

IT-Beratung Geoinformatik, Hattingen; Hochschule Bochum

MODELLIERUNG MENSCHLICHEN MOBILITÄTSVERHALTENS IN POI-EMPFEHLUNGSSYSTEMEN MIT HILFE VON METHODEN DER GEOINFORMATIK

Felix Maas, Benno Schmidt

Zusammenfassung: Das Ziel von Empfehlungssystemen besteht darin, dem Nutzer aus einer großen Menge an Artikeln solche zu empfehlen, die ihn am meisten interessieren könnten. E-Commerce- und Streamingdienste sind die bekanntesten Felder, in denen sie zum Einsatz kommen. In den letzten Jahren wurden sie auch in ortsbezogene soziale Netzwerke wie Foursquare oder Yelp integriert, um dem Nutzer personalisierte Empfehlungen über Points of Interest (POIs) zu geben. Zur Ermittlung der passenden Empfehlungen werden dabei häufig Methoden des maschinellen Lernens eingesetzt.

Dieser Beitrag zeigt auf, wie das Verständnis von menschlichem Mobilitätsverhalten im Zusammenspiel mit dem technischen Wissen aus der Geoinformatik zu einer verbesserten Qualität von POI-Empfehlungen beitragen kann. Er bezieht damit Disziplinen ein, welche bisher nur in stark eingeschränktem Maße in der Forschung und Entwicklung um das Themengebiet der POI-Empfehlungssysteme berücksichtigt wurden.

Schlüsselwörter: POI-Empfehlungssysteme, Empfehlungssysteme, Points of Interest, Data Mining, Machine Learning, Mobilität, ortsbezogene Dienste, ortsbezogene soziale Netzwerke

MODELING HUMAN MOBILITY BEHAVIOR IN POI RECOMMENDER SYSTEMS BY USING METHODS OF GEOINFORMATICS

Abstract: The goal of recommender systems is to recommend to the user from a large number of articles those ones that would interest him the most. E-commerce and streaming are the most popular fields in which recommender systems are used. In recent years, they have also been integrated into location-based social networks, such as Foursquare or Yelp, to give users personalized recommendations about points of interests. Methods of machine learning are often used to determine the appropriate recommendations.

This article points out how the understanding of human mobility behavior in combination with technical geoinformatics knowledge can contribute to improving the quality of POI recommendations. It thus includes disciplines which have hitherto been considered only to a very limited extent in research and development in the area of POI recommendation systems.

Keywords: POI recommender systems, recommender systems, points of interest, data mining, machine learning, mobility, location-based services, location-based social networks

Autoren

M. Eng. Felix Maas
IT-Berater Geoinformatik
Im Schneppenkamp 2
D-45525 Hattingen
E: felixfms@gmx.de

Prof. Dr. Benno Schmidt
Hochschule Bochum
Fachbereich Geodäsie
Lennershofstraße 140
D-44801 Bochum
E: benno.schmidt@hs-bochum.de

1 EINLEITUNG

Recommender Systems (dt. Empfehlungssysteme) zählen zu den interessantesten Entwicklungen im Umfeld von Big-Data- und Machine-Learning-Anwendungen. Bekanntheit erlangten sie vor allem im Bereich der E-Commerce-Anwendungen und Streamingdienste. Für jede Produkt-, Film- und Musikstückempfehlung auf Diensten wie *Amazon*, *Netflix* oder *Spotify* ist die Einbindung eines solchen Systems notwendig. In den letzten Jahren setzte sich die Forschung in diesem Rahmen auch mit ortsbezogenen Empfehlungen auseinander. In diesem Zuge entwickelte sich mit den POI Recommender Systems eine spezielle Art von Empfehlungssystemen, welche dem Nutzer vorrangig in ortsbezogenen sozialen Netzwerken wie *Foursquare*, *Yelp* oder *Gowalla* Points of Interest (POIs) vorschlagen, die er oder sie besuchen sollte.

Fortwährender Gegenstand der Forschung ist die Entwicklung neuer Algorithmen zur Optimierung der Empfehlungsqualität. In POI Recommender Systems wird diese insbesondere durch den Einbezug von Kontexteinflüssen gewährleistet. Während traditionelle Empfehlungssysteme sich bevorzugt mit dem Mining der Nutzerinteressen aus bisherigen Transaktionen mit dem jeweiligen System auseinandersetzen, benötigen POI-Empfehlungssysteme weitergehende Informationen über beispielsweise soziale Beziehungen zu anderen Nutzern, den Zeitpunkt, auf den sich die Empfehlung bezieht oder über den Aktionsraum des Nutzers sowie den Standort potenziell interessanter POIs.

Dieser Beitrag liefert Hinweise, welche Beiträge die Geoinformatik im Hinblick auf die Optimierung durch Berücksichtigung georäumlicher Aspekte liefern kann. Hierzu werden zunächst das zugrunde liegende POI Recommendation Problem und damit zusammenhängende Fachbegriffe erläutert (Kapitel 2). Nachfolgend werden bekannte Einflussfaktoren auf POI-Empfehlungen vorgestellt sowie Algorithmen zur Integration des geographischen Einflusses erläutert (Kapitel 3). Daraufhin werden interdisziplinäre Mobilitätsstudien untersucht, um das bestehende Konzept kritisch zu hinterfragen und Handlungsempfehlungen aufzustellen (Kapitel 4). In Kapitel 5 werden Methodiken der Geoinformatik vorgestellt, mit welchen diese Verbesserungsvorschläge in POI-Empfehlungsalgorithmen integriert

werden können. Im Anschluss folgt eine Diskussion über die Herausforderungen, welche mit der Einführung dieser Methodiken einhergehen (Kapitel 6). Zuletzt werden die wichtigsten Ergebnisse dieses Beitrags zusammengefasst (Kapitel 7) und künftige Forschungsthemen in Bezug auf POI-Empfehlungen aus Sicht der Geoinformatik vorgestellt (Kapitel 8).

2 DAS POI-EMPFEHLUNGSPROBLEM

Ein POI umfasst im Folgenden die Angabe zweier Koordinaten x und y bezogen auf ein Koordinatenreferenzsystem (z. B. geographische Länge und Breite). Weiterhin können einem POI thematische Attribute wie POI-Bezeichner, Kategorienangaben etc. zugeordnet sein. Mit Blick auf das POI Recommendation Problem ist die POI-Definition hier einfach gefasst. Alternative Begriffsdefinitionen wurden z. B. durch das W3C oder des OGC erarbeitet (POI Working Group 2010, Holt et al. 2013). In der praktischen Anwendung können dabei neben punkthaften auch andere Geometrietypen, wie z. B. Linien oder Polygone, relevant sein. Dieser Fall wird hier nicht weiter betrachtet.

Bereits besuchte POIs sollen von den Empfehlungen ausgeschlossen werden. Diese Menge der von einem Anwender u besuchten POIs wird nachfolgend als $S(u)$ bezeichnet. Zur Aufzeichnung der Menge $S(u)$ lassen sich Check-in-Ereignisse verarbeiten, die bei Betreten einer Örtlichkeit (x, y) durch den Systemnutzer ausgelöst werden. Insofern bildet $S(u)$ als Teilmenge aller vom System verwalteten POIs (nachfolgend als L bezeichnet) die nutzerspezifische Check-in-Historie ab (vgl. Zhao et al. 2016).

Eine POI Recommendation lässt sich mathematisch als nutzerspezifische Teilmenge der Differenzmenge $L - S(u)$ fassen. Praktisch interessieren dabei nicht alle Empfehlungen, sondern nur die „besten“ bezogen auf ein geeignetes Bewertungssystem, z. B. die POIs mit der höchsten Attraktivität oder dem größten erwarteten Nutzen für den Anwender. Dies erfordert die Einführung einer Bewertungsfunktion $r(l, u)$, die einem POI l aus der Menge L bezogen auf den Nutzer u ein Element einer geordneten Menge zuordnet. Das POI Recommendation Problem lässt sich schließlich formal als Tupel $\langle U, L, \{S(u)\}, r \rangle$ spezifizieren, wobei U die Menge aller Systemnutzer bezeichnet.

Eine Top- n -Recommendation $R_n(u)$ liefert bezogen auf das Problem $\langle U, L, \{S(u)\}, r \rangle$ die n besten POI Recommendations, wobei die Zahl $n > 0$ ist. Für eine detaillierte mathematische Beschreibung des POI-Recommendation-Problems siehe Schmidt & Maas (2020).

3 EINFLUSSFAKTOREN AUF POI-EMPFEHLUNGEN

In bisherigen POI-Empfehlungsalgorithmen werden insbesondere geographische, temporale und soziale Einflüsse berücksichtigt. Dabei spielt der geographische Einfluss die größte Rolle (Yu & Chen 2015). Die Faktoren wurden größtenteils durch die Auswertung von Check-in-Daten in ortsbezogenen sozialen Netzwerken detektiert. Die Erkenntnisse werden im Folgenden detaillierter mit Verweisen auf entsprechende Data-Mining-Studien erläutert:

- ▶ Der **geographische Einfluss** bezieht sich auf die Entdeckung des *geographischen Clusteringphänomens* im Check-in-Verhalten von Nutzern ortsbezogener sozialer Netzwerke. In dem Zusammenhang fanden Cho et al. (2011) heraus, dass Nutzer bevorzugt Orte in der Nähe ihrer Aktionszentren, wie z. B. ihrem Wohnstandort oder ihrer Arbeitsstätte, aufsuchen. Ye et al. (2011) entdeckten zusätzlich, dass die Check-in-Wahrscheinlichkeit eines unbesuchten POI von dessen Entfernung zu besuchten POIs abhängt. Erwähnenswert ist hierbei, dass POI-Empfehlungsansätze das geographische Clusteringphänomen unterschiedlich interpretieren. Während beispielsweise Ye et al. (2011) und Li et al. (2015) niedrige Distanzen zwischen einem Nutzer und einem neuen POI mit potenziell höherem Interesse verknüpfen, argumentieren Lian et al. (2014), dass ein unbesuchter POI in unmittelbarer Nähe anderer besuchter POIs eines Nutzers als für ihn unattraktiv zu bewerten ist, da er trotz seiner niedrigen Distanz in der Vergangenheit nie besucht wurde.
- ▶ Der **temporale Einfluss** bezeichnet zeitlich variierende Interessen an potenziellen POIs (Gao et al. 2013) bzw. POI-Kategorien (Hasan et al. 2013). So werden Einkaufszentren und Erholungsgebiete vor allem an den Wochenenden und Arbeitsstätten bevorzugt werktags besucht. Auch die Tageszeit beeinflusst die Check-in-Wahrscheinlichkeit.

lichkeiten. Beispielsweise werden Restaurants mittags und abends präferiert.

- Der **soziale Einfluss** bezieht sich auf die Beobachtung, dass Nutzer in ortsbezogenen sozialen Netzwerken, welche in einer Freundschaftsbeziehung zueinanderstehen, eher gemeinsame Interessen haben als andere. Dieser Einfluss ist jedoch im Gegensatz zu den anderen genannten Faktoren eher gering (Yu & Chen 2015).

Die Geoinformatik kann vor allem bei der Betrachtung des geographischen Einflusses eine erweiterte Sichtweise liefern. In bisherigen POI-Empfehlungsansätzen werden Wahrscheinlichkeitsdichtefunktionen zur Ermittlung der Check-in-Wahrscheinlichkeit in der Regel in Abhängigkeit der euklidischen Distanz eines unbesuchten POIs zu den Aktionszentren des Nutzers bzw. zu allen von ihm besuchten POIs angehalten. Die gängigsten Algorithmen greifen hierbei auf Gaußsche Normalverteilungen, exponentielle Zusammenhänge und Kerndichteschätzungen zurück (Zhao et al. 2016).

4 BEWERTUNG DES GEOGRAPHISCHEN EINFLUSSES AUS DER PERSPEKTIVE VON MOBILITÄTSSTUDIEN

Um einen Einblick in die Auswirkung geographischer Zusammenhänge auf POI-Empfehlungen zu erhalten und bisherige Ansätze zu bewerten, ist ein Verständnis der Motivation notwendig, warum die Entfernung bei der Entscheidung eines Ortsbesuchs für ein Individuum eine entscheidende Einflussgröße darstellt. Ein solcher Zugang kann durch Erklärungsmodelle menschlichen Mobilitätsverhaltens hinsichtlich der Zielwahl gegeben werden. Die Psychologie und die

Verkehrswissenschaften stellen in diesem Feld führende Disziplinen dar. Psychologische Studien heben insbesondere interne personenbezogene Einflüsse wie *Kontrollüberzeugungen, Einstellungen, Normen* und *Werte* hervor (Hunecke 2015). Weitere personenbezogene, jedoch nicht tiefenpsychologische Einflüsse auf das Mobilitätsverhalten stellen *soziodemografische Merkmale* sowie der *Lebensstil* von Individuen dar (Jarass 2012).

Laut verkehrswissenschaftlichen Studien, wie der von Hunecke & Schweer (2006), erklären vor allem externe strukturelle Einflüsse die Zielwahl. Wesentliche Determinanten stellen die *Raum- und Siedlungsstruktur* sowie das *verkehrsinfrastrikturale Angebot* dar. Die *Verkehrsmöglichkeiten* einer Person bestimmen weiterhin den Zugang und damit die Möglichkeit zur Nutzung von Teilen der Verkehrsinfrastruktur (Dick 2009). Auch der Wohnstandort, welcher nach Beckmann et al. (2006) als „zentrale[r] Knoten individueller Lebens- und Aktionsräume [im Lebensalltag]“ gilt, stellt ein entscheidendes Kriterium für die individuelle Wahrnehmung der genannten Größen dar. Durch all jene Faktoren entstehen Möglichkeiten und Zwänge der Mobilität (Hunecke & Schweer 2006, Jarass 2012). Diese wiederum bestimmen, wie gut ein Ziel für das Individuum erreichbar ist. Demnach kann die Erreichbarkeit als zentrale Einflussgröße auf die Zielwahl verstanden werden, welche alle in dem Zusammenhang genannten Parameter zusammenfasst (vgl. Abb. 1).

Die Erreichbarkeit lässt sich ebenso als Erklärungsansatz für die Modellierung des geographischen Einflusses in POI-Empfehlungssystemen interpretieren. Wie in Kapi-

tel 3 bereits erwähnt, spielt die Lage der Aktionszentren von Nutzern sowie die Lage möglicher Ziele bei der Modellierung eine entscheidende Rolle. Dies trifft auch auf die Erklärungsansätze aus genannten Mobilitätsstudien zu: Die Lage der Zielstandorte wird durch die Raum- und Siedlungsstruktur bestimmt. Aktionszentren können als übergeordnete Entität der Wohnstandorte betrachtet werden. Cho et al. (2011) stellen heraus, dass es sich bei einem Aktionszentrum beispielsweise auch um die Arbeitsstätte handeln kann.

Der entscheidende Interpretationsunterschied liegt in der Quantifizierung der Erreichbarkeit. Mobilitätsstudien heben hierbei vor allem den Einbezug der Verkehrsinfrastruktur sowie ihrer individuellen Nutzbarkeit hervor. Als essenzielle Metriken zur Quantifizierung der Erreichbarkeit werden zeitliche und monetäre Kosten erwähnt, welche bei der Nutzung des zur Verfügung stehenden Verkehrsangebots anfallen. Sie werden fortan als *Verkehrsaufwände* bezeichnet. POI-Empfehlungssysteme nutzen dagegen die stark simplifizierende euklidische Distanzmetrik.

Dass eine Berücksichtigung der verkehrstechnischen Anbindung bei der Zielwahl als sinnvoller erscheint als die Nutzung der euklidischen Distanz, wird durch das Fallbeispiel in Abbildung 2 verdeutlicht. Dargestellt werden Pkw-Routen zu zwei verschiedenen POIs A und B von einem Aktionszentrum AZ aus. Bei beiden POIs handelt es sich um Restaurants mit Rheinblick, stellen inhaltlich also ein ähnliches Interessenpotenzial dar. Die euklidische Distanz zu POI A ist zwar geringer als jene zu POI B, allerdings ist POI B sowohl zeitlich als auch monetär kostengünstiger

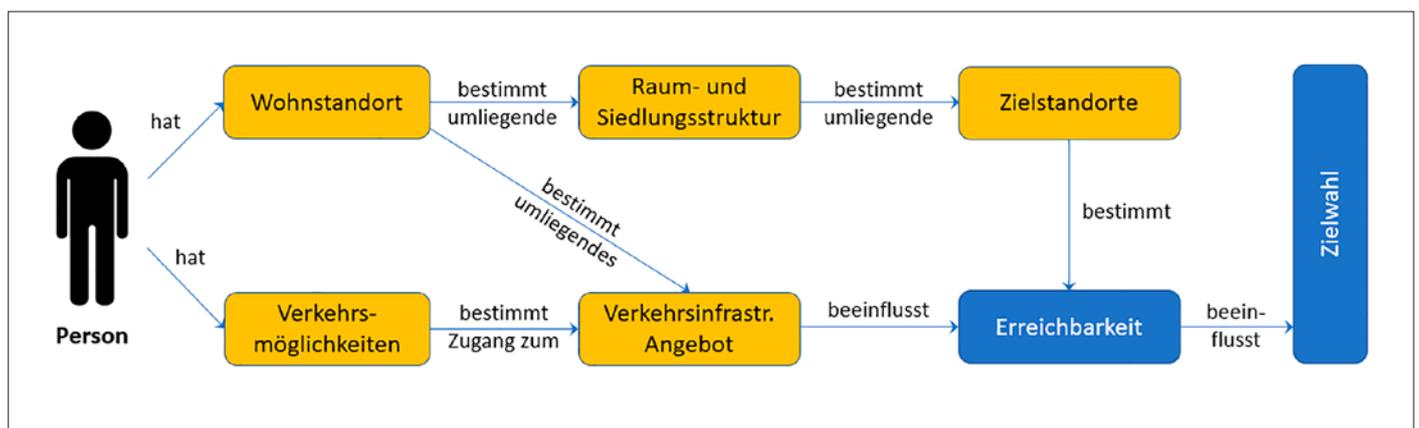


Abbildung 1: Einflussfaktoren auf die Erreichbarkeit von Zielstandorten

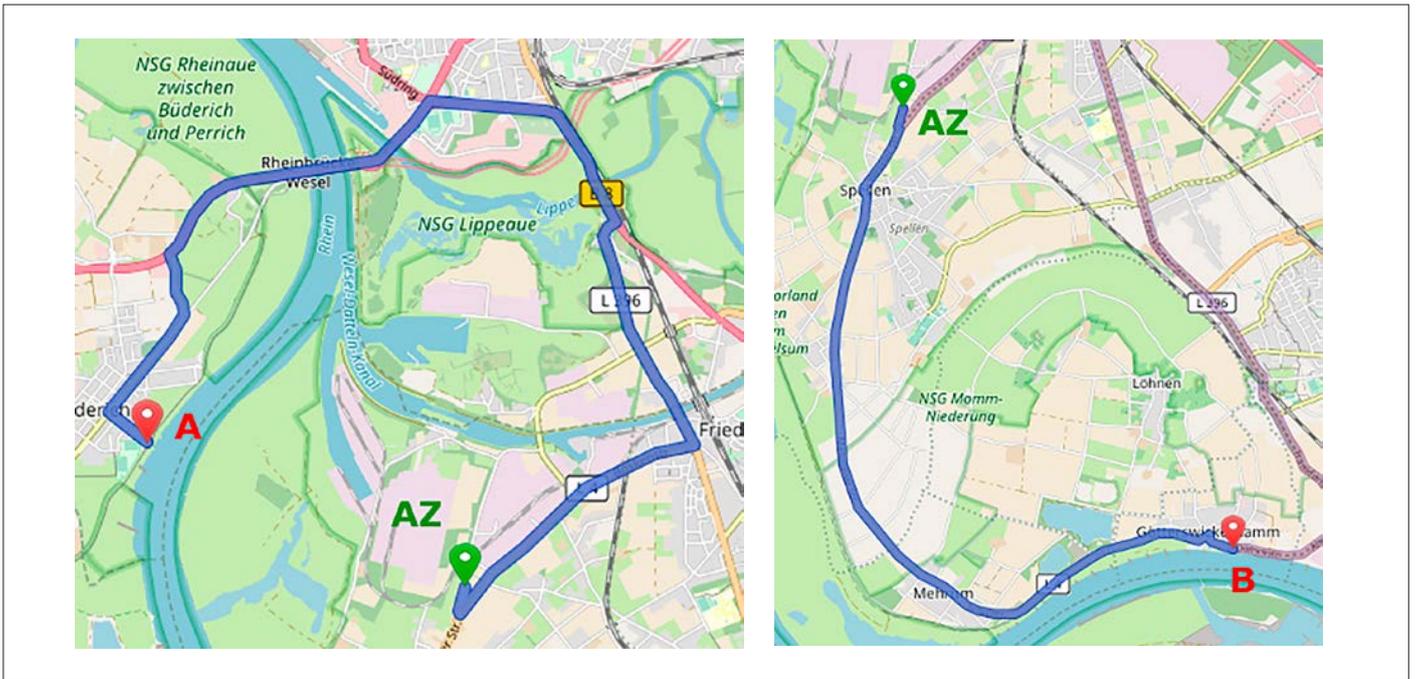


Abbildung 2: Routen von einem Ausgangspunkt AZ zu euklidisch unterschiedlich weit entfernten Punkten A und B (Map data copyrighted OpenStreetMap contributors and available from <https://www.openstreetmap.org>)

zu erreichen. Der wesentliche Grund für die Diskrepanz der beiden Metriken ist der Rhein. Als natürliche Grenze sorgt er dafür, dass der Nutzer des Systems zunächst eine Brücke aufsuchen muss, um POI A aufzusuchen. Dies ist beim Besuch von POI B nicht notwendig.

5 ALTERNATIVE METHODE ZUR BERÜCKSICHTIGUNG DER ERREICHBARKEIT IN POI-EMPFEHLUNGEN

Unter der Berücksichtigung der Erkenntnisse aus interdisziplinären Mobilitätsstudien wird deutlich, dass die euklidische Distanz zwischen dem Nutzer und potenziell interessanten POIs den Einfluss der Erreichbarkeit auf POI-Empfehlungen nur unzulänglich beschreiben kann. Wie im Folgenden deutlich wird, kann die Geoinformatik einen entscheidenden Beitrag zu einer den Mobilitätsstudien besser entsprechenden Modellierung leisten. Grundgedanke der Anpassung des geographischen Einflusses in POI-Empfehlungssystemen ist die Berücksichtigung der verkehrstechnischen Anbindung von POIs. Die Erreichbarkeit soll sich demnach aus dem Verkehrsaufwand ergeben, welcher mit dem präferierten Verkehrsmittel überwunden werden muss, um einen bestimmten POI zu besuchen.

Zunächst soll die anzupassende Schnittstelle genauer beschrieben werden. Wie in Kapitel 3 erwähnt, wird die Erreichbarkeit

in bisherigen POI-Empfehlungsansätzen probabilistisch mittels Wahrscheinlichkeitsdichtefunktionen modelliert. Die Besuchswahrscheinlichkeit eines neuen POI l_j wird dementsprechend mithilfe der beispielhaften Algorithmen (1) und (2) ermittelt.

$$\Pr(l_j|S(u)) = \prod_{l_y \in S(u)} a \cdot d(l_j, l_y)^b \quad (1)$$

$$\Pr(l_j|S(u)) = \frac{1}{n} \sum_{y=1}^n \hat{f}(d(l_j, l_y)) \quad (2)$$

Die Funktion (1), welche einen exponentiellen Zusammenhang zur Modellierung nutzt, wird im Ansatz von Ye (2011) verwendet, während sich Zhang & Chow (2013) der Kerndichteschätzung zur Ermittlung der Check-in-Wahrscheinlichkeit eines neuen POIs bedienen. Entsprechend den Ausführungen des Kapitels 2 stellt $S(u)$ die Menge aller vom Nutzer u besuchten POIs l_y dar. Die entscheidende Variable ist die euklidische Distanz d vom neuen POI l_j zu allen besuchten POIs l_y . Um den Verkehrsaufwand anstelle der euklidischen Distanz als Maß der Erreichbarkeit einzubeziehen, ist also der Algorithmus zur Ermittlung der Größe $d(l_j, l_y)$ anzupassen.

Um den im Weiteren vorgestellten Modellierungsansatz verständlich zu machen, soll zunächst der grundlegende Unterschied zwischen dem bestehenden und dem angestrebten Interpretationsansatz der

Erreichbarkeit verdeutlicht werden: Während die euklidische Distanz den geographischen Zusammenhang zwischen mehreren Standorten beschreibt, gibt die verkehrstechnische Anbindung Auskunft über deren topologische Beziehung. Verkehrsaufwände werden in der Geoinformatik bereits bei der Suche nach dem kürzesten Weg in Navigationssystemen als entscheidendes Maß verwendet. Als Datengrundlage werden Verkehrsnetze in Form von Graphen verwendet. Die Knoten stellen georeferenzierte Schnittpunkte, wie Haltestellen oder Kreuzungen, und die gewichteten Kanten Verkehrsaufwände zwischen den Knoten dar. Bei der Suche nach dem optimalen Pfad wird die Route im Graph von einem Start- zu einem Zielknoten ermittelt, welche die kürzeste Wegelänge oder Zeitdauer aufweist. Um dieses Problem effizient zu lösen, werden Routing-Algorithmen verwendet.

Eine derartige Lösung des Problems des kürzesten Pfads, auch als Shortest-Path-Problem bekannt, wird auch im Rahmen der Bewertung der Erreichbarkeit in POI-Empfehlungssystemen als zielführend eingestuft. Sie hätte im Fallbeispiel in Abbildung 2 zur Folge, dass, unter alleiniger Betrachtung des Erreichbarkeitseinflusses, POI B bevorzugt empfohlen werden würde. Wie Kapitel 4 zeigt, erscheint dieses Empfehlungsverhalten weitaus besser nachvollziehbar

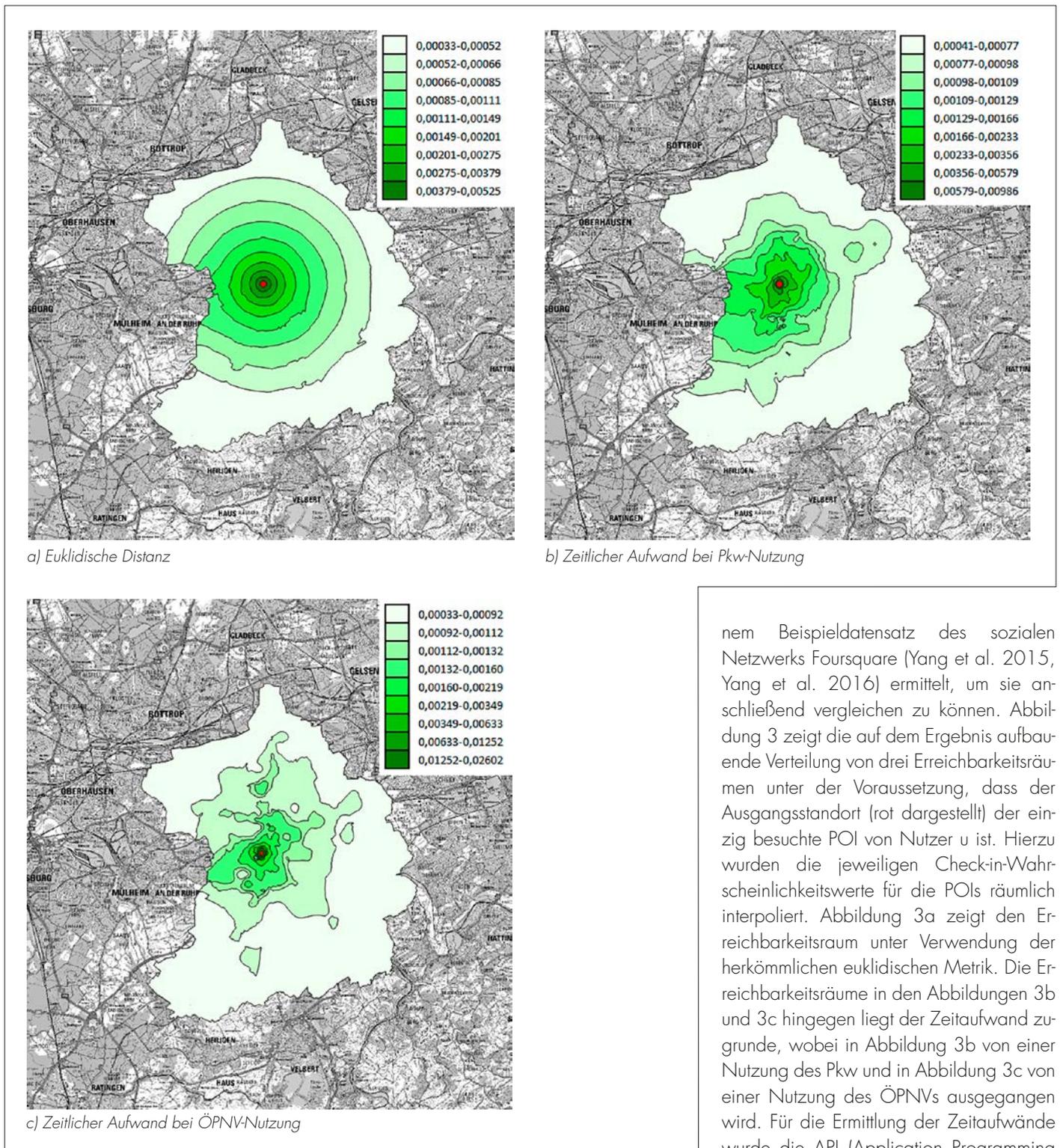


Abbildung 3: Erreichbarkeitsräume unter Verwendung verschiedener Algorithmen und Datengrundlagen (DTK100: Land NRW (2019) Datenlizenz Deutschland – Namensnennung – Version 2.0 (www.govdata.de/dl-de/by-2-0))

als die durch die reine Berücksichtigung der euklidischen Distanz entstehende präferierte Empfehlung von POI A.

Um den Effekt dieser alternativen Methode zur Modellierung des Erreichbarkeitseinflusses an einem umfangreicheren Beispiel zu verdeutlichen, wurde der reprä-

sentative POI-Empfehlungsansatz *Rank-GeoFM* (Li et al. 2015) nach dem vorgestellten Konzept modifiziert. Mithilfe des herkömmlichen und des modifizierten Ansatzes wurde die Distanz d von einem Ausgangs-POI in der Essener Innenstadt zu allen anderen POIs im Stadtgebiet aus ei-

nem Beispieldatensatz des sozialen Netzwerks Foursquare (Yang et al. 2015, Yang et al. 2016) ermittelt, um sie anschließend vergleichen zu können. Abbildung 3 zeigt die auf dem Ergebnis aufbauende Verteilung von drei Erreichbarkeitsräumen unter der Voraussetzung, dass der Ausgangsstandort (rot dargestellt) der einzig besuchte POI von Nutzer u ist. Hierzu wurden die jeweiligen Check-in-Wahrscheinlichkeitswerte für die POIs räumlich interpoliert. Abbildung 3a zeigt den Erreichbarkeitsraum unter Verwendung der herkömmlichen euklidischen Metrik. Die Erreichbarkeitsräume in den Abbildungen 3b und 3c hingegen liegt der Zeitaufwand zugrunde, wobei in Abbildung 3b von einer Nutzung des Pkw und in Abbildung 3c von einer Nutzung des ÖPNVs ausgegangen wird. Für die Ermittlung der Zeitaufwände wurde die API (Application Programming Interface) Here verwendet. Je höher die Sättigung der Grünwerte, desto wahrscheinlicher ist es, dass der Nutzer u einen POI in dem entsprechenden Gebiet besucht.

Wie in den Abbildungen zu sehen ist, weichen die Erreichbarkeitsräume deutlich voneinander ab. Während in Abbildung 3a die Check-in-Wahrscheinlichkeit gleichmäßig um den besuchten POI abnimmt, sind die Verteilungen in den Abbildungen 3b und 3c ungleichmäßig. Dies ist

auf die geographisch differenzierte Qualität des Verkehrsangebots zurückzuführen. Das Ergebnis zeigt also deutlich auf, dass die reine Entfernung nur bedingt die tatsächliche Erreichbarkeit von POIs über den Verkehr abbildet. Es ist also davon auszugehen, dass unter Berücksichtigung der Verkehrsaufwände signifikante Verbesserungen hinsichtlich der Qualität von POI-Empfehlungen zu erwarten sind.

6 HERAUSFORDERUNGEN BEI DER ALTERNATIVEN ERREICHBARKEITSERMITTLUNG

Trotz der Vorteile der vorgeschlagenen Methode birgt die alternative Modellierung des Erreichbarkeitseinflusses einige Herausforderungen, die ebenfalls im Rahmen der Geoinformatik zu lösen sind:

- ▶ **Datengrundlage.** Zur Berechnung der euklidischen Distanz zwischen zwei Punkten müssen ausschließlich deren Koordinaten bekannt sein. Diese sind in ortsbezogenen sozialen Netzwerken unmittelbar verfügbar. Bei der Ermittlung des Verkehrsaufwands sind zusätzlich Verkehrsnetze notwendig, um die optimalen Routen zu berechnen. Diese sind aus externen Datenquellen zu beziehen. Zudem sollten die Daten im Gültigkeitsbereich des sozialen Netzwerks flächendeckend und in ausreichender Qualität verfügbar sein.
- ▶ **Komplexität der Algorithmen.** Algorithmen zur Berechnung der euklidischen Distanz sind relativ simpel und zeiteffizient. Routing-Algorithmen hingegen sind weitaus komplexer.
- ▶ **Schätzung der individuellen Verkehrsmöglichkeiten.** Während die Ermittlung der euklidischen Distanz personenneutral erfolgen kann, ist dies bei der Ermittlung der topologischen Distanz nicht möglich. Wie bereits deutlich wurde, ist die Nutzung des verkehrsinfrastrukturellen Angebots von den Verkehrsmöglichkeiten und -präferenzen des Individuums abhängig. So ist ein Zielstandort von einem Ausgangsstandort nicht mit jedem Verkehrsmittel gleich gut erreichbar. Da die Verkehrspräferenzen nicht explizit von gängigen ortsbezogenen sozialen Netzwerken erhoben und etwa im Nutzerprofil nachgehalten werden, scheint eine implizite Schätzung der individuellen Verkehrsmöglichkeiten notwendig zu sein.

Da bekannte ortsbezogene soziale Netzwerke weltweit gültig sind, ist leicht nachzuvollziehen, dass eine flächendeckende und qualitativ ausreichende Verfügbarkeit der benötigten Daten nicht gewährleistet werden kann. Initiativen, wie OpenStreetMap auf der Seite von Open Source oder die Infrastructure for Spatial Information in the European Community (INSPIRE) auf der amtlichen Seite, beschäftigen sich jedoch bereits mit der landes- und staatsübergreifenden interoperablen Bereitstellung von Geodaten. Das ohnehin erfolgte Tracking der mobilen Endgeräte, mit denen ortsbezogene soziale Netzwerke genutzt werden, bietet einen zusätzlichen Zugang zu wertvollen Daten über Verkehrsnetze und -aufwände.

Der wohl schwierigste Punkt scheint die Schätzung der Verkehrspräferenzen des einzelnen Nutzers. Nach Jürgens & Kasper (2006) überwinden die meisten Personen den Weg zu nahegelegenen Zielen bevorzugt fußläufig. Schwieriger wird die Schätzung bei fernerer Zielen. Auch dort kann die Geoinformatik unterstützen. Laut Jürgens & Kasper (2006) werden die Verkehrsmöglichkeiten stark von den raumstrukturellen Begebenheiten des Wohnstandorts bestimmt. So ist die Nutzung des ÖPNVs in dicht und kompakt bebauten, nutzungsgemischten Gebieten, welche hauptsächlich in Innenstädten auftreten, weitaus verbreiteter als in homogenen Strukturen in ländlichen Gebieten. In diesen wird der Pkw bevorzugt. In der Geoinformatik werden Landschaftsstrukturmaße (Hennecke & Walz 2019) genutzt, um die Raumstruktur von Landschaften zu quantifizieren. Durch die Berechnung von Maßen wie der Punktdichte von Stationen des öffentlichen Verkehrs oder Diversitätsindizes können also Verkehrspräferenzen abgeleitet werden. Diese Schätzung kann durch den Einbezug von demografischen Informationen aus dem Nutzerprofil, welche ebenfalls einen Einfluss von Verkehrsmöglichkeiten haben, vermutlich weiter verbessert werden.

7 FAZIT UND AUSBLICK

Es zeigt sich, dass die Geoinformatik eine erweiterte und notwendige Perspektive im Hinblick auf die Optimierung der Empfehlungsqualität im sehr datenwissenschaftlich geprägten Feld der POI-Empfehlungssysteme liefern kann. Die wichtigste Erkenntnis stellt in dem Zusammenhang die Interpreta-

tion der Erreichbarkeit als topologisches Problem dar. So ist zu konstatieren, dass der Verkehrsaufwand, welcher sich mithilfe von Routing-Algorithmen ermitteln lässt, die Erreichbarkeit von Orten besser widerspiegelt als die euklidische Distanz. Es wird davon ausgegangen, dass die alternative Erreichbarkeitsmetrik insbesondere POI-Empfehlungen im lokalen Umfeld des Zielnutzers optimieren kann.

Zu erwähnen ist, dass die Reisedauer als Erreichbarkeitsmetrik bereits im Ortsempfehlungsansatz von Levandoski et al. (2012) berücksichtigt wurde. Durch den vorliegenden Beitrag wird dieses Vorgehen wissenschaftlich durch Mobilitätsstudien bestärkt. Andererseits wird durch die herangezogenen Mobilitätsstudien und den aufgeführten Vergleich der Erreichbarkeitsräume unter Nutzung verschiedener Verkehrsmittel deutlich, dass für eine zuverlässige Ermittlung der Reisedauer die individuellen Verkehrsmöglichkeiten der Nutzer zu berücksichtigen sind. Diesen Aspekt beachten Levandoski et al. (2012) nicht explizit. Der vorliegende Beitrag gibt darüber hinaus Aufschluss, wie die Schätzung der Verkehrsmöglichkeiten durch die Nutzung von Landschaftsstrukturmaßen erfolgen kann und welche Möglichkeiten aus Sicht der Geoinformatik bestehen, qualitativ ausreichende Graphen zur Ermittlung der topologischen Distanz zu nutzen.

Ob die Realisierung des vorgestellten Konzepts State-of-the-Art-Ansätze hinsichtlich der Empfehlungsqualität übertreffen kann, ist weitergehend zu evaluieren. Hierbei wird empfohlen, auf gängige Evaluierungsmethodiken und -metriken, wie sie in den Studien von Shani & Gunawardana (2011) oder Liu et al. (2017) vorgestellt werden, zurückzugreifen. Es ist davon auszugehen, dass sich die Qualität der POI-Empfehlungen vor allem mit der flächendeckenden Verfügbarkeit qualitativ hochwertiger Geodaten sowie mit der Schätzbarkeit der Verkehrsmöglichkeiten verbessern lässt (Maas 2019). Mit Blick auf den letztgenannten Aspekt wird vor allem die Berücksichtigung raumstruktureller Parameter sowie demografischer Informationen aus dem Nutzerprofil als zielführend erachtet.

Als weiterer Aspekt, welcher sich positiv auf die POI-Empfehlungsqualität auswirken könnte, sei die Erweiterung der empfohlenen Orte von punktförmigen Objekten auf Geometrien jeglicher Art genannt. Dies

kann beispielsweise bei großräumigeren Örtlichkeiten, wie zum Beispiel Einkaufszentren oder Parks, welche in ortsbezogenen sozialen Netzwerken lediglich als POI repräsentiert werden, hilfreich sein. Durch

eine flächenhafte Repräsentation (z. B. ein „Polygon of Interest“) ließe sich hier eine zutreffendere Aussage über die Erreichbarkeit von Standorten unter Verwendung der aufgeführten Methodiken treffen.

DANKSAGUNG

Die Autoren danken Prof. Dr.-Ing. Edmund Coersmeier vom Labor für Angewandte Informatik der Hochschule Bochum für seine fachlichen Anregungen.

Literatur

- Beckmann, K. J.; Hesse, M.; Holz-Rau, C.; Hunecke, M. (Hrsg.) (2006): *StadtLeben – Wohnen, Mobilität und Lebensstil*. VS Verlag für Sozialwissenschaften, Wiesbaden.
- Cho, E.; Myers, S. A.; Leskovec, J. (2011): Friendship and mobility: User movement in location-based social networks. In: *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining*, S. 1082-1090.
- Dick, M. (Hrsg.) (2009): *Mobilität als Tätigkeit: individuelle Expansion – alltägliche Logistik – kulturelle Kapazität*. Pabst Science Publishers, Lengerich.
- Hasan, S.; Zhan, X.; Ukkusuri, S. V. (2013): Understanding urban human activity and mobility patterns using large-scale location-based data from online social media. *2nd ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, 11.-14.08.2013, Chicago, IL.
- Hennecke, D.; Walz, U. (2019): Landschaftsstrukturmaße – Ein neuer Kursteil in OpenGeoEdu. In: *gis.Science*, 3/2019, S. 110-115.
- Holt, A.; Cotroneo, P.; Singh, R. (2013): Points of interest (POI) standards working group charter. OGC Project Document 12-142r1. Open Geospatial Consortium, Wayland, MA.
- Hunecke, M.; Schweer, I. R. (2006): Einflussfaktoren der Alltagsmobilität – Das Zusammenwirken von Raum, Verkehrsinfrastruktur, Lebensstil und Mobilitätseinstellungen. In: Beckmann, K. J.; Hesse, M.; Holz-Rau, C.; Hunecke, M. (Hrsg.): *StadtLeben – Wohnen, Mobilität und Lebensstil*. VS Verlag für Sozialwissenschaften, Wiesbaden, S. 147-166.
- Jarass, J. (2012): *Wohnstandortpräferenzen und Mobilitätsverhalten: Verkehrsmittelwahl im Raum Köln*. Springer VS College, Wiesbaden.
- Jürgens, C.; Kasper, B. (2006): Alltagsmobilität, Raum und Lebensstile. In: Beckmann, K. J.; Hesse, M.; Holz-Rau, C.; Hunecke, M. (Hrsg.): *StadtLeben – Wohnen, Mobilität und Lebensstil*. VS Verlag für Sozialwissenschaften, Wiesbaden, S. 125-141.
- Levandoski, J. J.; Sarwat, M.; Eldawy, A.; Mokbel, M. F. (2012): IARS: A Location-Aware Recommender System. In: *2012 IEEE 28th International Conference on Data Engineering*, S. 450-461.
- Lian, D.; Zhao, C.; Xie, X.; Sun, G.; Chen, E.; Rui, Y. (2014): GeoMF: Joint Geographical Modeling and Matrix Factorization for Point-of-Interest Recommendation. In: *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, S. 831-840.
- Li, X.; Cong, G.; Li, X.-L.; Pham, T.-A. N.; Krishnaswamy, S. (2015): Rank-GeoFM. In: Baeza-Yates, R.; Lalmas, M.; Moffat, A. & Ribeiro-Neto, B. (Eds.): *Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval – SIGIR '15*, S. 433-442.
- Liu, Y.; Pham, T.-A. N.; Cong, G.; Yuan, Q. (2017): An experimental evaluation of point-of-interest recommendation in location-based social networks. In: *Proceedings of the VLDB Endowment*, 10 (10), S. 1010-1021.
- Maas, F. (2019): *Integration von Faktoren der räumlichen Mobilität in Point-of-Interest-Empfehlungssystemen in Location-based Social Networks*. Masterarbeit (M. Eng. Geoinformatik), Hochschule Bochum.
- POI Working Group (2010): *World Wide Web Consortium (W3C)*. <https://www.w3.org/2010/POI/> (Zugriff 11/2019).
- Schmidt, B.; Maas, F. (2020): *Mathematical description of the point-of-interest recommendation problem*. Report no. 19-01, Bochum University of Applied Sciences, Geovisualization Lab. https://www.researchgate.net/publication/337604939_Mathematical_Description_of_the_Point-of-Interest_Recommendation_Problem (Zugriff 01/2020).
- Shani, G.; Gunawardana, A. (2011): Evaluating Recommendation Systems. In: Ricci, F.; Rokach, L.; Shapira, B. & Kantor, P. B. (Eds.): *Recommender Systems Handbook*. Springer, Boston, MA, S. 257-297.
- Yang, D.; Zhang, D.; Chen, L.; Qu, B. (2015): NationTelescope: Monitoring and visualizing large-scale collective behavior in LBSNs. In: *Journal of Network and Computer Applications*, 55, S. 170-180.
- Yang, D.; Zhang, D.; Qu, B. (2016): Participatory Cultural Mapping Based on Collective Behavior Data in Location-Based Social Networks. In: *ACM Transactions on Intelligent Systems and Technology*, 7 (3), S. 1-23.
- Ye, M.; Yin, P.; Lee, W.-C.; Lee, D.-L. (2011): Exploiting geographical influence for collaborative point-of-interest recommendation. *34th ACM SIGIR Conference on Research and Development in Information Retrieval*, 24.-28.07.2011, Beijing, China, S. 325-334.
- Yu, Y.; Chen, X. (2015): A Survey of Point-of-Interest Recommendation in Location-Based Social Networks. In: *Trajectory-Based Behavior Analytics, Papers from the 2015 AAAI Workshop*, S. 53-60.
- Zhang, J.-D.; Chow, C.-Y. (2013): iGSLR – Personalized geo-social location recommendation. In: *Proceedings of the 21st ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, S. 334-343.
- Zhao, S.; King, I.; Lyu, M. R. (2016): A survey of point-of-interest recommendation in location-based social networks. <https://arxiv.org/pdf/1607.00647v1.pdf> (Zugriff 11/2019).