

Entwicklung einer Anwendung zur Nachfrageprognose und Reallokation von E-Rollern

Masterarbeit

zur Erlangung des akademischen Grades „Master of Science (M. Sc.)“ im
Studiengang Wirtschaftsingenieur der Fakultät für Elektrotechnik und Informatik,
Fakultät für Maschinenbau und der Wirtschaftswissenschaftlichen Fakultät der
Leibniz Universität Hannover

vorgelegt von

Name: Wielert

Vorname: Henrik



Prüfer: Prof. Dr. Michael H. Breitner

Hannover, den 28.03.2020

Inhaltsverzeichnis

Abbildungsverzeichnis	II
Tabellenverzeichnis	III
Abkürzungsverzeichnis	IV
1. Einleitung	1
2. Grundlagen und Definitionen	2
2.1 Technische Ausgestaltung der E-Roller	2
2.2 Definitionen	4
3. Literaturübersicht: Reallokationsstrategien	7
4. Algorithmen zur Clusteranalyse	12
4.1 K-MEANS	14
4.2 DBSCAN	16
4.3 HDBSCAN	17
5. Erstellung einer Web-Anwendung zur Nachfrageprognose von E-Rollern	23
5.1 Anforderungen an die Applikation	23
5.2 Struktur der Anwendung	25
5.2.1 Aufbau des Programms	25
5.2.2 Verwendete Python-Libraries	28
5.3 Zusammensetzung der Datenbasis	30
5.4 Machine Learning	33
5.4.1 Aufruf der Clusteranalyse	33
5.4.2 Erstellung der Regressionen	36
5.5 Grafische Oberfläche der Anwendung	40
5.5.1 Erstellung einer Interaktiven Karte	40
5.5.2 Visualisierung der statistischen Werte	45
5.5.3 Erstellung einer Web-Anwendung	51
5.6 Szenarioanalyse zur Überprüfung der Prognose	53
6. Reallokationsstrategie für das Sharing von E-Rollern	57

6.1	Anforderungen an eine Reallokationsstrategie	58
6.2	Theoretisches Reallokationsmodell für E-Roller	60
7.	Diskussion.....	63
8.	Limitation.....	66
9.	Fazit und Handlungsempfehlungen	67
	Literaturverzeichnis	69
	Anhang A: Quellcode.....	I
	Anhang B: Ausschnitte der Datentabellen	XXXIX
	Anhang C: Optimierungsmodell zur Reallokation	XLI
	Ehrenwörtliche Erklärung.....	

Abbildungsverzeichnis

Abbildung 1	E-Roller der Provider in Deutschland v.l.n.r. Lime, Tier, Circ und Voi.....	3
Abbildung 2	Verschiedene Strategien zur Reallokation	10
Abbildung 3	Verteilung der Abfahrten von E-Rollern über einen Zeitraum von 5 Tagen	12
Abbildung 4	Erzeugung dichtebasierter Cluster mit DBSCAN.....	13
Abbildung 5	Iterationsbeispiel K-Means	15
Abbildung 6	Veranschaulichung der Definitionen bei DBSCAN	17
Abbildung 7	Baumdiagramm der Clustererzeugung mit HDBSCAN.....	19
Abbildung 8	Clusterzusammensetzung in Abhängigkeit des Dichte-Levels	20
Abbildung 9	Gliederung der Datenbasis in Cluster bei HDBSCAN.....	21
Abbildung 10	Struktur der programmierten Applikation	27
Abbildung 11	E-Roller-Datensatz	32
Abbildung 12	Gegenüberstellung verschiedener Parametereinstellungen bei HDBSCAN (v.l.n.r. 1,5%, 1%, 0,5%)	35
Abbildung 13	Zusammenführung der Datensätze zur Regression	37
Abbildung 14	Ablauf der Polynom-Regression	38
Abbildung 15	Layer der Cluster der Abfahrten eines Wochentags	42

1. Einleitung

Die Weltbevölkerung nimmt zu und im gleichen Zuge zieht es die Menschen in die Städte. Mit der zunehmenden Urbanisierung in Deutschland steigt die Einwohnerzahl der Städte immer stärker an. Im Zuge dieses Wachstums steigen die Anforderungen an die Mobilitätskonzepte innerhalb der Städte. Aufgrund des Klimawandels und lokaler Luftverschmutzung werden in vielen Städten neue Verkehrskonzepte benötigt. Dabei müssen die Konzepte eine Reduzierung der Emissionen von CO₂ und NO_x mit sich führen. Bei der Ausgestaltung dieser Verkehrskonzepte muss eine Kombination aus verschiedenen Verkehrsmitteln berücksichtigt werden. Vor diesem Hintergrund werden öffentliche Verkehrsmittel einen stärkeren Fokus in der Planung erfahren. Jedoch werden weiterhin individuelle Verkehrsmittel, wie Fahrräder, E-Autos und auch E-Roller benötigt. Allerdings wird ein Großteil dieser Verkehrsmittel über Sharing Angebote genutzt werden.

Bereits im Sommer 2019 mit dem Inkrafttreten der Elektrokleinstfahrzeuge-Verordnung (eKFV) etablierten einige Provider ihre Flotten an E-Rollern in deutschen Innenstädten. Da Sharing Systeme für Fahrräder bereits weit verbreitet waren, erfolgte der Roll-Out geplant und zügig.

Mein Interesse ist genau an dieser Stelle geweckt worden. Ausgehend von einer zweigeteilten Meinung in der Bevölkerung begann ich zu hinterfragen, wie eine Verbesserung der Zuordnung von Nutzer zu E-Roller in einem Abbau von gesellschaftlicher Kritik und einer Effizienzsteigerung der Sharing Systeme münden kann. Um den Status Quo zu verbessern muss eine Umverteilung über das Sharing-Netzwerk stattfinden. Diese Umverteilung wird Rebalancing oder Reallokation genannt. Eine effizient koordinierte Reallokation kann die Auslastung des Sharing-Netzwerkes steigern und eine effiziente Allokation von Nutzern zu E-Rollern ermöglichen.

Der Aufbau einer effizienten Reallokationsstrategie erfordert dabei eine genaue Kenntnis des Nachfrageverhaltens der Nutzer, sowie Aufschluss über deren Präferenzen. Daher ist es für Provider zwingend notwendig die Daten des operativen Betriebs bestmöglich aufzunehmen und auszuwerten.

Die Forschungsfragen dieser Arbeit richten sich genau an dieser Thematik aus:

- 1. Kann eine Nachfrageprognose für die Nutzung von E-Rollern unter Einbeziehung von Wetterdaten erstellt werden?**
- 2. Kann die erstellte Nachfrageprognose als Entscheidungsunterstützungswerkzeug für die Reallokation dienen?**
- 3. Lässt sich auf Basis einer Nachfrageprognose ein Reallokationsmodell in einem Free Floating Sharing System für E-Roller effizient gestalten?**

Diese Arbeit befasst sich mit der Entwicklung einer Web-Applikation, welche auf Basis einer Nachfrageprognose wichtige Erkenntnisse über die Reallokation von E-Rollern identifiziert. Der folgende Abschnitt befasst sich zunächst mit der groben Vorstellung der Ausgestaltung eines E-Rollers. Weiterhin werden wichtige Begriffsdefinitionen aus dem Kontext dieser Arbeit vorgenommen. Das dritte Kapitel führt eine Übersicht der Reallokationsthematik im Kontext von Sharing Systemen an. Die Erläuterung und mathematische Herleitung unterschiedlicher Ansätze von Algorithmen zur Clusteranalyse erfolgt im vierten Abschnitt. Anschließend wird in Kapitel 5 die entwickelte Web-Applikation detailliert in ihrer Funktionsweise dargestellt. Dabei wird weniger eine Vorstellung des Quellcodes vorgenommen, als anschaulich hergeleitet, wie bei der Entwicklung vorgegangen wurde. Abschließend erfolgt zudem in diesem Kapitel die Überprüfung der erzeugten Prognose anhand eines Szenarios. Im sechsten Abschnitt erfolgt daraufhin eine Analyse der Anforderungen an eine Reallokation und die theoretische Entwicklung eines Optimierungsmodells zur Abbildung einer Reallokationsstrategie. Weiterhin folgt die Diskussion der Ergebnisse der Arbeit und der Forschungsfragen. Daraufhin werden Limitationen bei der Bearbeitung erläutert. Abschließend wird das Fazit vorgenommen und Handlungsempfehlungen werden ausgesprochen.

2. Grundlagen und Definitionen

In diesem Kapitel wird zunächst grundlegend auf die Anforderungen an die technische Ausgestaltung eines E-Rollers eingegangen. Dabei erfolgt ausschließlich die Betrachtung von elektrisch betriebenen Tretrollern. In einem zweiten Abschnitt werden Definitionen verschiedener relevanter Begriffe aus dem Kontext dieser Arbeit vorgenommen.

2.1 Technische Ausgestaltung der E-Roller

Es existieren verschieden ausgestattete Varianten von E-Rollern auf dem Markt. Jedoch unterliegen alle E-Roller die auf deutschen Straßen fahren der eKFV. Die rechtlich verbindliche Verordnung regelt die genauen Anforderungen an die technische Ausstattung der E-Roller und ist am 15. Juni 2019 in Kraft getreten.

Der Gegenstand der Betrachtung sind dabei Kraftfahrzeuge mit einem elektrischen Antrieb und einer Höchstgeschwindigkeit von 6 – 20 km/h (§1 Absatz 1 eKFV). Die maximale Leistung eines E-Rollers ohne Selbststabilisierung darf 500 W nicht überschreiten. Wenn zur Selbststabilisierung mindestens 60% der Leistung benötigt werden, gilt 1.400 W als maximal zulässige Leistung (§ 1 Absatz 1 Satz 3 eKFV). E-Roller in Sharing Systemen sind in der Regel mit einer solchen Stabilisierungstechnik ausgestattet. Ein E-Roller darf eine Höhe von 140 cm, eine Breite von 70 cm sowie eine Länge von 200 cm nicht überschreiten (§ 1 Absatz 1 Satz 4 eKFV). Darüber hinaus ist

Aufgrund fehlender Datensätze aus den Sommermonaten fehlen unter Umständen wichtige Informationen über das Nutzerverhalten aus diesem Zeitraum. Da die ersten verfügbaren Daten aus dem September 2019 stammen, können somit wenig Schlüsse auf das Fahrtverhalten der Nutzer bei hohen Temperaturen und langen Tagen getroffen werden. Somit sollten weitere Forschungen in Sommermonaten stattfinden.

Der Betrachtungsraum dieser Arbeit ist die Stadt Hannover. Hannover bietet günstige Voraussetzungen für E-Roller Sharing, da die Stadt nur geringe Höhenunterschiede aufweist und eine vergleichsweise gute Infrastruktur für Radfahrer bietet. Eine Übertragbarkeit der Analysemethode wäre gewährleistet, jedoch können Städte mit unterschiedlichen Rahmenbedingungen gänzlich unterschiedliche Anforderungen an die Reallokation aufweisen.

Im Rahmen dieser Arbeit stehen lediglich die Daten eines Providers von E-Rollern in einem FFSS zur Verfügung. Daher kann aus der Datenbasis keine Aussage über den Wettbewerb für Sharing-Systeme mit E-Rollern getroffen werden. Es empfiehlt sich die Daten weiterer Provider in die Betrachtung einzubeziehen, um ein umfassendes Bild über die wirkliche Verteilung der Nachfrageorte zu erhalten.

9. Fazit und Handlungsempfehlungen

Die Aufnahme und Analyse der in einem FFSS erzeugten Daten ist zwingend notwendig für den Provider, da die Reallokation maßgeblich zum Erfolg des Unternehmens beiträgt. Auf die operative Reallokation wirken viele externe Faktoren maßgeblich ein. Vor diesem Hintergrund müssen Daten sowohl über interne Datenquellen als auch über externe Datenquellen bezogen werden. Im Kontext von FFSS sind vor allem Wetterdaten sowie geographische Daten von hoher Bedeutung. Auf der Basis der resultierenden Daten können in der Folge Analysen des Sharing-Netzwerkes durchgeführt werden.

In dieser Arbeit wird ein Data Mining Ansatz in Kombination mit Machine Learning Techniken vorgestellt. Dabei erfolgt die Untersuchung der Fahrtdaten von E-Rollern in Kombination mit Wetterdaten. Die durchgeführte Clusteranalyse erzeugt zeitlich stabile, dichtebasierte Cluster, die Auswertungen über Ballungszentren eines Sharing-Netzwerkes erlauben. Aufbauend auf dieser Datenanalyse können Prognosen über das Verhalten der Nutzer innerhalb des Sharing-Netzwerkes getroffen werden. Dies haben die Auswertungen in Kapitel 5.6 bewiesen.

Allerdings basieren die Prognosen auf Fahrtdaten anstelle von Nachfragedaten. Somit kann keine direkte Aussage über eine Nachfrageprognose getroffen werden. Wird der Auswertung jedoch die Annahme zugrunde gelegt, dass die Nachfrage und das Fahtaufkommen proportional zu einander stehen, so erlaubt die Prognose des Fahrt-

volumens eine bedingte Prognose des Nachfragevolumens. Folglich können erzeugten Prognosen als Entscheidungsunterstützungswerkzeug genutzt werden.

Da E-Roller in FFSS erst seit kurzer Zeit in den Städten betrieben werden, existieren derzeit nur wenige Veröffentlichungen zu dieser Thematik. Insbesondere im Bereich der Reallokation der E-Roller über das Netzwerk existieren bisher keine Veröffentlichungen. Da auch providerseitig keine Informationen über die Reallokationsstrategie herausgegeben werden, kann somit keine Aussage über die Effizienz der unterschiedlichen Strategien erfolgen. Das im Rahmen dieser Arbeit entwickelte Optimierungsmodell zur Reallokation stellt eine theoretische Herleitung aus Kennzahlen eines FFSS dar. Dabei wird jedoch kein Anspruch auf die Anwendbarkeit des Modells postuliert.

Aufgrund der Neuartigkeit des Sharing von E-Rollern, erfolgt ein hoher Forschungsaufwand zu jeglichen Themen in Bezug auf E-Roller. Derzeitige Veröffentlichungen fokussieren besonders die Analyse von Nutzergruppen. Allerdings werden auch technische Weiterentwicklungen der E-Roller selbst, aber auch der Ladeinfrastruktur, erforscht.

In der Zukunft werden E-Roller ein fester Bestandteil des städtischen Lebens. Begünstigt wird dies durch die Verdrängung von PKW aus den Innenstädten. Dabei werden sich die Provider in einem wachsenden Wettbewerb befinden. Um die Marktposition zu sichern oder auszubauen müssen die Provider die Daten ihres Sharing-Netzwerkes effizient zu einer Kostenreduktion im Bereich der Reallokation nutzen.

Infolge dessen besteht dringender Forschungsbedarf an Reallokationsstrategien und -modellen. In Bezug auf die Reallokationsstrategie können besonders Case-Studies über verschiedene Varianten und Mischformen der Reallokation wichtige Erkenntnisse über die Anforderungen bereithalten. Diese können anschließend genutzt werden, um eine effiziente Ausgestaltung der Reallokation zu untersuchen.

Zwar existieren bereits erste Forschungen darüber, welche Verkehrsmittel durch den E-Roller anteilig substituiert werden, jedoch sollte auch weitere Forschung in diese Richtung stattfinden. In diesem Kontext können besonders Case-Studies in nahezu PKW-freien Innenstädten wie Brüssel wesentliche Erkenntnisse bereithalten.

Darüber hinaus sollten weitere Forschungsbestrebungen in der Einbeziehung externer Daten bestehen. In diesem Kontext sind besonders sozio-demographische Daten interessant, mit dem Fokus zukünftige Entwicklungen der Märkte zu evaluieren. Zudem bestünde die Möglichkeit über Nutzeranalysen einen Aufschluss über weitere mögliche Nutzergruppen zu erhalten.