

Berechnung populärer Routen

Calculating Popular Routes

Volker Klingspor

Fachbereich Wirtschaft, Hochschule Bochum · volker.klingspor@hs-bochum.de

Zusammenfassung: Obwohl Routenplaner mit unterschiedlichen Zielsetzungen in großer Zahl bereits existieren, mangelt es an Planern, die Zweiradfahrer zielgerichtet über schöne Strecken führen. Zusätzlich zu Planern, die Eigenschaften der Strecke für die Bewertung von Wegen verwenden, basiert die hier eingeführte Bewertung auf der Häufigkeit, mit der Strecken befahren werden. Dazu wird mithilfe öffentlicher Tracksammlungen berechnet, welche Wege populär sind und diese Wege werden in der Routenplanung bevorzugt. Es wird ein Häufigkeitsmaß definiert, mit dem Wege bewertet werden, und eine Gewichtung vorgestellt, mit der die Popularität in das vorhandene Kantengewicht einfließt, um daraus „kürzeste“ Wege im Sinne der Graphentheorie berechnen zu können.

Schlüsselwörter: Routenplanung, populäre Routen, Wegehäufigkeiten

Abstract: Although there already are numerous route planners with various objectives, hardly any of them purposefully guide bikers via beautiful/scenic routes. In addition to only using the route's characteristics for rating the track, the planners operate a routing algorithm based on the frequency the routes are taken. Therefore, publicly available track data are used to calculate which paths are popular and the algorithm gives preference to these paths for route planning. A frequency measurement is defined that helps to evaluate routes leading to a weighting system in which the track popularity enhances the edge weights in order to be able to calculate "shortest" ways defined by graph theory.

Keywords: Route planning, popular routes, way frequency

1 Motivation und Stand der Technik

Die Wahl einer Route hängt bei Zweiradfahrern von vielen unterschiedlichen Faktoren ab. So stellte Bernhard Leder in einer Umfrage fest, dass für Motorradfahrer die Kurvigkeit, das Höhenprofil, die Landschaft, Oberflächeneigenschaften der Straße wie z. B. der Grip, die zulässige Höchstgeschwindigkeit und das Aufkommen von LKW-Verkehr Einfluss auf die Streckenwahl hat (Navratil, 2012). Für Fahrradfahrer gelten ähnlichen Kriterien jedoch mit anderer Priorisierung. So spielt das Gripniveau eine geringe Rolle, der Straßenbelag jedoch sehr wohl. Und während Motorradfahrer eher Straßen mit höherer Höchstgeschwindigkeit wählen, bevorzugen Fahrradfahrer eher Straßen mit niedriger Höchstgeschwindigkeit, um dem motorisierten Verkehr aus dem Weg zu gehen.

Einige Straßeneigenschaften wie das Höhenprofil und die zulässige Höchstgeschwindigkeit lassen sich unmittelbar den Kartendaten entnehmen. Für die Kurvigkeit gibt es ebenso Berechnungsfunktionen (Navratil, 2012) wie für landschaftliche Eigenschaften (Hochmair & Navratil, 2008). Neben der Bewertung der Straßeneigenschaften gibt es auch Ansätze, die den Besuch attraktiver Ziele in die Planung mit einbeziehen (Baker & Verstockt, 2017).

Sollen alle diese Kriterien in die Berechnung einer Route einfließen, ergibt sich daraus ein Mehr-Kriterien-Kürzester-Pfad-Problem, das NP-schwierig ist (Mooney & Winstanley, 2006). Daher wird dies häufig durch lineare Kombination aller Kriterien der Wegebewertung zu

einem einfachen Kürzeste-Pfade-Probleme reduziert, wodurch jedoch das Problem unrealistisch vereinfacht wird. Oder es wird mithilfe genetischer bzw. evolutionärer Algorithmen gelöst (Mooney & Winstanley, 2006).

Im produktiven Einsatz sind derzeit nur wenige zielgruppenspezifischen Routenplaner zu finden. Mit Kurviger (<https://kurviger.de/>) existiert z. B. ein Motorradroutenplaner, der kurvige Strecken bevorzugt, für Fahrradfahrer gibt es einige regionale Routenplaner wie den Radroutenplaner NRW (<http://www.radroutenplaner.nrw.de/>). Allerdings verwenden diese Routenplaner nur einzelne der oben genannten Kriterien: Kurviger priorisiert Landstraßen und kurvige Straßen, der Radroutenplaner NRW kann neben kürzesten Strecken auch solche mit weniger Steigungen und Strecken, die das Radnetz, das Knotenpunktnetz oder Themenrouten bevorzugen, berechnen.

Betrachtet man die berechneten Routen, so sind diese sowohl bei Kurviger als auch beim Radroutenplaner NRW deutlich länger als die schnellste bzw. kürzeste Route. Kurviger berechnet bei 20 zufällig ausgewählten Streckenplanungsaufgaben Routen, die 19,6 % mehr Zeit erfordern, als die schnellste autobahnfreie Route¹. Der Radroutenplaner NRW berechnet ebenfalls deutlich längere Routen, wenn nicht die kürzeste, häufig über Hauptverkehrsstraßen führende Route ausgewählt ist. Bei abermals 20 exemplarischen Routing-Aufgaben über das Radnetz ist der Zeitbedarf im Durchschnitt 24,7 % größer, beim Routen über das Knotenpunktnetz sogar 36 % länger.

Diese Streckenverlängerung ist unproblematisch, solange Zweiradfahrer als Betreiber eines Hobbys betrachtet werden, deren primäres Ziel nicht das zügige Erreichen eines Zielortes, sondern das Verbringen einer bestimmten Zeit auf einer schönen Strecke ist. So argumentiert auch Navratil (2012), dessen Algorithmus zur Berechnung kurviger Routen je nach verwendetem Berechnungsmodell Routen mit 24 % bzw. 64 % größerem Zeitbedarf berechnet. Verwenden jedoch Zweiradfahrer – und dies gilt sowohl für Fahrrad- als auch für Motorradfahrer – ihr Fahrzeug als Verkehrsmittel im engeren Sinne, so muss die berechnete Route nicht nur schön, sondern auch effizient sein.

Um Routen zu finden, die zwar attraktiv, aber trotzdem effizient sind, wird hier ein Ansatz beschrieben, bei dem Strecken bevorzugt werden, die auch von anderen Verkehrsteilnehmern häufig gewählt werden. Damit vertraut das Verfahren auf die Ortskenntnis von Personen, die Streckendaten im Internet zu Verfügung stellen. Die Häufigkeit, mit der Straßen bzw. Streckenabschnitte in den Daten vorkommen, wird analysiert, und die Bewertung der Straßen so geändert, dass bei der Routenplanung populäre Straßen bevorzugt werden. Um einzelnen Ausreißern in den Daten, wie z. B. temporären Umleitungen, wenig Raum zu geben, wird für das Verfahren möglichst umfangreiches Datenmaterial herangezogen.

Die eigentliche Routenberechnung basiert auf bekannten Verfahren der Graphentheorie. Sobald das Straßennetz als Graph repräsentiert ist, kann mithilfe des Dijkstra-Algorithmus (Dijkstra, 1959) oder des A*-Algorithmus (Hart, Nilsson, & Raphael, 1968) relativ effizient eine kürzeste Route berechnet werden. Für diese recht alten Algorithmen existieren Verbesserungen, die durch Vorverarbeitung eine deutlich effizientere Routenberechnung ermöglichen (Geisberger, Sanders, Schultes, & Delling, 2008).

¹ Die diesen Daten zugrundeliegenden Routingaufgaben sowie die im Folgenden beschriebenen Programme sind unter <https://github.com/vklBo/popularRoutes> verfügbar.

Zielsetzung dieser Arbeit ist die prototypische Implementierung aller Bestandteile eines Planers populärer Routen und dessen Evaluierung anhand realistischer Datenmengen. In Abschnitt 2 dieser Arbeit wird zunächst das Modell für die Neubewertung der Kanten beschrieben. In Abschnitt 3 zeigen wir, wie dieses Modell in einem praktischen Versuch angewendet wurde, um in Abschnitt 4 die Ergebnisse zu diskutieren und zukünftige Aufgaben zu beschreiben.

2 Modell zur Bewertung populärer Routen

2.1 Modellierung des Straßennetzes

Ein Graph $G = (V, E)$ besteht aus einer Menge Knoten V und einer Menge Kanten E . Jede Kante $e = (v_1, v_2) \in E$ verbindet zwei Knoten v_1 und v_2 aus der Knotenmenge. Zur Darstellung eines Straßennetzes als Graph wird jede Kreuzung, jede Einmündung, aber auch jede Geschwindigkeitsänderung und ähnliches als ein Knoten des Graphen repräsentiert. Jede befahrbare Straße wird als Kante zwischen solchen Knoten dargestellt. Damit folgen wir der Repräsentation der OpenStreetMap (OSM) (Openstreetmap, 2018). Alternative Repräsentationen, bei denen die Kreuzungen und Einmündungen als Kanten und Straßen als Knoten dargestellt werden (Porta, Crucitti, & Latora, 2006), wären nur sinnvoll, wenn die Routenberechnung vorwiegend auf Eigenschaften der Kreuzungen beruhen würde.

Eine Bewertungsfunktion $w: E \rightarrow \mathbb{R}^+$ ordnet jeder Kante e des Graphen einen Wert $w(e)$ zu. Um kürzeste Weg zwischen zwei Knoten zu finden, belegt man den Wert jeder Kante mit der geographischen Länge der entsprechenden Straße, und wendet dann einen der oben genannten Algorithmen zur Berechnung des Weges an, dessen Summe der Kantengewichte minimal ist. Für schnellste Wege belegt man jede Kante mit dem Zeitbedarf für das Befahren der Straße, z. B. indem man die Weglänge durch die dort zulässige Höchstgeschwindigkeit teilt.

2.2 Berechnung der Popularität einer Kante

Da die Berechnung, die in diesem Artikel entwickelt wird, auf der Popularität der Route basieren soll, wird eine Bewertungsfunktion benötigt, die die Popularität jeder einzelnen Kante berücksichtigt. Dazu definieren wir die Häufigkeit, mit der Kante $e \in E$ in einer Tracksammlung T vorkommt, als Häufigkeitsfunktion $H_T: E \rightarrow \mathbb{N}_0$. Mit dieser Funktion ordnen wir jeder Kante im Graphen zu, in wie vielen Tracks einer Sammlung diese Kante befahren wurde.

Diese Häufigkeit ist ein absoluter Wert, der abhängig von der Größe der Tracksammlung und der betrachteten Region, sowie der Anzahl der Kanten in der betrachteten Region ist, und damit als Maß für die Popularität nicht gut geeignet. Um ein geeigneteres Maß zu erhalten, muss die absolute Häufigkeit auf ein relatives Maß mit einem definierten Intervall abgebildet werden.

Die relative Häufigkeit $\frac{H_T(e)}{\max_{e \in E} H_T(e)}$ wäre ein solches Maß im Intervall $[0, 1]$. Da die Kantenhäufigkeit jedoch nicht annähernd gleichverteilt, sondern stark linksschief ist – selten besuchte Kanten kommen deutlich häufiger vor als häufig besuchte Kanten – ist das Maß wenig geeignet: Die meisten Straßen würden als wenig populäre Straßen eingestuft, und nur wenige sehr häufig befahrene Straßen wären als populär bewertet.

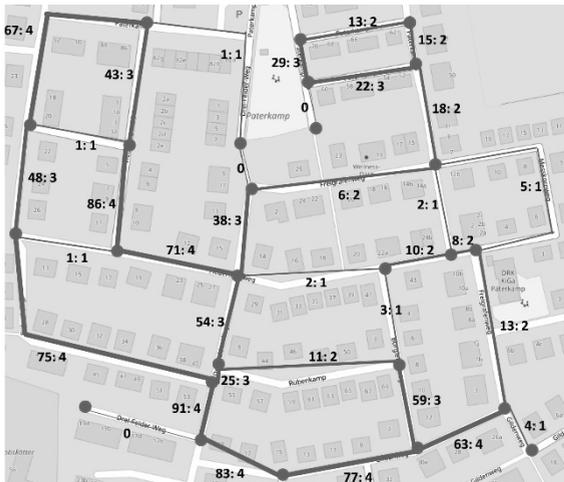


Abb. 1:
Graph eines Kartenausschnitts von Lüdinghausen (OpenStreetMap, 2018) mit exemplarischen Kantenhäufigkeiten und zugeordneten Quantilen

Zur Behebung der Schiefe verwenden wir empirische Quantile (Henze, 2013). Für die in $H_T(E)$ repräsentierten Häufigkeiten werden p -Quantile $H_{i/p}(E)$, $i = \{1, \dots, p\}$ gebildet. $H_{i/p}(E)$ stellt die maximale Häufigkeit dar, die dem unteren i/p -ten Anteil der Kantenhäufigkeit entspricht.

Für eine Kante e ist ihre Bewertung $popularity(e)$ dann das oberste Quantil, in das die Häufigkeit dieser Kante fällt, also

$$popularity(e) = \max_{1 \leq i \leq p} i: H_T(e) \leq H_{i/p}(E).$$

Dieser Wert ist ein diskreter Wert zwischen 1 und der Anzahl Quantile p , wobei p frei gewählt werden kann. Der Wert 0 wird für gar nicht befahrene Straße reserviert. Damit eignet sich die Bewertung sehr gut zum Speichern im Graphen, vor allem, wenn für p 2er-Potenzen gewählt werden, da nur $\log_2(p)$ Bit zur Speicherung jeder Bewertung benötigt werden, was in Anbetracht der sehr großen Anzahl der Kanten, die zur Repräsentation eines realistisch großen Straßennetzes erforderlich sind, nicht irrelevant ist.

Abbildung 1 zeigt einen Kartenausschnitt mit exemplarischen Kantenhäufigkeiten. Für diese wurde 4-Quantile berechnet, als Quantilgrenzen ergeben sich 5 für das erste Quantil, 18 für das zweite, 59 für das dritte und 91 für das vierte Quantil. Entsprechend werden die Kanten diesen vier Quantilen zugeordnet, was ebenfalls in der Abbildung dargestellt ist.

2.3 Regionale Beschränkung der Popularität

Die im Abschnitt zuvor beschriebene $popularity$ -Berechnung zieht alle Kanten des Graphen zur Berechnung der Quantile heran. Dies hat den Vorteil, dass die Quantile nur einmal berechnet werden müssen. Die Anzahl der verfügbaren Tracks ist in verschiedenen Regionen jedoch sehr unterschiedlich: Im Sauerland sind jede Woche viele Hunderte Motorradfahrer unterwegs, von denen viele auch Tracks aufzeichnen und veröffentlichen, in Lappland ist die Anzahl sehr viel niedriger. Entsprechend sind selbst die attraktivsten Straßen in Lappland sehr viel schlechter bewertet als die mittelmäßigen Straßen im Sauerland.

Daher ist für die Berechnung der Popularität eine regionale Gewichtung erforderlich. Im einfachsten Fall werden für jede Kante die Quantile nur basierend auf den Häufigkeiten der Kanten in der Umgebung berechnet, also z. B. der Kanten im Umkreis eines definierten Radius. Diese Berechnung ist jedoch sehr aufwendig, da für jede zu bewertende Kante zunächst alle Kanten der Umgebung ermittelt werden müssen, für diese müssen dann die Quantile berechnet werden, und erst dann kann die Popularität bestimmt werden. Der Graph der OSM-Europakarte enthält ca. 50 Millionen mit Kraftfahrzeugen befahrbare Kanten, deren Popularität auf diesem Weg berechnet werden müsste.

Um die Berechnung effizienter zu gestalten, werden die Quantile für eine festgelegte Menge von Regionen vorab berechnet. Dazu wird ein Längen- und Breitengrad-paralleles Raster über die Karte gelegt, und für jedes Rasterelement werden die Quantile berechnet und gespeichert. Die Anzahl der Rasterelemente ist konstant und wird allein durch die Rastergröße und die Größe des verwendeten Kartenausschnitts bestimmt. Für die Berechnung der Popularität einer Kante kann als Referenzquantil jenes verwendet werden, in dessen Rasterelement diese Kante fällt. Auf diese Weise erhält man eine Bewertung, die relativ zur Region ist.

Allerdings ergeben sich bei diesem Ansatz Verwerfungen an den Grenzen der Rasterelemente. So können zwei gleich häufig befahrene, räumlich nahe Straßen unterschiedlich bewertet werden, wenn sie in zwei verschiedenen Rasterelementen mit unterschiedlichen Quantilgrenzen liegen. Um dies zu verhindern, wird nicht nur ein einziges Rasterelement ausgewählt, sondern die vier Rasterelemente, deren Mittelpunkte der zu bewertenden Kante am nächsten liegen. Aus den Quantilgrenzen der vier Rasterelemente werden unter Verwendung einer linearen inversen Distanzgewichtung (Bill, 2016, p. 532) neue, nur für diese eine Kante gültige, Quantilgrenzen berechnet.

Seien $H_{i/p}(R_j)$ die p -Quantile einer Region R_j . Weiterhin sei $d(e, R_j)$ die geographische Distanz zwischen einer Kante e und dem Mittelpunkt der Region R_j . Dann werden die neuen Quantilgrenzen $H_{i/p}^*(e)$ berechnet als

$$H_{i/p}^*(e) = \left(\frac{\sum_{j=1}^4 \left(\frac{1}{d(e, R_j)} \cdot H_{i/p}(R_j) \right)}{\sum_{j=1}^4 \left(\frac{1}{d(e, R_j)} \right)} \right).$$

Da durch die Neuberechnung der Quantile die Quantilgrenzen sinken können, kann es vorkommen, dass die Kantenhäufigkeit einer sehr beliebten Kante größer ist, als die oberste Quantilgrenze. In diesem Fall wird die Kante als maximal populär mit dem Wert p bewertet.

2.4 Bestimmung der Gewichtsfunktion

Für die eigentliche Routenberechnung benötigen wir schließlich noch eine Funktion, die Kanten anhand des Zeitbedarfs für das Befahren der Kante bewertet, andererseits aber populäre Straßen bevorzugt.

Sei $w_{time}: E \rightarrow R^+$ eine Gewichtsfunktion, die für jede Kante den Zeitbedarf für das Befahren dieser Kante berechnet, also die Gewichtsfunktion, um schnellste Wege zu finden.

Sei $w_{pop}: E \rightarrow [0, 1]$ eine Funktion, die jeder Kante e ein Gewicht anhand ihrer Popularität zuweist, wobei populäre Kanten kleine Werte erhalten, also gilt

$$e_1, e_2 \in E: H_T(e_1) \geq H_T(e_2) \Rightarrow w_{pop}(e_1) \leq w_{pop}(e_2).$$

Dann ist

$$w(e) = w_{pop}(e) \cdot w_{time}(e)$$

eine Bewertung, die bei der Berechnung kürzester Wege populäre schnelle Routen finden lässt. Der Popularitätsfaktor verlängert dabei unter Umständen die gefundene Route, da er den Zeitbedarf einer Kante virtuell reduziert. Gesucht ist daher eine Funktion w_{pop} , die einen geeigneten Ausgleich zwischen Popularität und Zeitbedarf darstellt.

Als Funktion für die Gewichtung einer Kante im Routing wurde durch praktische Versuche als Einstieg für eine spätere Optimierung folgende Funktion gewählt, bei der der Teiler p der Maximalwert der Kantenpopularität, also die Anzahl der Quantile ist:

$$w_{pop}(e) = \left(\frac{1 - \text{popularity}(e)}{2p} \right)^2 \in [0.25, 1].$$

Die Funktion ist so gewählt, dass die populärsten Straßen einen Bonus von 0,25 erhalten, unpopuläre Straßen erhalten eine Bewertungskorrektur von 1, also keinen Bonus. Bei kleineren Werten als maximalem Bonus zeigt sich der Dijkstra-Algorithmus als ineffizient, da aufgrund der geringen Kantenkosten sehr viele Kanten expandiert werden. Bei größeren Werten erweist sich der Einfluss der Bewertung populärer Straßen als relativ gering.



Abb. 2:

Graph mit Quantilen, den daraus berechneten Gewichten sowie dem kürzesten und dem populärsten Pfad

Abbildung 2 zeigt den bereits in Abbildung 1 dargestellten Graph mit den zugeordneten Quantilen und den daraus berechneten Gewichten, wobei zur Vereinfachung als Zeitgewichtung die konstante Funktion $w_{time}(e) = 1$ verwendet wurde. Der einfach gestrichelte Pfad kennzeichnet den kürzesten, der doppelt gestrichelte den populärsten Weg.

Die Trennung der Berechnung von Popularität und Kantengewicht ermöglicht es, den Einfluss der Popularität auf die Routenberechnung einfach und benutzerorientiert zu verändern, da die Popularität in einem Vorverarbeitungsschritt einmalig berechnet wird, das Kantenge-

wicht im eigentlichen Routingverfahren dann aber dynamisch bei jeder Routenberechnung neu berechnet werden kann.

3 Evaluation am Beispiel eines Motorradroutenplaners

In diesem Abschnitt weisen wir nach, dass der gewählte Ansatz realistisch durchführbar ist und zeigen exemplarische Routingergebnisse. Dazu wurden eine relevante Anzahl Motorradtracks² analysiert, die Kantenhäufigkeiten ermittelt, daraus die Popularität berechnet, und die Gewichtsfunktion in einem Routenplaner implementiert. Im Folgenden wird der exakte Versuchsaufbau beschrieben. Dabei werden auch die Größenordnungen der Daten und des Zeitbedarfs für die Berechnungen dargestellt, um nachzuweisen, dass das entwickelte Verfahren auch für realistische Datenmengen handhabbar ist.

3.1 Verwendete Hard- und Software

Die Evaluation wurde auf einem PC mit Intel Xeon E3-1225 Prozessor mit 3.3 GHz Prozessortakt und 32 GByte RAM mit Ubuntu-Linux als Betriebssystem durchgeführt. Zur Speicherung der Daten wurde PostgreSQL 9.5 mit PostGIS in der Version 2.2.1 verwendet. Der Import der Kartendaten erfolgte mit Osmosis Version 0.44. Als Kartenmaterial wurde die OSM-Europakarte vom 15.01.2018 (Geofabrik, 2018) verwendet. Für die Berechnung der Kantenpopularität und die eigentliche Routenberechnung wurde die Open-Source-Software Graphhopper in der Version 9.5 (Graphhopper, 2017) um eigene Programmierung ergänzt.

3.2 Import der Kartendaten und Berechnung der Popularität

Für die Zählung der Kantenhäufigkeiten ist der Import aller relevanten, also befahrbaren Straßen inklusive deren geographischen Positionen erforderlich. Dieser Import wurde mit Osmosis in 24 Stunden durchgeführt. Es wurden 51.579.478 befahrbare Kanten aus dem Europa-Kartensatz in die Datenbank importiert.

Für die Zählung der Kantenhäufigkeit wurden von acht einschlägigen Internetpräsenzen mit Motorradtouren mithilfe von Linux-Programmen oder einfachen selbst programmierten Skripten 193.580 Trackdateien geladen. Eine Sichtung der Tracks ergab, dass sie hinsichtlich Länge der aufgezeichneten Strecke und Dichte der Trackpunkte sehr unterschiedlich waren. Um eine Mindestqualität sicherzustellen, wurden alle Tracks eliminiert, deren Parsen syntaktische Fehler ergab, deren Diameter³ kleiner als 10 km waren oder die nicht im Bereich der Karte lagen. Es verblieben 84.643 Trackdateien, deren Tracks mithilfe der Graphhopper-Software nachgefahren wurde, indem für jedes Paar aufeinanderfolgender Trackpunkte eine Route berechnet wurde. Die Häufigkeit aller von der Route berührten Kanten wurde inkrementiert. Die Berechnung der Kantenhäufigkeit hat inklusive des Parsens der Tracks 55:32 Stunden benötigt. Mit Abschluss dieses Verarbeitungsschrittes enthielt die Datenbank 7.375.348

² Für die Studie wurden Motorradrouten verwendet, weil bei Fahrradrouten die Unterscheidung der Zielgruppen Rennrad-, Touren- und Mountainbikefahrer relevant ist, für deren Unterscheidung aber Metainformationen über die Tracks gesammelt werden müssten, worauf wir bisher aus Aufwandgründen bisher verzichten mussten.

³ Unter dem Diameter verstehen wir die Länge der Diagonale der Bounding-Box des Tracks.

Kanten, die durch mindestens einen Track berührt wurden, die Kante mit der höchsten Häufigkeit kam in 1463 Tracks vor und insgesamt 96.550.577 Berührungen wurden gezählt.

Im nächsten Schritt wurde ein Raster über Europa gelegt, und für jedes Rasterelement aus den lokalen Häufigkeiten Quantile gebildet. Als Rastergröße wurde jeweils ein halbes Längen- und Breitengrad gewählt, wodurch die Rasterelemente in Mitteleuropa eine Diagonale von ca. 65 km Länge hatten. Für den gewählten Kartenausschnitt Europa ergeben sich, unter Berücksichtigung der Situation, dass über Wasserflächen und einigen extrem dünn besiedelten Landflächen keine Häufigkeiten berechnet werden müssen, 4734 zu bewertende Regionen, für die alle Kanten dieser Region ermittelt und Quantile gebildet wurden. Als Anzahl für die Quantile wurde $p=7$ gewählt. Da zusätzlich die Häufigkeit 0 gespeichert werden muss, werden drei Bits zur Speicherung der Häufigkeit benötigt. Die Berechnung der Quantile benötigte in unserem Experiment 4:43 Stunden. Um die berechneten Quantile und weitere Eigenschaften einer Region anzeigen zu können, wurde die Graphhopper-Software um die Einblendung dieser Daten in der Karte ergänzt.

Der letzte Vorbereitungsschritt ist die Berechnung der Popularität jeder einzelnen Kante. Da gar nicht besuchte Kanten immer die Popularität 0 haben, werden nur besuchte Kanten betrachtet. Für jede Kante werden, wie in Abschnitt 2.3 beschrieben, mithilfe der inversen Distanzgewichtung spezifische Quantilgrenzen berechnet und geprüft, in welches Quantil die Kante fällt. Um die Popularität zu visualisieren, wurde das Graphhopper-Frontend um eine Komponente erweitert, mit der für jede Popularitätsstufe die Kartenkanten unterschiedlich eingefärbt werden. Die Berechnung der Kantenpopularität der gut 7 Mio. zu bewertenden Kanten erforderte 50:36 Stunden.

Als Funktion für die Gewichtung einer Kante wurde die bereits im Abschnitt 2.4 definierte Funktion $w_{pop}(e)$ verwendet, wobei durch die gewählte Quantilanzahl p den Wert 7 hat. Zusätzlich wurde das Standardkantengewicht für Kanten auf Motorradfahrer angepasst und nicht nur der Zeitbedarf bewertet: Autobahnen und Ortsdurchfahrten wurden abgewertet.

3.3 Ergebnisse

In einer ersten Studie wurden manuell 35 verschiedene Routen unterschiedlicher Länge in Europa berechnet, und mit den Routen anderer Routenplaner verglichen. Die verwendeten Routenplaner waren Google Maps (<https://www.google.de/maps>), Motoplaner (<http://motoplaner.de/>) und Kurviger mit den Einstellungen „schnell“ und „kurvig“. Alle Planer wurden in der Grundeinstellung verwendet, nur die Nutzung von Autobahnen war bei Google Maps und Motoplaner jeweils ausgeschaltet, um die Ergebnisse vergleichbar zu halten.

Abbildung 3 zeigt die durchschnittliche Routenlänge und den durchschnittlichen Zeitbedarf für die berechneten Routen der verschiedenen Routenplaner. Bei der Analyse des Ergebnisses ist zu beachten, dass Motoplaner, Google Maps und Kurviger die Fahrzeit etwas konservativer als „Popular Routes“ berechnen, weswegen in den nachfolgenden Betrachtungen für „Popular Routes“ ein empirisch ermittelter Korrekturfaktor von 12 % auf die Fahrzeit aufgeschlagen wurde. Die Distanzen von „Popular Routes“ und Kurviger (schnell) sind etwa 4 % größer als die von Google Maps, die von Kurviger (kurvig) sind mit 17 % deutlich länger. Die Fahrzeiten von „Popular Routes“ sind etwa 9 % länger als die Fahrzeiten von Google Maps und Kurviger (schnell), Kurviger (kurvig) erzeugt Routen, die im Mittel nochmal 21 % mehr Zeit benötigen als „Popular Routes“.

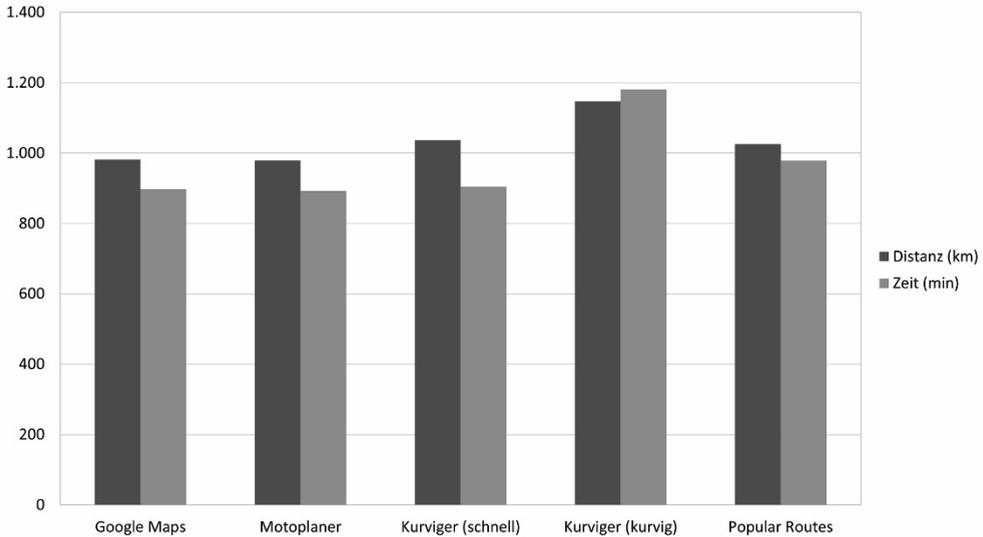


Abb. 3: Vergleich verschiedener Routenplaner am Beispiel zufälliger Routen

4 Diskussion und zukünftige Aufgaben

Mit dieser Arbeit stellen wir einen Rahmen für Routenplaner vor, der die Popularität von Straßen in die Berechnung von Routen mit einfließen lässt. Die Arbeit stellt eine erste Studie vor mit der gezeigt wird, dass die Berechnung einer solchen Popularität auch bei großen Datenmengen und realistischen Kartenausschnitten in vertretbarer Zeit möglich ist. Auch zeigt sie, dass die berechneten Routen zumindest dem äußeren Anschein nach für die in den Versuchen gewählte Zielgruppe Motorradfahrer attraktiv sind, sie sind unerheblich länger als schnellste Routen, und führen durch attraktive Regionen.

Die Bewertung der Qualität der gefundenen Routen basiert in diesem Artikel nur auf der Länge und des Zeitbedarfs für die Route. Ob die Route wirklich attraktiver ist, wird so nicht erkannt. Auffällig ist jedoch, dass einige der verglichenen Routen deutlich voneinander abweichen, z. B. routen manche Routenplaner in den italienischen Süden entlang der Adria-Küste, während andere Algorithmen Strecken durch das Apennin berechnen. Für eine objektive Analyse könnte zukünftig berechnet werden, wie groß der Anteil populärer Kanten im Vergleich zu den Routen anderer Routenplaner ist. Andererseits zeigt sich die Qualität der Routen vor allem im tatsächlichen Gebrauch des Routenplaners durch die Zielgruppe. Dazu müsste der Routenplaner zumindest prototypisch in Produktion genommen werden.

Viele Parameter bei Berechnung der Popularität und ihrer Gewichtung in der Routenplanung sind bisher empirisch ermittelt worden und daher zunächst exemplarisch zu verstehen. Sie sollten durch eine systematische Analyse evaluiert und optimiert werden. Um die Popularität präziser berechnen zu können, könnten zukünftig nicht nur die Häufigkeiten gezählt werden, sondern auch Metadaten über die verwendeten Tracks betrachtet werden. Dies können z. B. Bewertungen anderer Nutzer, Quelle der Tracks und Qualität der Tracks sein. Die festen Rasterelemente könnten durch funktionale Regionen, die z. B. mit dem SANET-Tool

(<http://sanet.csis.u-tokyo.ac.jp/>) gebildet werden können, ersetzt werden. Die Kombination der Popularitätsgewichtung mit anderen Bewertungen, wie Kurvigkeit oder landschaftliche Attraktivität, sollte ebenfalls betrachtet werden. Im Idealfall ersetzt aber das hier vorgestellte Verfahren zusätzliche Bewertungen, da alle diese Faktoren implizit in der Popularität einer Straße aufgehen sollten.

Insgesamt erscheint der hier vorgestellte Ansatz als geeignet, für Zweiräder Routen zu berechnen, die ein zielorientiertes Verwenden dieser Verkehrsmittel ermöglicht. Aus diesem Grund ist Zielsetzung des Autors, den Ansatz um die Berechnung von Fahrradrouten zu ergänzen, insbesondere auch, um die Attraktivität des Fahrrads als regelmäßiges Verkehrsmittel zu steigern.

Literatur

- Baker, K., & Verstockt, S. (2017). Cultural Heritage Routing: A Recreational Navigation-based Approach in Exploring Cultural Heritage. *Journal on Computing and Cultural Heritage*, 10(4), 1–20.
- Bill, R. (2016). *Grundlagen der Geo-Informationssysteme*. 6. Ed. Berlin/Offenbach: Wichmann.
- Dijkstra, E. W. (1959). A Note on Two Problems in Connexion with Graphs. *Numerische Mathematik*, 1, 269–271.
- Geisberger, R., Sanders, P., Schultes, D., & Delling, D. (2008). Contraction Hierarchies: Faster and Simpler Hierarchical Routing in Road Networks. In: C. C. McGeoch (Ed.), *Lecture Notes in Computer Science: Experimental Algorithms* (pp. 319–333). Springer.
- Geofabrik (2018). *Europakarte OSM, Lizenz CC-BY-SA 2.0*. Retrieved from <http://download.geofabrik.de/europe-latest.osm.pdf>.
- Graphhopper (2017). Fast and Flexible Route Planning. Apache license, Version 2.0. Retrieved from <https://github.com/graphhopper/>.
- Hart, P., Nilsson, N., & Raphael, B. (1968). A formal basis for the heuristic determination of minimum cost paths. *Numerische Mathematik*, 2, 100–107.
- Henze, N. (2013). *Stochastik für Einsteiger: Eine Einführung in die faszinierende Welt des Zufalls*. 10. Ed. Wiesbaden: Springer Fachmedien.
- Hochmair, H., & Navratil, G. (2008). Computation of scenic routes in street networks. *Geospatial Crossroads@ GI_Forum'08: Proceedings of the Geoinformatics Forum Salzburg*. (pp. 124–133).
- Mooney, P., & Winstanley, A. (2006). An evolutionary algorithm for multicriteria path optimization problems. *Int. Journal of Geographical Information Science*, 20(4), 401–423.
- Navratil, G. (2012). Curviness as a Parameter for Route Determination. In T. Jekel & A. Koller (Eds.), *GI_Forum 2012: Geovisualisation, society and learning* (pp. 355–364). Berlin/Offenbach: Wichmann.
- Openstreetmap (2018). *Lizenz CC-BY-SA 2.0*. Retrieved from <https://www.openstreetmap.de/karte.html>.
- Porta, S., Crucitti, P., & Latora, V. (2006). The network analysis of urban streets: A dual approach. *Physica A: Statistical Mechanics and its Applications*, 369(2), 853–866.