H. Tirri, "Developing GIS-supported location-based services", in *Proceedings of the second International Conference on Web Information Systems Engineering*, S. 66–75. DOI: 10.1109/WISE.2001.996708.
- [77] K. Tran Dang, N. Trong Phan und N. Chan Ngo, "An OpenLS privacy-aware middleware supporting location-based applications", *International Journal of Pervasive Computing and Communications*, Jg. 9, Nr. 4, S. 311–345, 2013. DOI: 10.1108/IJPC-09-2013-0024.
- [78] A. Cupper, G. Treu und C. Linnhoff-Popien, "TraX: A device-centric middleware framework for location-based services", *IEEE Communications Magazine*, Jg. 44, Nr. 9, S. 114–120, 2006, ISSN: 0163-6804. DOI: 10.1109/MCOM.2006.1705987.
- [79] M. Rao, V. S. Ramamurthy und B. Raj, *Standards, Spatial Framework and Technologies for National GIS*, National Institute of Advanced Studies, Hrsg., 2015.
- [80] Open Geospatial Consortium Inc., Hrsg., *OpenGIS geography markup language (GML) encoding standard*, 2007. Adresse: http://portal.opengeospatial.org/files/?artifact_id=20509 (besucht am 08.02.2017).

- [81] TomTom International B.V., Hrsg., *OpenLR™ White Paper*, 2012.
- [82] CEN, Hrsg., *SA 2016-06 Location Referencing Harmonization*, 2016. Adresse: <http://www.cencenelec.eu/News/Tenders/Tenders/call-2016-06.pdf> (besucht am 08.02.2017).
- [83] G. Vivo, *D7.3.1 LDM API and Usage Reference*, Safespot, Hrsg., 2010. Adresse: http://www.safespot-eu.org/documents/SF_D7.3.1_Annex2_LDM_API_and_Usage_Reference_v0.7.pdf (besucht am 01.09.2017).
- [84] ETSI, *Local Dynamic Map (LDM)-Rational for and guidance on standardization*, 2011. Adresse: http://www.etsi.org/deliver/etsi_tr/102800_102899/102863/01.01.01_60/tr_102863v010101p.pdf (besucht am 01.09.2017).
- [85] H. Shimada, A. Yamaguchi, H. Takada und K. Sato, "Implementation and Evaluation of Local Dynamic Map in Safety Driving Systems", *Journal of Transportation Technologies*, Jg. 05, Nr. 02, S. 102–112, 2015, ISSN: 2160-0473. DOI: 10.4236/jtts.2015.52010.
- [86] T. Eiter, H. Füreder, F. Kasslatner, J. Parreira und P. Schneider, "Towards a semantically enriched local dynamic map", in *Proceedings of ITS World Congress*, 2016. Adresse: https://publik.tuwien.ac.at/files/publik_257396.pdf (besucht am 01.09.2017).
- [87] ADASIS Forum, Hrsg., *ADASIS v2 Protocol*, 2013.
- [88] H. Pu, "Dynamic eHorizon with Traffic Light Information for Efficient Urban Traffic", in *Advanced Microsystems for Automotive Applications 2015*, Ser. Lecture Notes in Mobility, T. Schulze, B. Müller und G. Meyer, Hrsg., Cham: Springer International Publishing, 2016, S. 15–24, ISBN: 978-3-319-20854-1. DOI: 10.1007/978-3-319-20855-8_2.
- [89] Y. Horita und R. S. Schwartz, "Extended electronic horizon for automated driving", in *14th International Conference on ITS Telecommunications (ITST)*, IEEE, 2015, S. 32–36, ISBN: 978-1-4673-9382-9. DOI: 10.1109/ITST.2015.7377396.
- [90] D. Burgstahler, S. Zöller, M. Möbus, T. Walter, T. Rückelt und R. Steinmetz, "Navigate. KOM: Datenbankbasierter Informationsansatz für Fahrassistenzsysteme", *GMM-Fachbericht-AmE 2015–Automotive meets Electronics*, 2015.
- [91] D. Burgstahler, C. Peusens, D. Boehnstedt und R. Steinmetz, "Horizon.KOM: A First Step Towards an Open Vehicular Horizon Provider", in *VEHITS 2016*, M. Helfert und O. Gusikhin, Hrsg., Setubal, Portugal: SCITEPRESS - Science and Technology Publications Lda, 2016, S. 79–84, ISBN: 978-989-758-185-4. DOI: 10.5220/0005799700790084.
- [92] V. Protschky, S. Feld und M. Wälischmiller, "Traffic Signal Adaptive Routing", *2015 IEEE 18th International Conference on Intelligent Transportation Systems*, 2015, ISSN: 2153-0017. DOI: 10.1109/ITSC.2015.81.
- [93] A. Corti, V. Manzoni, S. M. Savaresi, M. D. Santucci und O. Tanna, "A Centralized Real-Time Driver Assistance System for Road Safety Based on Smartphone", in *Advanced Microsystems for Automotive Applications 2012*, S. 221–230. DOI: 10.1007/978-3-642-29673-4_20.
- [94] D. K. Jackson, "Systems and Methods for monitoring and reporting road quality", US9108640, 2015.
- [95] S. Grubwinkler, "Fahrprofilbasierte Energieverbrauchsprädiktion für vernetzte Elektrofahrzeuge", Dissertation, Technische Universität München, München, 2017.
- [96] W. Ertel, *Grundkurs Künstliche Intelligenz: Eine praxisorientierte Einführung*, 4. Aufl. 2016, Ser. Computational Intelligence. Wiesbaden: Springer Fachmedien Wiesbaden und Imprint: Springer Vieweg, 2016, ISBN: 3658135492.
- [97] J. Lohrer und M. Lienkamp, "An approach for predicting vehicle velocity in combination with driver turns", *Automotive and Engine Technology*, Nr. 1, S. 27–33, 2016, ISSN: 2365-5135. DOI: 10.1007/s41104-016-0012-2.
- [98] U. Fayyad, G. Piatetsky-Shapiro und P. Smyth, "From data mining to knowledge discovery in databases", *AI magazine*, Jg. 17, Nr. 3, S. 37, 1996. DOI: 10.1609/aimag.v17i3.1230.

- [99] G. Krempel, I. Žliobaite, D. Brzeziński, E. Hüllermeier, M. Last, V. Lemaire, T. Noack, A. Shaker, S. Sievi, M. Spiliopoulou u. a., “Open challenges for data stream mining research”, *ACM SIGKDD explorations newsletter*, Jg. 16, Nr. 1, S. 1–10, 2014. DOI: 10.1145/2674026.2674028.
- [100] M. Wittmann, J. Lohrer, J. Betz, B. Jäger, M. Kugler, M. Klöppel, A. Waclaw, M. Hann und M. Lienkamp, “A Holistic Framework for Acquisition, Processing and Evaluation of Vehicle Fleet Test Data”, *IEEE 20th International Conference on Intelligent Transportation Systems (ITSC)*, 2017.
- [101] T. Pesce, S. Schickram, J. Ritzer u. a., *Verbundvorhaben eFlott: Online-Analyse des Nutz- und Ladeverhaltens von Elektrofahrzeugen im Flottenversuch: Abschlussbericht*, 2012.
- [102] Bayern Innovativ, Hrsg., *Mobilitätsuntersuchungen mit MINI Elektrofahrzeugen im Kontext von Privat- und Flottennutzern in urbanen Zentren*, 2013. Adresse: <http://www.elektromobilitaet-verbindet.de/projekte/mobilitaetsuntersuchungen-mini.html> (besucht am 28.02.2017).
- [103] M. Kugler, C. Frank, S. Osswald, M. Miramontes, J. Kinigadner und M. Lienkamp, *sun2car GAP - Untersuchung des Kundenannahmeverhaltens zur Integration von Erneuerbarer Energie in Elektrofahrzeuge innerhalb der e-GAP Modellkommune Garmisch-Partenkirchen: Abschlussbericht*, 2016.
- [104] B. Jäger, M. Hann und M. Lienkamp, *VEM - Virtuelle Elektromobilität im Taxi- und Gewerbeverkehr München: Neuartiger Ansatz zur Untersuchung von technischen, wirtschaftlichen und ökologischen Gesichtspunkten einer elektrifizierten Fahrzeugflotte: Abschlussbericht*, Technische Universität München, Hrsg., 2016. DOI: 10.2314/GBV:871992922.
- [105] PostgreSQL Global Development Group, Hrsg., *PostgreSQL*. Adresse: <https://www.postgresql.org/> (besucht am 02.03.2017).
- [106] PostGIS Project, Hrsg., *PostGIS - Spatial and Geographic objects for PostgreSQL*. Adresse: <http://postgis.net/> (besucht am 02.03.2017).
- [107] J. Ritzer, “Methoden zur Messung der Bewegungszustände von Fahrzeugen mit mobilen Endgeräten”, Dissertation, Technische Universität München, München, 2016.
- [108] W.-C. Lee und B.-W. Cheng, “Comparison of portable and onboard navigation system for the effects in real driving”, *Safety science*, Jg. 48, Nr. 10, S. 1421–1426, 2010. DOI: 10.1016/j.ssci.2010.06.004.
- [109] Oxford Technical Solutions Limited, Hrsg., *User Manual - RTv2 GNSS-aided inertial measurement systems*, 2015.
- [110] A. Schultze, M. Hauck und M. Lienkamp, “Creation of driving cycles with lateral and longitudinal dynamics for analyzing energy losses in chassis”, in *Proceedings 16th Stuttgart International Symposium*, Stuttgart, 2016, S. 1199–1212, ISBN: 978-3-658-13254-5. DOI: 10.1007/978-3-658-13255-2_89.
- [111] L. Fahrmeir, C. Heumann, R. Künstler, I. Pigeot und G. Tutz, *Statistik: Der Weg zur Datenanalyse*, Ser. Springer-Lehrbuch. Berlin und Heidelberg: Springer Spektrum, 2016, ISBN: 978-3-662-50371-3.
- [112] OpenStreetMap Foundation, Hrsg., *OpenStreetMap Lizenzvereinbarung*. Adresse: <http://www.openstreetmap.org/copyright> (besucht am 01.03.2017).
- [113] C. Moeller, Hrsg., *osm2po*. Adresse: <http://osm2po.de/> (besucht am 02.03.2017).
- [114] Landesamt für Digitalisierung, Breitband und Vermessung, Hrsg., *GeoportalBayern*. Adresse: <https://geoportal.bayern.de/geoportalbayern/> (besucht am 07.03.2017).
- [115] Landesamt für Digitalisierung, Breitband und Vermessung, Hrsg., *Geländemodell*. Adresse: <http://www.ldbv.bayern.de/produkte/3dprodukte/gelaende.html> (besucht am 07.03.2017).

- [116] N. Jet Propulsion Laboratory, Hrsg., *Shuttle Radar Topography Mission*. Adresse: <http://www2.jpl.nasa.gov/srtm/> (besucht am 07. 03. 2017).
- [117] W. Cleveland und S. Delvin, "Locally Weighted Regression: An Approach to Regression Analysis by Local Fitting", *Journal of the American Statistical Association*, Jg. 83, Nr. 403, S. 596–610, 1988. DOI: 10.1080/01621459.1988.10478639.
- [118] N. Pelekis und Y. Theodoridis, *Mobility Data Management and Exploration*. New York, NY: Springer New York, 2014, ISBN: 978-1-4939-0391-7. DOI: 10.1007/978-1-4939-0392-4.
- [119] M. A. Quddus, W. Y. Ochieng und R. B. Noland, "Current map-matching algorithms for transport applications: State-of-the art and future research directions", *Transportation Research Part C: Emerging Technologies*, Jg. 15, Nr. 5, S. 312–328, 2007, ISSN: 0968090X. DOI: 10.1016/j.trc.2007.05.002.
- [120] Y. Lou, C. Zhang, Y. Zheng, X. Xie, W. Wang und Y. Huang, "Map-matching for low-sampling-rate GPS trajectories", in *Proceedings of the 17th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, 2009, S. 352–361, ISBN: 978-1-60558-649-6. DOI: 10.1145/1653771.1653820.
- [121] HERE Global B.V., Hrsg., *HERE WeGo*. Adresse: <https://wego.here.com/> (besucht am 03. 08. 2017).
- [122] R. Brunauer und K. Rehl, "Big Data in der Mobilität—FCD Modellregion Salzburg", in *Big Data*, Springer, 2016, S. 235–267, ISBN: 978-3-658-11589-0. DOI: 10.1007/978-3-658-11589-0_12.
- [123] HERE Global B.V., Hrsg., *REST APIs & Platform Extensions*. Adresse: <https://developer.here.com/develop/rest-apis> (besucht am 03. 08. 2017).
- [124] C. A. Quiroga und D. Bullock, "Travel time studies with global positioning and geographic information systems: an integrated methodology", *Transportation Research Part C: Emerging Technologies*, Jg. 6, Nr. 1-2, S. 101–127, 1998, ISSN: 0968090X. DOI: 10.1016/S0968-090X(98)00010-2.
- [125] F. Maier, "Abschnittsweise Regressionsanalyse zur Schätzung von Verkehrskenngrößen: Anwendung auf lokale Daten, Abschnittsdaten und Informationen aus Positionsmeldungen", Dissertation, Technische Universität München, München.
- [126] M. Treiber und A. Kesting, *Verkehrsdynamik und -simulation: Daten Modelle und Anwendungen der Verkehrsflussdynamik*, Ser. Springer-Lehrbuch. Heidelberg [u.a.]: Springer, 2010, ISBN: 978-3-642-05228-6.
- [127] Z. Yang und L. Yuncai, "Missing traffic flow data prediction using least squares support vector machines in urban arterial streets", in *Computational Intelligence and Data Mining, 2009. CIDM '09. IEEE Symposium on*, 2009, S. 76–83, ISBN: 978-1-4244-2765-9. DOI: 10.1109/CIDM.2009.4938632.
- [128] Q. Ye, W. Y. Szeto und S. C. Wong, "Short-Term Traffic Speed Forecasting Based on Data Recorded at Irregular Intervals", *IEEE Transactions on Intelligent Transportation Systems*, Jg. 13, Nr. 4, S. 1727–1737, 2012, ISSN: 1524-9050. DOI: 10.1109/TITS.2012.2203122.
- [129] M. T. Asif, J. Dauwels, C. Y. Goh, A. Oran, E. Fathi, M. Xu, M. M. Dhanya, N. Mitrovic und P. Jaillet, "Spatiotemporal Patterns in Large-Scale Traffic Speed Prediction", *IEEE Transactions on Intelligent Transportation Systems*, Jg. 15, Nr. 2, S. 794–804, 2014, ISSN: 1524-9050. DOI: 10.1109/TITS.2013.2290285.
- [130] C. Kühnel, "Verkehrsdatenerfassung mittels Floating Car Observer auf zweistreifigen Landstraßen", Dissertation, Universität Kassel, Kassel, 2012.
- [131] G. Andrienko und N. Andrienko, "Poster: Dynamic time transformation for interpreting clusters of trajectories with space-time cube", in *2010 IEEE Symposium on Visual Analytics Science and Technology*, 2010, S. 213–214, ISBN: 978-1-4244-9487-3. DOI: 10.1109/VAST.2010.5653580.

- [132] M. Lin und W.-J. Hsu, "Mining GPS data for mobility patterns: A survey", *Pervasive and Mobile Computing*, Jg. 12, S. 1–16, 2014, ISSN: 15741192. DOI: 10.1016/j.pmcj.2013.06.005.
- [133] A. M. Hendawi und M. F. Mokbel, "Predictive spatio-temporal queries: a comprehensive survey and future directions", in *Proceedings of the First ACM SIGSPATIAL International Workshop on Mobile Geographic Information Systems*, 2012, S. 97–104.
- [134] J. Zhang, Y. Zheng, D. Qi, R. Li und X. Yi, "DNN-based prediction model for spatio-temporal data", in *Proceedings of the 24th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, 2016, S. 92, ISBN: 978-1-4503-4589-7. DOI: 10.1145/2996913.2997016.
- [135] L. Etienne, T. Devogele und A. Bouju, "Spatio-temporal Trajectory analysis of mobile objects following the same itinerary", *Advances in Geo-Spatial Information Science*, 2012, ISSN: 978-0-203-12578-6.
- [136] Z. Yan, D. Chakraborty, C. Parent, S. Spaccapietra und K. Aberer, "Semantic Trajectories: Mobility Data Computation and Annotation", *ACM Transactions on Intelligent Systems and Technology*, 2009, ISSN: 2157-6904. DOI: 10.1145/2483669.2483682.
- [137] J. Yang, "Labeling Spatial Trajectories in Road Network using Probabilistic Graphical Models", Dissertation, Technische Universität München, München, 2016.
- [138] Y. Zheng, Q. Li, Y. Chen, X. Xie und W.-Y. Ma, "Understanding mobility based on GPS data", in *Proceedings of the 10th international conference on Ubiquitous computing*, 2008, S. 312–321, ISBN: 978-1-60558-136-1. DOI: 10.1145/1409635.1409677.
- [139] Y. Endo, H. Toda, K. Nishida und J. Ikedo, "Classifying spatial trajectories using representation learning", *International Journal of Data Science and Analytics*, Jg. 2, Nr. 3-4, S. 107–117, 2016, ISSN: 2364-415X. DOI: 10.1007/s41060-016-0014-1.
- [140] Y. Endo, H. Toda, K. Nishida und A. Kawanobe, "Deep Feature Extraction from Trajectories for Transportation Mode Estimation", in *Advances in Knowledge Discovery and Data Mining*, Ser. Lecture Notes in Computer Science, J. Bailey, L. Khan, T. Washio, G. Dobbie, J. Z. Huang und R. Wang, Hrsg., Bd. 9652, Cham: Springer International Publishing und Imprint: Springer, 2016, S. 54–66, ISBN: 978-3-319-31749-6. DOI: 10.1007/978-3-319-31750-2_5.
- [141] S. Khalid und A. Naftel, "Motion Trajectory Clustering for Video Retrieval Using Spatio-temporal Approximations", in *Visual Information and Information Systems*, Ser. Lecture Notes in Computer Science, D. Hutchison, T. Kanade, J. Kittler u. a., Hrsg., Bd. 3736, Berlin, Heidelberg: Springer Berlin Heidelberg, 2006, S. 60–70, ISBN: 978-3-540-30488-3. DOI: 10.1007/11590064_6.
- [142] T. Nawaz, A. Cavallaro und B. Rinner, "Trajectory clustering for motion pattern extraction in aerial videos", in *2014 IEEE International Conference on Image Processing (ICIP)*, 2014, S. 1016–1020, ISBN: 978-1-4799-5751-4. DOI: 10.1109/ICIP.2014.7025203.
- [143] K. Kim, D. Lee und I. Essa, "Gaussian process regression flow for analysis of motion trajectories", in *Computer Vision (ICCV), 2011 IEEE International Conference on*, 2011, S. 1164–1171, ISBN: 978-1-4577-1102-2. DOI: 10.1109/ICCV.2011.6126365.
- [144] M. Vlachos, G. Kollios und D. Gunopulos, "Discovering similar multidimensional trajectories", in *Data Engineering, 2002. Proceedings. 18th International Conference on*, 2002, S. 673–684, ISBN: 0-7695-1531-2. DOI: 10.1109/ICDE.2002.994784.
- [145] G.-P. Roh und S.-w. Hwang, "TPM: Supporting pattern matching queries for road-network trajectory data", in *Proceedings of the 14th International Conference on Extending Database Technology*, 2011, S. 554–557, ISBN: 978-1-4503-0528-0. DOI: 10.1145/1951365.1951439.
- [146] L. Zhang, G. Yang, Z. Wang und W. Huang, "Trajectory Clustering Based on Spatial Generalization", *Journal of Information & Computational Science*, Nr. 9(2), S. 315–321, 2012.

- [147] A. Peterson, "The origin-destination matrix problem-analysis and computations", Dissertation, Linköpings Universitet, Norrköping, 2007.
- [148] L. Moreira-Matias, J. Gama, M. Ferreira, J. Mendes-Moreira und L. Damas, "Time-evolving O-D matrix estimation using high-speed GPS data streams", *Expert Systems with Applications*, Jg. 44, S. 275–288, 2016, ISSN: 09574174. DOI: 10.1016/j.eswa.2015.08.048.
- [149] J. Krumm, "Real time destination prediction based on efficient routes", in *SAE 2006 World Congress & Exhibition*. DOI: 10.4271/2006-01-0811.
- [150] J. Alvarez-Garcia, J. Ortega, L. Gonzalez-Abril und F. Velasco, "Trip destination prediction based on past GPS log using a Hidden Markov Model", *Expert Systems with Applications*, Jg. 37, Nr. 12, S. 8166–8171, 2010, ISSN: 09574174. DOI: 10.1016/j.eswa.2010.05.070.
- [151] S. Gambs, M.-O. Killijian und Cortez, Miguel Nunez Del Prado, "Towards temporal mobility markov chains", in *1st International Workshop on Dynamicity Collocated with OPODIS 2011, Toulouse, France, 2012*, 2–pages.
- [152] C. Manasseh und R. Sengupta, "Predicting driver destination using machine learning techniques", in *Intelligent Transportation Systems - (ITSC), 2013 16th International IEEE Conference on*, 2013, S. 142–147, ISBN: 978-1-4799-2914-6. DOI: 10.1109/ITSC.2013.6728224.
- [153] M. Chen, Y. Liu und X. Yu, "Nlpm: A next location predictor with markov modeling", in *Pacific-Asia Conference on Knowledge Discovery and Data Mining*, 2014, S. 186–197, ISBN: 978-3-319-06605-9. DOI: 10.1007/978-3-319-06605-9_16.
- [154] J. Yang, J. Xu, M. Xu, N. Zheng und Y. Chen, "Predicting next location using a variable order Markov model", in *Proceedings of the 5th ACM SIGSPATIAL International Workshop on GeoStreaming (IWGS) 2014*, C. Zhang, A. Basalamah und A. Hendawi, Hrsg., New York: ACM, 2014, S. 37–42, ISBN: 9781450331395. DOI: 10.1145/2676552.2676557.
- [155] Q. Liu, S. Wu, L. Wang und T. Tan, "Predicting the Next Location: A Recurrent Model with Spatial and Temporal Contexts", in *Thirtieth AAAI Conference on Artificial Intelligence*, 2016, S. 194–200, ISBN: 9781577357605.
- [156] A. de Brébisson, É. Simon, A. Auvolet, P. Vincent und Y. Bengio, "Artificial neural networks applied to taxi destination prediction", *CoRR*, Nr. abs/1508.00021, 2015.
- [157] J. Lv, Q. Li und X. Wang, "T-CONV: A Convolutional Neural Network For Multi-scale Taxi Trajectory Prediction", *CoRR*, Nr. abs/1611.07635, 2016.
- [158] R. Simmons, B. Browning, Z. Yilu und V. Sadekar, "Learning to Predict Driver Route and Destination Intent", in *Intelligent Transportation Systems Conference, 2006. ITSC '06. IEEE*, 2006, S. 127–132, ISBN: 1-4244-0093-7. DOI: 10.1109/ITSC.2006.1706730.
- [159] A. Carlsson und H.-C. Reuss, "An incremental learning method for foresight information used in predictive driving strategies", *7. Internationales Stuttgarter Symposium*, S. 377–396, 2007.
- [160] A. Carlsson und H.-C. Reuss, "Cost-Efficient Provision of Accurate Foresight Route Information", *4th International Symposium - Networks for Mobility*, 2008.
- [161] A. Piegsa und H.-C. Reuss, "The Learning Vehicle: A virtual co-driver as trip companion?", *Themenheft Forschung - Intelligente Fahrzeuge*, Nr. 7, S. 48–58, 2011.
- [162] J. Froehlich und J. Krumm, "Route Prediction from Trip Observations", *World Congress - Society of Automotive Engineers (SAE)*, 2008.
- [163] V. S. Tiwari, A. Arya und S. Chaturvedi, "Route prediction using trip observations and map matching", in *Advance Computing Conference (IACC), 2013 IEEE 3rd International*, 2013, S. 583–587, ISBN: 978-1-4673-4527-9. DOI: 10.1109/IAAdCC.2013.6514292.

- [164] E. Necula, "Mining GPS Data to Learn Driver's Route Patterns", in *Symbolic and Numeric Algorithms for Scientific Computing (SYNASC), 2014 16th International Symposium on*, 2014, S. 264–271, ISBN: 978-1-4799-8447-3. DOI: 10.1109/SYNASC.2014.43.
- [165] C. Ress, A. Etermad, D. Kuck und M. Boerger, "Electronic Horizon-Supporting ADAS applications with predictive map data", in *Proceedings of the 13th ITS World Congress, London, 2006*.
- [166] D. Burgstahler, A. Xhoga, C. Peusens, M. Möbus, D. Böhnstedt und R. Steinmetz, "RemoteHorizon. KOM: Dynamic Cloud-based eHorizon", in *AmE 2016 - Automotive meets Electronics; 7th GMM-Symposium*, 2016, ISBN: 978-3-8007-4167-0.
- [167] S. V. Bechtolsheim, L. Dunn, A. Hecht, M. Schmitt, J. Feigen und M. Roser, "Map data architecture platform for advanced driver assistent systems", EP 2 042 833 A1, 2009.
- [168] R. Gee und U. Stählin, "Prädiktiver eHorzion", DE 10 2010 007 260 A1, 2011.
- [169] M. Haug, J. Moennich und G. Schaaf, "Verfahren zum Bestimmen eines am ehesten wahrscheinlichen Fahrpfads", DE102009028299A1, 2011.
- [170] G. Ghisio, "System and method for estimating the most probable path of a vehicle travelling on a road", US20130173150 A1, 2012.
- [171] P. Engel, J. W. Balkema und A. Varchmin, "Verfahren und Anordnung zum Bestimmen eines am ehesten wahrscheinlichen Fahrpfads eines Fahrzeugs", DE102011078946A1, 2013.
- [172] K. Mattmüller, R. Haimerl, M. Rauch und R. Varnhagen, "Verbesserung eines Fahrhorizontsystems durch Selbstlernen des höchstwahrscheinlichen Fahrpfades", DE102014223 331A1, 2016.
- [173] Y. Tao, C. Faloutsos, D. Papadias und B. Liu, "Prediction and indexing of moving objects with unknown motion patterns", in *Proceedings of the 2004 ACM SIGMOD international conference on Management of data*, 2004, S. 611–622, ISBN: 1-58113-859-8. DOI: 10.1145/1007568.1007637.
- [174] H. A. Karimi und X. Liu, "A predictive location model for location-based services", in *Proceedings of the 11th ACM international symposium on Advances in geographic information systems*, 2003, S. 126–133, ISBN: 1-58113-730-3. DOI: 10.1145/956676.956693.
- [175] J. Krumm, "A Markov Model for Driver Turn Prediction", *SAE 2008 World Congress*, 2008.
- [176] G. Xue, Z. Li, H. Zhu und Y. Liu, "Traffic-Known Urban Vehicular Route Prediction Based on Partial Mobility Patterns", in *Parallel and Distributed Systems (ICPADS), 2009 15th International Conference on*, 2009, S. 369–375, ISBN: 978-1-4244-5788-5. DOI: 10.1109/ICPADS.2009.129.
- [177] A. Monreale, F. Pinelli und R. Trasarti, "WhereNext: a location predictor on trajectory pattern mining", in *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*, 2009, ISBN: 978-1-60558-495-9. DOI: 10.1145/1557019.1557091.
- [178] H. Jeung, M. L. Yiu, X. Zhou und C. S. Jensen, "Path prediction and predictive range querying in road network databases", *The VLDB Journal*, Jg. 19, Nr. 4, S. 585–602, 2010, ISSN: 1066-8888. DOI: 10.1007/s00778-010-0181-y.
- [179] G. Gidófalvi, C. Borgelt, M. Kaul und T. B. Pedersen, "Frequent Route Based Continuous Moving Object Location- and Density Prediction on Road Networks", in *Proceedings of the 19th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, Ser. GIS '11, New York, NY, USA: ACM, 2011, S. 381–384, ISBN: 978-1-4503-1031-4. DOI: 10.1145/2093973.2094028.
- [180] B. Mao, J. Cao, Z. Wu, G. Huang und J. Li, "Predicting Driving Direction with Weighted Markov Model", in *ADMA 2012: Advanced Data Mining and Applications*, 2012, S. 407–418, ISBN: 978-3-642-35527-1. DOI: 10.1007/978-3-642-35527-1_34.

- [181] M. H. S. Eldaw, M. Levene und G. Roussos, "Collective Suffix Tree-based Models For Location Prediction", in *Proceedings of the 2013 ACM conference on Pervasive and ubiquitous computing adjunct publication*, 2013, S. 441–450. DOI: 10.1145/2494091.2495976.
- [182] J. Zhou, A. K. H. Tung, W. Wu und W. S. Ng, "A "semi-lazy" approach to probabilistic path prediction in dynamic environments", in *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining*, 2013, S. 748–756, ISBN: 978-1-4503-2174-7. DOI: 10.1145/2487575.2487609.
- [183] H. Huo, S.-y. Chen, B. Xu und L. Liu, "A Trajectory Prediction Method for Location-Based Services", in *Web Technologies and Applications*, Ser. Lecture Notes in Computer Science, R. Cai, K. Chen, L. Hong, X. Yang, R. Zhang und L. Zou, Hrsg., Bd. 9461, Cham: Springer-Verlag New York Inc, 2016, S. 127–138, ISBN: 978-3-319-28120-9. DOI: 10.1007/978-3-319-28121-6_12.
- [184] J. Zhang, Y. Zheng und D. Qi, "Deep Spatio-Temporal Residual Networks for Citywide Crowd Flows Prediction", *CoRR*, Nr. abs/1610.00081, 2016.
- [185] R. Trasarti, R. Guidotti, A. Monreale und F. Giannotti, "MyWay: Location prediction via mobility profiling", *Information Systems*, Jg. 64, S. 350–367, 2017, ISSN: 03064379. DOI: 10.1016/j.is.2015.11.002.
- [186] J. Weston, S. Chopra und A. Bordes, "Memory networks", *CoRR*, Nr. abs/1410.3916, 2014.
- [187] S. Sukhbaatar, J. Weston, R. Fergus u. a., "End-to-end memory networks", in *Advances in neural information processing systems*, 2015, S. 2440–2448, ISBN: 9781510825024.
- [188] A. Graves, G. Wayne und I. Danihelka, "Neural turing machines", *CoRR*, Nr. abs/1410.5401, 2014.
- [189] I. Goodfellow, Y. Bengio und A. Courville, *Deep Learning*. MIT Press, 2016, ISBN: 9780262035613.
- [190] D. Rumelhart, G. Hinton und R. Williams, "Learning representations by back-propagating errors", *Nature*, Jg. 323, Nr. 6088, S. 533–538, 1986. DOI: 10.1038/323533a0.
- [191] L. Bottou, "Large-Scale Machine Learning with Stochastic Gradient Descent", in *Proceedings of COMPSTAT'2010*, Ser. SpringerLink: Springer e-Books, Y. Lechevallier und G. Saporta, Hrsg., Heidelberg: Springer-Verlag Berlin Heidelberg, 2010, S. 177–186, ISBN: 978-3-7908-2603-6. DOI: 10.1007/978-3-7908-2604-3_16.
- [192] R. Gudishala und C. Wilmot, "Comparison of Time-Dependent Sequential Logit and Nested Logit for Modeling Hurricane Evacuation Demand", *Transportation Research Record: Journal of the Transportation Research Board*, Jg. 2312, S. 134–140, 2012, ISSN: 0361-1981. DOI: 10.3141/2312-14.
- [193] B. Letham, C. Rudin und D. Madigan, "Sequential event prediction", *Machine Learning*, Jg. 93, Nr. 2-3, S. 357–380, 2013, ISSN: 0885-6125. DOI: 10.1007/s10994-013-5356-5.
- [194] M. Treiber, A. Hennecke und D. Helbing, "Congested traffic states in empirical observations and microscopic simulations", *Physical Review E*, Jg. 62, Nr. 2, S. 1805–1824, 2000. DOI: 10.1103/PhysRevE.62.1805.
- [195] P. Ondruska und I. Posner, "The Route Not Taken: Driver-Centric Estimation of Electric Vehicle Range", in *Twenty-Fourth International Conference on Automated Planning and Scheduling*, 2014, ISBN: 978-1-57735-660-8.
- [196] W. Schakel, G. Klunder, B. van Arem, E. Harmsen und M. Hagenzieker, "Reducing travel delay by in-car advice on speed, headway and lane use based on downstream traffic flow conditions – a simulation study", *15th Edition of the Euro Working Group on Transportation*, 2012.
- [197] M. Back, "Prädiktive Antriebsregelung zum energieoptimalen Betrieb von Hybridfahrzeugen", Dissertation, Universität Karlsruhe, Karlsruhe, 2006.

- [198] J.-W. Biermann, A. Barkow und F. Töpler, "Vorausschauende Betriebsstrategie für Hybridfahrzeuge der Subkompaktklasse", *ATZ Elektronik*, Jg. 4, Nr. 06, S. 10–19, 2009, ISSN: 2192-8878. DOI: 10.1007/BF03223983.
- [199] C. Minett, M. Salomons, W. Daamen, B. van Arem und S. Kuijpers, "Eco-Routing: Comparing the Fuel Consumption of Different Routes between an Origin and Destination Using Field Test Speed Profiles and Synthetic Speed Profiles", in *2011 IEEE Forum on Integrated and Sustainable Transportation Systems*, 2011, ISBN: 978-1-4577-0992-0. DOI: 10.1109/FISTS.2011.5973621.
- [200] D. Karbowski, S. Pagerit und A. Calkins, "Energy consumption prediction of a vehicle along a user-specified real-world trip", in *Proceedings from the Electric Vehicle Symposium (EVS26)*, 2012, ISBN: 9781622764211.
- [201] P. Themann und L. Eckstein, "Prediction of driver behaviour by situational models in a cooperative environment to optimize energy efficiency", *9th ITS European Congress, Dublin, 4/7 Juni*, 2013.
- [202] Forschungsgesellschaft Kraftfahrtwesen mbH Aachen, Hrsg., *PELOPS White Paper*, 2012. Adresse: http://www.fka.de/pdf/pelops_whitepaper.pdf (besucht am 01.09.2017).
- [203] P. Themann, A. Zlocki und L. Eckstein, "Energieeffiziente Fahrzeuglängsführung durch V2X-Kommunikation", *ATZ-Automobiltechnische Zeitschrift*, S. 27–33, ISSN: 2192-8800. DOI: 10.1007/s35148-014-0452-9.
- [204] A. Engstle, A. Zinkl, A. Angermaier und W. Schelter, "Elektrifizierter Antriebsstrang - mehr Effizienz durch vorausschauendes Energiemanagement", in *Fahrerassistenzsysteme und Effiziente Antriebe*, Ser. SpringerLink : Bücher, W. Siebenpfeiffer, Hrsg., Wiesbaden: Springer Vieweg, 2015, S. 146–154, ISBN: 3658081619. DOI: 10.1007/978-3-658-08161-4_18.
- [205] T. van Keulen, G. Naus, B. de Jager, R. van de Molengraft, M. Steinbuch und E. Aneke, "Predictive cruise control in hybrid electric vehicles", *World Electric Vehicle Journal Vol. 3*, 2009.
- [206] N. J. Kohut, "Integrating traffic data and model predictive control to improve fuel economy", *12th IFAC Symposium on Control in Transportation Systems*, Nr. 42-15, S. 155–160, 2009. DOI: 10.3182/20090902-3-US-2007.0032.
- [207] L. Daeheung, C. Suk Won, R. Aymeric, K. Namwook und K. Dominik, "Optimal Control Strategy for PHEVs using Prediction of Future Driving Schedule", *World Electric Vehicle Journal Vol. 5*, 2012, ISSN: 2032-6653.
- [208] P. Markschläger, H.-G. Wahl, F. Weberbauer und M. Lederer, "Assistenzsystem für mehr Kraftstoffeffizienz", in *Vernetztes Automobil*, W. Siebenpfeiffer, Hrsg., Wiesbaden: Springer Fachmedien Wiesbaden, S. 146–153, ISBN: 365804019X. DOI: 10.1007/978-3-658-04019-2_21.
- [209] T. Radke, "Energieoptimale Längsführung von Kraftfahrzeugen durch Einsatz vorausschauender Fahrstrategien", Dissertation, Karlsruher Institut für Technologie, Karlsruhe, 2013.
- [210] H. G. Wahl und F. Gauterin, "An iterative dynamic programming approach for the global optimal control of hybrid electric vehicles under real-time constraints", in *Intelligent Vehicles Symposium (IV), 2013 IEEE*, 2013, S. 592–597, ISBN: 978-1-4673-2754-1. DOI: 10.1109/IVS.2013.6629531.
- [211] J. Wollaeger, S. A. Kumar, S. Onori, D. Filev, U. Ozguner, G. Rizzoni und S. Di Cairano, "Cloud-computing based velocity profile generation for minimum fuel consumption: A dynamic programming based solution", in *American Control Conference (ACC), 2012*, 2012, S. 2108–2113, ISBN: 978-1-4577-1095-7.
- [212] B. Saerens, "Optimal Control Based Eco-Driving: Theoretical Approach and Practical Applications", Dissertation, KU Leuven, Heverlee, Belgien, 2012.

- [213] G. Zhao und C. Wu, "Mathematical Modeling of Driver Speed Control With Individual Differences", *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, Jg. 43, Nr. 5, S. 1091–1104, 2013, ISSN: 2168-2216. DOI: 10.1109/TSMC.2013.2256854.
- [214] S. Schroedl und Z. Wenbing, "Predicting driving speed using neural networks", in *Intelligent Transportation Systems, 2003. Proceedings. 2003 IEEE*, Bd. 1, 2003, S. 402–407, ISBN: 0-7803-8125-4. DOI: 10.1109/ITSC.2003.1251985.
- [215] R. Bartholomaeus, M. Klingner und M. Lehnert, "Prediction of power demand for hybrid vehicles operating in fixed-route service", in *Proceedings of the 17th IFAC World Congress*, 2008, S. 6–11. DOI: 10.3182/20080706-5-KR-1001.00951.
- [216] G. Ciccicarese, T. Donateo und C. Palazzo, "On-board prediction of future speed profile for energy management of hybrid electric vehicles", *International Journal of Automotive Technology and Management*, Nr. 12, S. 232–251, 2012, ISSN: 1470-9511. DOI: 10.1504/IJATM.2012.050357.
- [217] T. Donateo, "Intelligent Usage of Internal Combustion Engines in Hybrid Electric Vehicles", in *Internal Combustion Engines*, K. Lejda, Hrsg., InTech, 2012, ISBN: 978-953-51-0856-6. DOI: 10.5772/48120.
- [218] D. Krajzewicz, G. Hertkorn, C. Rössel und P. Wagner, "SUMO (Simulation of Urban MObility) - an open-source traffic simulation", in *4th Middle East Symposium on Simulation and Modelling*, A. Al-Akaidi, Hrsg., 2002, S. 183–187. Adresse: <http://elib.dlr.de/6661/>.
- [219] C. Sun, X. Hu, S. J. Moura und F. Sun, "Comparison of Velocity Forecasting Strategies for Predictive Control in HEVs", in *ASME 2014 Dynamic Systems and Control Conference*, 2014, V002T20A003, ISBN: 978-0-7918-4619-3. DOI: 10.1115/DSCC2014-6031.
- [220] C. Sun, F. Sun und H. He, "Investigating adaptive-ECMS with velocity forecast ability for hybrid electric vehicles", *Applied Energy*, Jg. 185, S. 1644–1653, 2017, ISSN: 03062619. DOI: 10.1016/j.apenergy.2016.02.026.
- [221] A. Rezaei und J. B. Burl, "Prediction of Vehicle Velocity for Model Predictive Control", *IFAC-PapersOnLine*, Jg. 48, Nr. 15, S. 257–262, 2015, ISSN: 24058963. DOI: 10.1016/j.ifacol.2015.10.037.
- [222] X. Geng, H. Liang, H. Xu, B. Yu und M. Zhu, "Human-driver speed profile modeling for autonomous vehicle's velocity strategy on curvy paths", in *2016 IEEE Intelligent Vehicles Symposium (IV)*, 2016, S. 755–760, ISBN: 978-1-5090-1821-5. DOI: 10.1109/IVS.2016.7535472.
- [223] M. Kerper, C. Wewetzer und M. Mauve, "Analyzing vehicle traces to find and exploit correlated traffic lights for efficient driving", in *Intelligent Vehicles Symposium (IV), 2012 IEEE*, 2012, S. 310–315, ISBN: 978-1-4673-2119-8. DOI: 10.1109/IVS.2012.6232143.
- [224] M. Kerper, C. Wewetzer, A. Sasse und M. Mauve, "Learning Traffic Light Phase Schedules from Velocity Profiles in the Cloud", in *New Technologies, Mobility and Security (NTMS), 2012 5th International Conference on*, 2012, S. 1–5, ISBN: 978-1-4673-0228-9. DOI: 10.1109/NTMS.2012.6208704.
- [225] M. Kerper und C. Wewetzer, "Verfahren und Vorrichtung zum Bereitstellen einer Fahrempfehlung für einen Streckenabschnitt", DE102010054077A1, 2012.
- [226] M. Kerper, "Ursachen- und wirkungsorientierte Analyse gesammelter Fahrprofile zur taktischen Fahrtoptimierung", Dissertation, Heinrich-Heine-Universität, Düsseldorf, 2013.
- [227] T. Mauk, "Selbstlernende zuverlässigkeitsorientierte Prädiktion energetisch relevanter Größen im Kraftfahrzeug", Dissertation, Universität Stuttgart, Stuttgart, 2011.
- [228] D. Karbowski, V. Smis-Michel und V. Vermeulen, "Using trip information for PHEV fuel consumption minimization", in *2013 World Electric Vehicle Symposium and Exhibition (EVS27)*, 2013, S. 1–12, ISBN: 978-1-4799-3832-2. DOI: 10.1109/EVS.2013.6914710.

- [229] J. Park, Y. L. Murphey, J. Kristinsson, R. McGee, M. Kuang und T. Phillips, "Intelligent speed profile prediction on urban traffic networks with machine learning", in *The 2013 International Joint Conference on Neural Networks (IJCNN)*, 2013, S. 1–7, ISBN: 978-1-4673-6129-3. DOI: 10.1109/IJCNN.2013.6707119.
- [230] X. Wang, J. Park, Y. L. Murphey, J. Kristinsson, M. Kuang und T. Phillips, "Intelligent Trip Modeling on Ramps using ramp classification and knowledge base", in *Proceedings of the 2014 International Joint Conference on Neural Networks*, 2014, S. 967–974, ISBN: 978-1-4799-1484-5. DOI: 10.1109/IJCNN.2014.6889971.
- [231] J. Lemieux und Y. Ma, "Vehicle Speed Prediction Using Deep Learning", in *2015 IEEE Vehicle Power and Propulsion Conference (VPPC)*, 2015, S. 1–5, ISBN: 978-1-4673-7637-2. DOI: 10.1109/VPPC.2015.7353037.
- [232] J. Yoon, B. Noble und M. Liu, "Surface street traffic estimation", in *Proceedings of the 5th international conference on Mobile systems, applications and services*, E. W. Knightly, Hrsg., New York, NY: ACM, 2007, S. 220–232, ISBN: 1595936149. DOI: 10.1145/1247660.1247686.
- [233] M. Montazeri-Gh und A. Fotouhi, "Traffic condition recognition using the -means clustering method", *Scientia Iranica*, Jg. 18, Nr. 4, S. 930–937, 2011, ISSN: 10263098. DOI: 10.1016/j.scient.2011.07.004.
- [234] H. Xi, T. Ying und H. Xingui, "An Intelligent Multifeature Statistical Approach for the Discrimination of Driving Conditions of a Hybrid Electric Vehicle", *Intelligent Transportation Systems, IEEE Transactions on*, Jg. 12, Nr. 2, S. 453–465, 2011, ISSN: 1524-9050. DOI: 10.1109/TITS.2010.2093129.
- [235] K. Brundell-Freij und E. Ericsson, "Influence of street characteristics, driver category and car performance on urban driving patterns", *Transportation Research Part D: Transport and Environment*, Jg. 10, Nr. 3, S. 213–229, 2005, ISSN: 13619209. DOI: 10.1016/j.trd.2005.01.001.
- [236] J. Kang, T. Ma, F. Ma und J. Huang, "Link-based emission model for eco routing", in *11th International Conference on ITS Telecommunications (ITST), 2011*, 2011, S. 207–212, ISBN: 978-1-61284-668-2. DOI: 10.1109/ITST.2011.6060054.
- [237] H. Yu, F. Tseng und R. McGee, "Driving pattern identification for EV range estimation", in *2012 IEEE International Electric Vehicle Conference*, IEEE, 2012, S. 1–7, ISBN: 978-1-4673-1561-6. DOI: 10.1109/IEVC.2012.6183207.
- [238] H. Rahimi-Eichi, P. B. Jeon, M.-Y. Chow und T.-J. Yeo, "Incorporating big data analysis in speed profile classification for range estimation", in *Proceedings, 2015 IEEE International Conference on Industrial Informatics (INDIN)*, Piscataway, NJ: IEEE, 2015, S. 1290–1295, ISBN: 978-1-4799-6649-3. DOI: 10.1109/INDIN.2015.7281921.
- [239] G. Agamennoni, J. Nieto, E. M. Nebot u. a., "A Bayesian approach for driving behavior inference", in *Intelligent Vehicles Symposium (IV), 2011 IEEE*, 2011, S. 595–600, ISBN: 978-1-4577-0891-6. DOI: 10.1109/IVS.2011.5940407.
- [240] S. Worrall, G. Agamennoni, J. Nieto und E. Nebot, "A Context-Based Approach to Vehicle Behavior Prediction", *Intelligent Transportation Systems Magazine, IEEE*, Jg. 4, Nr. 3, S. 32–44, 2012, ISSN: 1939-1390. DOI: 10.1109/MITS.2012.2203230.
- [241] D. A. Johnson und M. M. Trivedi, "Driving style recognition using a smartphone as a sensor platform", in *14th International IEEE Conference on Intelligent Transportation Systems*, 2011, S. 1609–1615, ISBN: 978-1-4577-2197-7. DOI: 10.1109/ITSC.2011.6083078.
- [242] M. Liebner, M. Baumann, F. Klanner und C. Stiller, "Driver intent inference at urban intersections using the intelligent driver model", in *Intelligent Vehicles Symposium (IV), 2012 IEEE*, 2012, S. 1162–1167, ISBN: 978-1-4673-2119-8. DOI: 10.1109/IVS.2012.6232131.

- [243] F. Gross, J. Jordan, F. Wening, F. Klanner und B. Schuller, "Route and Stopping Intent Prediction at Intersections From Car Fleet Data", *IEEE Transactions on Intelligent Vehicles*, Jg. 1, Nr. 2, S. 177–186, 2016, ISSN: 2379-8904. DOI: 10.1109/TIV.2016.2617625.
- [244] J. Lohrer und M. Lienkamp, "Building representative velocity profiles using FastDTW and spectral clustering", in *14th International Conference on ITS Telecommunications (ITST)*, IEEE, 2015, S. 45–49, ISBN: 978-1-4673-9382-9. DOI: 10.1109/ITST.2015.7377398.
- [245] E. Kreyszig, *Advanced engineering mathematics*. John Wiley & Sons, 2010, ISBN: 9780470458365.
- [246] W. Leutzbach, *Einführung in die Theorie des Verkehrsflusses*, Ser. Hochschultext. Berlin, Heidelberg: Springer Berlin Heidelberg, 1972, ISBN: 3540057242.
- [247] D. J. Berndt und J. Clifford, "Using Dynamic Time Warping to Find Patterns in Time Series", *AAAIWS'94 Proceedings of the 3rd International Conference on Knowledge Discovery and Data Mining*, S. 359–370, 1994.
- [248] X. Wang, A. Mueen, H. Ding, G. Trajcevski, P. Scheuermann und E. Keogh, "Experimental comparison of representation methods and distance measures for time series data", *Data Mining and Knowledge Discovery*, Jg. 26, Nr. 2, S. 275–309, 2013, ISSN: 1573-756X. DOI: 10.1007/s10618-012-0250-5.
- [249] R. Giusti und G. E. Batista, "An empirical comparison of dissimilarity measures for time series classification", in *Intelligent Systems (BRACIS), 2013 Brazilian Conference on*, 2013, S. 82–88, ISBN: 9781479912711. DOI: 10.1109/BRACIS.2013.22.
- [250] E. Keogh und C. A. Ratanamahatana, "Exact indexing of dynamic time warping", *Knowledge and Information Systems*, Jg. 7, Nr. 3, S. 358–386, 2005, ISSN: 0219-1377. DOI: 10.1007/s10115-004-0154-9.
- [251] S. Salvador und P. Chan, "Toward accurate dynamic time warping in linear time and space", *Intelligent Data Analysis*, Jg. 11, Nr. 5, S. 561–580, 2007, ISSN: 1088-467X.
- [252] E. J. Keogh und M. J. Pazzani, "Scaling up dynamic time warping for datamining applications", in *Proceedings of the sixth ACM SIGKDD international conference on Knowledge discovery and data mining*, 2000, S. 285–289, ISBN: 1-58113-233-6. DOI: 10.1145/347090.347153.
- [253] R. Xu und D. Wunsch, "Survey of clustering algorithms", *IEEE transactions on neural networks*, Jg. 16, Nr. 3, S. 645–678, 2005, ISSN: 1045-9227. DOI: 10.1109/TNN.2005.845141.
- [254] U. von Luxburg, "A tutorial on spectral clustering", *Statistics and computing*, Jg. 17, Nr. 4, S. 395–416, 2007, ISSN: 1573-1375. DOI: 10.1007/s11222-007-9033-z.
- [255] A. Y. Ng, M. I. Jordan, Y. Weiss u. a., "On spectral clustering: Analysis and an algorithm", in *NIPS'01 Proceedings of the 14th International Conference on Neural Information Processing Systems: Natural and Synthetic*, Bd. 14, S. 849–856.
- [256] L. Zelnik-Manor und P. Perona, "Self-tuning spectral clustering", in *Advances in neural information processing systems*, 2004, S. 1601–1608.
- [257] W.-Y. Chen, Y. Song, H. Bai, C.-J. Lin und E. Y. Chang, "Parallel spectral clustering in distributed systems", *IEEE transactions on pattern analysis and machine intelligence*, Jg. 33, Nr. 3, S. 568–586, 2011, ISSN: 0162-8828. DOI: 10.1109/TPAMI.2010.88.
- [258] B. F. J. Manly, *Multivariate statistical methods: A primer*. CRC Press, 2004, ISBN: 9781482285987.
- [259] K. P. Murphy, *Machine learning: A probabilistic perspective*. 2012, ISBN: 9780262306164.
- [260] E. Acuña und C. Rodriguez, "The Treatment of Missing Values and its Effect on Classifier Accuracy", in *Classification, clustering, and data mining applications*, Ser. Studies in classification, data analysis, and knowledge organization, D. Banks, Hrsg., Berlin: Springer, 2004, S. 639–647, ISBN: 978-3-540-22014-5. DOI: 10.1007/978-3-642-17103-1_60.

- [261] D. Wentura und M. Pospeschill, *Multivariate Datenanalyse: Eine kompakte Einführung*, Ser. Lehrbuch. Wiesbaden: Springer, 2015, ISBN: 978-3-531-17118-0.
- [262] K. Backhaus, B. Erichson, W. Plinke und R. Weiber, *Multivariate Analysemethoden: Eine anwendungsorientierte Einführung*, 14., überarb. u. aktualisierte Aufl. 2016. Berlin: Springer, 2016, ISBN: 978-3-662-46075-7.
- [263] D. G. Kleinbaum und M. Klein, *Logistic regression: A self-learning text*, 3rd ed., Ser. Statistics for biology and health. New York: Springer, 2010, ISBN: 978-1-4419-1741-6.
- [264] S. A. Czepiel, *Maximum likelihood estimation of logistic regression models: Theory and implementation*, 2002. Adresse: <https://czep.net/stat/mlelr.pdf> (besucht am 01.09.2017).
- [265] J. Behnke, *Logistische Regressionsanalyse*. Wiesbaden: Springer Fachmedien Wiesbaden, 2015, ISBN: 978-3-658-05081-8. DOI: 10.1007/978-3-658-05082-5.
- [266] N. Döring und J. Bortz, *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften*. Berlin, Heidelberg: Springer, 2006, ISBN: 978-3-540-33305-0.
- [267] E. Käfer, "Situationsklassifikation und Bewegungsprognose in Verkehrssituationen mit mehreren Fahrzeugen", Dissertation, Universität Bielefeld, Bielefeld, 2013.
- [268] B. Heißing, M. Ersoy und S. Gies, *Fahrwerkhandbuch*. Wiesbaden: Springer Fachmedien Wiesbaden, 2013, ISBN: 978-3-658-01991-4. DOI: 10.1007/978-3-658-01992-1.
- [269] E. Mehrjerdian, "Potenziale einer Lenkmomentenempfehlung unter fahrdynamischen Gesichtspunkten", Dissertation, Technische Universität Dortmund, 2015.
- [270] D. Ammon, "Vehicle system dynamics challenges on the way to autonomous driving", 4. *Internationales Münchner Fahrwerk-Symposium*, 2013.
- [271] J. M. Giron-Sierra, *Digital signal processing with Matlab examples: Signals and communication technology*. Singapore: Springer, 2017, ISBN: 9789811025396.
- [272] S. Matz, "Nutzerorientierte Fahrzeugkonzeptoptimierung in einer multimodalen Verkehrsumgebung", Dissertation, Technische Universität München, München, 2015.
- [273] J. Lohrer, M. Förth und M. Lienkamp, "A data-driven predictive energy management strategy for plug-in hybrid vehicles", in *2017 International Conference on Mechanical, System and Control Engineering (ICMSC)*, 2017, S. 297–304, ISBN: 978-1-5090-6530-1. DOI: 10.1109/ICMSC.2017.7959490.
- [274] J. Gonder und T. Markel, "Energy Management Strategies for Plug-In Hybrid Electric Vehicles", *SAE 2007 World Congress*, 2007. DOI: 10.4271/2007-01-0290.
- [275] S. G. Wirasingha und A. Emadi, "Classification and Review of Control Strategies for Plug-In Hybrid Electric Vehicles", *IEEE Transactions on Vehicular Technology Vol. 60, No. 1*, 2011, ISSN: 0018-9545. DOI: 10.1109/TVT.2010.2090178.
- [276] F. R. Salmasi, "Control Strategies for Hybrid Electric Vehicles: Evolution, Classification, Comparison, and Future Trends", *IEEE Transactions on Vehicular Technology, Vol. 56, No. 5*, 2007, ISSN: 0018-9545. DOI: 10.1109/TVT.2007.899933.
- [277] A. Panday und H. O. Bansal, "A Review of Optimal Energy Management Strategies for Hybrid Electric Vehicle", *International Journal of Vehicular Technology*, 2014. DOI: 10.1155/2014/160510.
- [278] C. Zhang und A. Vahidi, "Real-Time Optimal Control of Plug-in Hybrid Vehicles with Trip Preview", *American Control Conference*, 2010, ISSN: 2378-5861. DOI: 10.1109/ACC.2010.5531308.
- [280] Q. Gong, Y. Li und Z.-R. Peng, "Trip-Based Optimal Power Management of Plug-in Hybrid Electric Vehicles", *Transactions on Vehicular Technology Vol. 57 No. 6*, 2008, ISSN: 1939-9359. DOI: 10.1109/TVT.2008.921622.

- [281] M. Struwe, *Variational methods: Applications to nonlinear partial differential equations and Hamiltonian systems*, 4. ed., Ser. Ergebnisse der Mathematik und ihrer Grenzgebiete. Berlin: Springer, 2010, Bd. Folge 3, 34, ISBN: 9783642093296.
- [282] S. Onori, L. Serrao und G. Rizzoni, “Adaptive Equivalent Consumption Minimization Strategy for Hybrid Electric Vehicles”, in *ASME 2010 Dynamic Systems and Control Conference*, 2010, S. 499–505, ISBN: 978-0-7918-3884-6. DOI: 10.1115/DSCC2010-4211.
- [283] M. Winkler, S. Geulen, M. Josevski, M. Tegethoff, D. Abel und B. Vöcking, “Online Parameter Tuning Methods for Adaptive ECMS Control Strategies in Hybrid Electric Vehicles”, *FISITA World Automotive Congress*, 2014.
- [284] P. Naghshtabrizi, J. Kristinsson, H. Yu und R. McGee, “Distance Until Charge prediction and fuel economy impact for Plug-in Hybrid Vehicles”, in *American Control Conference (ACC), 2011*, 2011, S. 2739–2744, ISBN: 978-1-4577-0081-1. DOI: 10.1109/ACC.2011.5991117.
- [285] B. Bader, O. Torres, J. A. Ortega, G. Lux und J. L. Romeral, “Reduction of the prediction horizon of predictive energy management for a plug-in HEV in hilly terrain”, *WIT Transactions on the Built Environment*, Nr. 130, S. 397–408, 2013, ISSN: 1743-3509. DOI: 10.2495/UT130311.
- [286] Mapbox, Hrsg., *Mapbox Lizenzvereinbarung*. Adresse: <https://www.mapbox.com/tos/> (besucht am 01.03.2017).
- [287] P. Feig, A. König, K. Gschwendtner, J. Lohrer, J. Schatz und M. Lienkamp, “Deriving system parameters of a 360 degree low speed autonomous emergency braking driver assistance system for parking and maneuvering based on naturalistic driving studies”, in *2017 IEEE International Conference on Vehicular Electronics and Safety (ICVES)*, 2017, S. 156–161, ISBN: 1509056785. DOI: 10.1109/ICVES.2017.7991918.

Studentische Arbeiten

Während der Anfertigung dieser Dissertation entstanden unter wesentlicher wissenschaftlicher, fachlicher und inhaltlicher Anleitung des Autors die nachfolgend aufgeführten studentischen Arbeiten, in denen verschiedene Fragestellungen zum Thema untersucht wurden. Die entstandenen Ergebnisse sind teilweise in die vorliegende Arbeit eingeflossen. Vielen Dank an die Bearbeiter für Ihre Unterstützung im Rahmen dieses Forschungsprojekts.

- [279] M. Förth, "Entwicklung einer Betriebsstrategie für ein PHEV auf Basis einer Fahrprofilprädiktion", Masterarbeit, Technische Universität München, 2016.
- [288] X. Su, "Analyse der Mobilitätsdaten von Fahrzeugschwärmen zur Verkehrslagebestimmung", Semesterarbeit, Technische Universität München, 2014.
- [289] R. Wachter, "Entwicklung einer skalierbaren Quelle-Ziel-Matrix auf Basis von Mobilitätsdaten", Bachelorarbeit, Technische Universität München, 2015.
- [290] H.-P. Huber, "Potentialanalyse von innovativen Diensten für vernetzte Fahrzeuge", Masterarbeit, Technische Universität München, 2015.
- [291] B. Lohse, "Analyse und Klassifizierung von Fahrprofilen auf unterschiedlichen Streckensegmenten", Bachelorarbeit, Technische Universität München, 2015.
- [292] C. Zhang, "Entwicklung eines kartenbasierten Prädiktionsmodells für das Fahrverhalten von PKW im urbanen Raum", Masterarbeit, Technische Universität München, 2015.
- [293] C. Zhang, "Analyse von Einflussfaktoren auf das Geschwindigkeitsverhalten im Stadtverkehr", Semesterarbeit, Technische Universität München, 2015.
- [294] W. Christl, "Entwicklung eines Trip-Prädiktionsmodells auf Basis von Mobilitätsdaten", Bachelorarbeit, Technische Universität München, 2016.
- [295] J. Bostner, "Entwicklung einer Erweiterung von digitalen Karten um Energieverbrauchsprofile für eine Flottensimulation", Semesterarbeit, Technische Universität München, 2016.
- [296] J. Yu, "Development of a predictive thermal management for electrical machine overload optimization", Masterarbeit, Technische Universität München, 2016.
- [297] A. Ranjan, "Development of a map-based prediction model for driving behaviour of passenger cars in an urban environment", Final Year Project (Imperial College London), Technische Universität München, 2016.
- [298] T. Scholz, "Prädiktion von fahrdynamischen und energetischen Größen für ein Elektrofahrzeug", Masterarbeit, Technische Universität München, 2016.
- [299] D. Ergene, "Sensitivitätsanalyse zur Erstellung repräsentativer Fahrprofile", Bachelorarbeit, Technische Universität München, 2017.

Eigene Veröffentlichungen

Während der Anfertigung dieser Dissertation sind Veröffentlichungen entstanden, in denen Teilaspekte dieser Arbeit vorgestellt wurden.

- [5] H.-P. Huber, J. Lohrer, J. Hutzschenreuther und M. Lienkamp, “Potenzialanalyse von innovativen Diensten für vernetzte Fahrzeuge”, *ZfAW, Zeitschrift für die gesamte Wertschöpfungskette der Automobilwirtschaft*, Nr. 3, S. 12–22, 2015.
- [52] T. Tang, D. Soto-Setzke, C. Kohl, T. Köhn, J. Lohrer und J. Betz, “EE-Architektur für mobile Dienste”, *ATZextra*, Jg. 19, Nr. 14, S. 40–45, 2014, ISSN: 2195-1454. DOI: 10.1365/s35778-014-1356-8.
- [97] J. Lohrer und M. Lienkamp, “An approach for predicting vehicle velocity in combination with driver turns”, *Automotive and Engine Technology*, Nr. 1, S. 27–33, 2016, ISSN: 2365-5135. DOI: 10.1007/s41104-016-0012-2.
- [100] M. Wittmann, J. Lohrer, J. Betz, B. Jäger, M. Kugler, M. Klöppel, A. Waclaw, M. Hann und M. Lienkamp, “A Holistic Framework for Acquisition, Processing and Evaluation of Vehicle Fleet Test Data”, *IEEE 20th International Conference on Intelligent Transportation Systems (ITSC)*, 2017.
- [244] J. Lohrer und M. Lienkamp, “Building representative velocity profiles using FastDTW and spectral clustering”, in *14th International Conference on ITS Telecommunications (ITST)*, IEEE, 2015, S. 45–49, ISBN: 978-1-4673-9382-9. DOI: 10.1109/ITST.2015.7377398.
- [273] J. Lohrer, M. Förth und M. Lienkamp, “A data-driven predictive energy management strategy for plug-in hybrid vehicles”, in *2017 International Conference on Mechanical, System and Control Engineering (ICMSC)*, 2017, S. 297–304, ISBN: 978-1-5090-6530-1. DOI: 10.1109/ICMSC.2017.7959490.
- [287] P. Feig, A. König, K. Gschwendtner, J. Lohrer, J. Schatz und M. Lienkamp, “Deriving system parameters of a 360 degree low speed autonomous emergency braking driver assistance system for parking and maneuvering based on naturalistic driving studies”, in *2017 IEEE International Conference on Vehicular Electronics and Safety (ICVES)*, 2017, S. 156–161, ISBN: 1509056785. DOI: 10.1109/ICVES.2017.7991918.

Anhang

A Literaturübersicht zur Trip-Prädiktion

Tabelle A.1 : Literaturübersicht Trip-Prädiktion

Quelle	Jahr	Typ	Methode	Datenbasis	Abstaktion	Nutzer
[149]	2006	Z	Satz von Bayes	4 300 Fahrten	Raster	K
[19]P	2008	Z	Kriterien	-	Adresse	I
[150]	2010	Z	HMM	645 Fahrten	Punkte	I
[151]	2012	Z	HMM + Zeit	-	POI	I
[152]	2013	Z	Decision Tree wP	269 Fahrten	OD-Matrix	I
[153]	2014	Z	Markov Modell	> 1,7 Mio. Fahrten	Punkte	I+K
[154]	2014	Z	DBSCAN + VOMM	> 17 000 Fahrten	POI	K
[156]	2015	Z	BRNN	> 1,7 Mio. Fahrten	Punkte	K
[157]	2016	Z	CNN	> 1,7 Mio. Fahrten	Raster	K
[155]	2016	Z	RNN	Gowalla	Punkte	I
[158]	2006	R	HMM + Zeit	46 Fahrten	Segmente	I
[160]	2008	R	Kriterien	-	Segmente	I
[162]	2008	R	Hausdorff + Hier. Cluster	27 000 Fahrten	-	I
[163]	2013	R	Hausdorff + Hier. Cluster	27 000 Fahrten	Segmente	I
[164]	2014	R	HMM	269 Fahrten	Segmente	I
[174]	2003	T	Fahrwahrscheinlichkeit	-	Segmente	I
[173]	2004	T	Bewegungsfunktion	-	-	I
[165]	2006	T	-	-	Segmente	I
[175]	2008	T	Markov Modell	27 000 Fahrten	Segmente	I/K
[167]P	2009	T	Straßenklassen	-	Segmente	I
[176]	2009	T	VOMM + Zeit	ShanghaiGrid	Segmente	K
[177]	2009	T	Decision Tree	4 500 Fahrten	Raster	K
[178]	2010	T	Maximum Likelihood / Greedy	> 100 000 Fahrten	Segmente	K
[168]P	2011	T	Straßenklassen + Tabelle	-	Segmente	I
[179]	2011	T	Präfixbaum	17 000 Fahrten	Segmente	K
[169]P	2011	T	Dyn. Kartenattribute	-	Segmente	I
[170]P	2012	T	Dyn. Umgebungsdaten	-	Segmente	I
[180]	2012	T	Gewichtetes Markov Modell	> 250 000 Fahrten	Raster + POI	K
[181]	2013	T	DBScan + Suffixbaum	Nokia	Raster	I/K
[171]P	2013	T	Fahrwahrscheinlichkeit	-	Segmente	I+K
[182]	2013	T	Dyn. Referenztrajektorien	SGP Taxi	Raster	K
[91]	2015	T	Straßenklassen	-	Segmente	I
[183]	2016	T	Backtracking	> 18 000 Fahrten	Segmente	K
[172]P	2016	T	Wahrscheinlichkeit + Lernen	-	Segmente	I+K
[184]	2016	T	Deep Residual Learning	TaxiBJ & BikeNYC	Raster	K
[185]	2017	T	Distanzmaß	> 326 000 Fahrten	-	I+K

P = Patent, Z = Ziel, R = Route, T = Trip, K = Kollektiv, I = Individuell

B Literaturübersicht zur Fahrprofilprädiktion

Tabelle B.1 : Literaturübersicht Fahrprofilprädiktion

Quelle	Jahr	Typ	Methode	Datenbasis	H	Ziel
[194]	2000	M	Physikalisch	-	K+L	F
[195]	2014	M	IDM	-	L	P
[196]	2012	M	IDM	-	K	F
[197]	2006	M	Physikalisch	-	L	P
[198]	2009	M	Physikalisch + Verkehr	-	L	P
[199]	2011	M	Physikalisch + Verkehr	-	L	P
[200]	2012	M	Physikalisch + Verkehr + Stopps	-	L	P
[201]	2013	M	PELOPS + C2C + C2I	-	L	P
[204]	2014	M	Physikalisch + Verkehr + Fahrer	-	L	P
[213]	2014	M	Kognitive Mechanismen	-	L	F
[32]	2009	M	Regressionsmodelle	Einzelne Straßen	K	F
[205]	2009	M	Physikalisch	-	L	P
[206]	2009	M	Modellprädiktive Regelung	-	L	P
[207]	2012	M	DP	-	L	P
[208]	2012	M	DP	-	L	P
[210]	2013	M	Iterative Dynamic Programming	-	L	P
[211]	2012	M	DP + Back-End	-	L	P
[212]	2012	M	Optimal Control	-	L	P
[232]	2007	E	Klassifizierung	40 Manöver	-	V
[233]	2011	E	k-means	Fixe Route > 6 Monate	-	V
[234]	2011	E	SVM	unbekannt	K	V
[28]	2001	E	PCA	5 Fzg, 18 945 km	K	P
[235]	2005	E	PCA	5 Fzg, 18 945 km	K	P
[236]	2011	E	Entscheidungsbaum	unbekannt	K	P
[237]	2012	E	Fuzzy c-means	unbekannt	L	P
[24]	2014	E	Manuelle Klassifikation	>1 000 Manöver	K	P
[238]	2015	E	Markov Ketten	>23 000 km	L	P
[95]	2017	E	SVM	>450 000 km	L	P
[239]	2011	E	Probabilistisches Modell	unbekannt	L	F
[241]	2011	E	DTW	>200 Manöver	K	F
[242]	2012	E	IDM	>200 Manöver	K	F
[243]	2016	E	Random Forest	>50 000 Manöver	K	F
[214]	2003	D	Neuronales Netz	>200 Manöver	K	F
[215]	2008	D	SVD + Nachbar	131 Manöver	K	F
[216]	2012	D	SUMO	-	K	F
[219]	2014	D	RBF-Neuronales Netz	Fahrzyklen	K	F
[221]	2015	D	Autoregressive-Model	Fixe Route 40 Fahrten	K	F
[222]	2016	D	Neuronales Netz	>40 000 km	K	F
[29]	2016	D	DTW + Hierarchisches Clustern	Fixe Route 10 Fahrten	L	F
[227]	2011	D	Selbstlernender Ansatz	Fixe Route 85 Fahrten	L	F
[228]	2013	D	Markov Ketten	> 1 000 Fahrten	L	F
[229]	2013	D	Mehrlagiges Perceptron	Fixe Route 92 Fahrten	L	F
[231]	2015	D	Stacked Autoencoder	Fixe Route > 700 Fahrten	L	F

H = Horizont M = Mathematisches Modell, E = Eigenschaften, D = Datenbasiert, K = Kurzfristig, L = Langfristig, F = Fahrverhalten, P = Energie, V = Verkehrsfluss

C Übersicht zu Rechenzeiten

Für die Ermittlung der Rechenzeiten wird der Algorithmus auf einer virtuellen Maschine mit Windows 7 (64 Bit), Intel X5650 (2,67 GHz x10), 64 GB und Matlab R2017b durchgeführt.

Trip-Prädiktion

Abb. C.1 veranschaulicht die benötigte Rechenzeit für die Erstellung eines Simulationsmodells für die Trip-Prädiktion. Die Zeit ist abhängig von der Anzahl der Segmente in und um das Gebiet, sowie der Anzahl der verwendeten Fahrten. Für diese Auswertung wurden 1200 Modelle herangezogen.

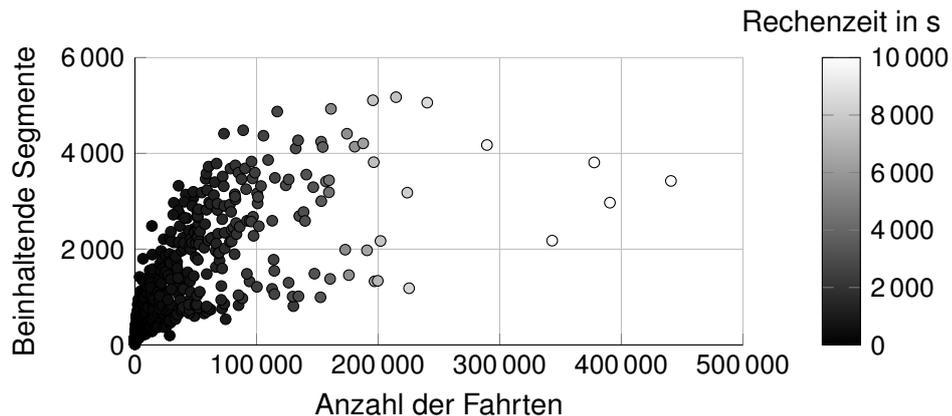


Abbildung C.1 : Zeit für die Erstellung eines Trip-Prädiktion Modells

Die mittlere Rechenzeit für die Prädiktion des nächsten Streckenabschnitts beträgt 1,33 s. Für diese Auswertung wurden 200 Modelle herangezogen.

Fahrprofilprädiktion

Abb. C.2 veranschaulicht die benötigte Rechenzeit für die Erstellung eines Simulationsmodells für die Fahrprofilprädiktion. Die Zeit ist abhängig von der Anzahl der Regressionskoeffizienten, sowie der Anzahl der verwendeten Fahrten. Für diese Auswertung wurden 75 Modelle herangezogen.

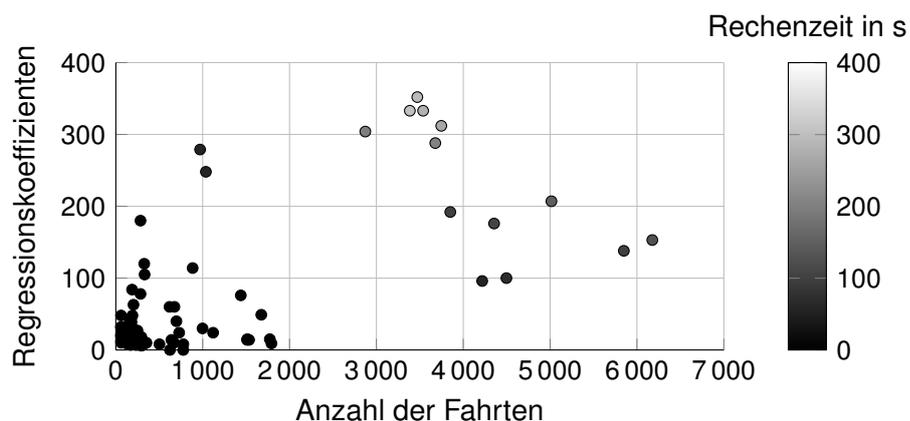


Abbildung C.2 : Zeit für die Erstellung eines Fahrprofilprädiktion Modells

Die Rechenzeit für die Prädiktion des Fahrprofils auf dem nächsten Streckenabschnitt ist nahezu unabhängig von dem erstellten Modell. Es handelt sich um einfache mathematische Operationen. Die mittlere Rechenzeit beträgt 0,0012 s.