

Institut für Formale Methoden der Informatik
Universität Stuttgart
Universitätsstraße 38
D-70569 Stuttgart

Studienarbeit Nr. 2356

Untersuchung der Struktur großer Straßennetzwerke

Frederik Hartmann

Studiengang: Informatik
Prüfer: Prof. Dr. Stefan Funke
Betreuer: Prof. Dr. Stefan Funke

begonnen am: 1. November 2011
beendet am: 2. Mai 2012

CR-Klassifikation: F.2.2, K.1

Inhaltsverzeichnis

1. Einleitung	4
2. Transformation der OSM Karte	5
2.1. Anforderungen der Algorithmik an die Graphenstruktur	5
2.2. OSM Graphstruktur	5
2.2.1. Knoten	6
2.2.2. Wege und Kanten	6
2.2.3. Relationen	6
2.3. Transformationsverfahren	7
2.3.1. Import	7
Knoten	8
Wege	8
2.3.2. Export	9
2.4. Bereinigung	11
2.5. Laufzeit	11
2.6. Ergebnis	11
3. Bestimmung der starken Zusammenhangskomponenten	13
3.1. Erwartete Struktur	13
3.2. Vorgehensweise zur Bestimmung der SZHK	14
3.3. Analyse	15
3.4. Ergebnis	17
4. Geometrische Spannereigenschaften	18
4.1. Erläuterung des Konzepts	18
4.2. Adaption	18
4.3. Naiver Berechnungsansatz	19
4.4. Verbesserungsideen	19
4.5. Schätzung ausgewählter Kennzahlen des Graphen	21
4.5.1. Mittelwert und Nachbarknoten	21
4.5.2. Verteilung der Wege	22
4.5.3. Klassifikation der Wege	23
4.5.4. Verteilung der Knoten	25
4.6. Ermittlung der Anbindung	26
4.6.1. Algorithmus	26
4.6.2. Knotenreduktion	26
4.6.3. Clustering	27

4.6.4. Laufzeit	28
4.7. Ergebnis	29
5. Zusammenfassung & Ausblick	31
A. Anhang	32
A.1. Datenbankdesign	32
A.1.1. Tabelle:nodes	32
A.1.2. Tabelle:node_tags	32
A.1.3. Tabelle:way_tags	32
A.1.4. Tabelle:way_splits	32
A.1.5. Tabelle:h_ways	33
A.2. Tabelle der Anzahl und Größe der starken Zusammenhangskomponenten . .	33
A.3. Verteilung der Wege aus der Stichprobe	35
Literaturverzeichnis	42

1. Einleitung

In den letzten Jahren hat die Bedeutung von Geoinformationen durch das Aufkommen von navigationsfähigen Smartphones und personalisierter Werbung stark zugenommen. Diese Anwendungsbereiche, aber auch intelligentes Routing oder Verbesserungen im Straßenbau, benötigen exakte, detaillierte Karten, die Straßendaten mit Points of Interests oder Stauinformationen vernetzen.

Im Jahr 2004 wurde durch die Gründung von OpenStreetMap eine Plattform für eine Open Source Straßenkarte geschaffen, an der sich jeder mit einem GPS Empfänger oder GPS fähigem Smartphone beteiligen kann.

Der Open Source Ansatz kann jedoch auf Grund fehlender Organisationsstrukturen auch ein Problem sein. Dies lässt sich alleine an 199 unterschiedlichen Straßentypen erkennen, von denen jedoch nur 20-30 Typen weitere Verbreitung erfahren. Auch werden viele Straßen uneinheitlich eingepflegt. Dennoch ist die Qualität der OSM bereits heute in manchen Bereichen gleichwertig oder besser als kommerzielle Alternativen.[DZ10]

Um die OpenStreetMap algorithmisch verwenden zu können, muss zunächst eine Transformation und eine Bereinigung der Karte durchgeführt werden, da das Ursprungsformat nicht für die algorithmische Bearbeitung geeignet ist und viele nicht benötigte Informationen enthält.

Die vorliegende Studienarbeit beschäftigt sich im Kapitel 2 mit der Umwandlung in ein algorithmisch gut verwendbares Datenformat und der Bereinigung des Graphen.

In den Kapiteln 3 und 4 geht es um die Ermittlung von grundlegenden Eigenschaften des Graphen, die zur Optimierung und Laufzeitabschätzung von Algorithmen benötigt werden. Durch die Analyse der Verzerrung der Weglängen und Distanzen werden Hinweise auf problematische Gebiete im Graphen gegeben.

2. Transformation der OSM Karte

2.1. Anforderungen der Algorithmik an die Graphenstruktur

Um Graphdaten in der Algorithmik verwenden zu können, ist ein schlankes und klar definiertes Datenformat notwendig. Dieses Datenformat sollte unabhängig von der Rechnerarchitektur und der verwendeten Programmiersprache sein, damit es ohne weitere Transformationen oder Bearbeitung gelesen werden kann. Auf Grund der unterschiedlichen Darstellung von Datentypen in verschiedenen Programmiersprachen sollte kein binäres Format verwendet werden. Bei der Größe der Graphen von einigen hundert Megabyte bis zu mehreren Gigabyte sind XML Datenstrukturen ebenfalls problematisch. Diese lassen sich nur unter größerem Zeit- und Platzaufwand parsen und verschwenden auf Grund der Struktur zu viel Platz.

Durch die obigen Anforderungen und dem Wunsch, Graphen unkompliziert austauschen zu wollen, hat sich in der Abteilung Algorithmik am Institut der Theoretischen Informatik das unten vorgestellte Plaintextformat durchgesetzt, das auch in der vorliegenden Studienarbeit verwendet wird.

```
1 N                               #Anzahl der Knoten
2 M                               #Anzahl der Kanten
3 N*(ID osmID lat lon Elevation) #Menge der Knoten
4 M*(SourceID DestinationID Distance) #Menge der Kanten
```

Die zugehörigen Datenformate(in Stringrepräsentation):

```
1 Integer
2 Integer
3 Integer Long Double Double Integer
4 Integer Integer Integer
```

Die Kanten sind implizit gerichtet und dürfen mehrdeutig sein, ein Multigraph ist also zulässig. Nach den oben angegebenen Knoten und Kantenattributen können noch weitere Eigenschaften hinzugefügt werden, wie z.B. die Maximalgeschwindigkeit oder der Kantentyp.

2.2. OSM Graphstruktur

Die am einfachsten zu verarbeitenden Daten des OSM Projektes liegen in einem XML Format vor, das entweder komplett oder aufgeteilt in verschiedenen Regionen zur Verfügung steht

[OSM12b]. Das Datenformat dieses Exportes wird im folgenden Abschnitt erklärt[OSM12a]. Die verwendeten Snippets stammen aus dem OSM Export des Saarlands.

2.2.1. Knoten

Positionen werden durch den Tag "node" definiert. Sie stellen die Ankerpunkte für alle anderen Typen dar.

```

1 <node id="1488862382" lat="49.2332313" lon="7.0439174" version="1" changeset="9715446"
  user="Anakin78" uid="547619" timestamp="2011-11-01T18:55:55Z">
2   <tag k="name" v="Netto" />
3   <tag k="shop" v="supermarket" />
4 </node>
```

Die Informationen, die jeder Knoten enthält (Zeile 1), sind die GPS Koordinaten, eine eindeutige ID und Informationen zur letzten Bearbeitung des Knoten. Um weitergehende Informationen an den Knoten anzuhängen, werden Tags hinzugefügt. Diese Tags(Z. 2-3) haben immer zwei Attribute, das Attribut "k" definiert die Eigenschaft, das Attribut "v" den Wert der Eigenschaft.

2.2.2. Wege und Kanten

Wege werden durch den Tag "way" definiert, Kanten kennt die OSM Datenstruktur nicht als eigenständigen Tag, diese werden implizit durch die Liste der Knoten des Weges definiert.

```

1 <way id="28957912" version="1" changeset="356695" uid="35829" user="tower27"
  timestamp="2008-12-13T17:19:33Z">
2   <nd ref="318392781"/>
3   <nd ref="318392783"/>
4   <nd ref="318392785"/>
5   <tag k="highway" v="residential" />
6   <tag k="name" v="Lortzingweg" />
7   <tag k="oneway" v="yes" />
8 </way>
```

Jeder Weg enthält als Attribut eine eindeutige ID und Informationen zur letzten Änderung. Die enthaltenen "nd" Elemente(Z. 2-4) stellen die Liste der Knoten dar. Tags fügen weitere Informationen wie den Straßentyp, den Namen oder Informationen zur Kennung im "Traffic Message Channel"(Staudaten) zum Weg hinzu. Die im Normalfall ungerichteten Kanten bestehen immer zwischen einem Knoten und seinem Nachfolger.

Wege sind in der OSM Struktur nicht nur als physische Wege zu verstehen, sie werden auch zur Abgrenzung von Gebäuden, Feldern oder anderen Gebieten bzw. Linien verwendet.

2.2.3. Relationen

Relationen werden durch den Tag "relation" definiert. Sie werden verwendet, um Knoten und Wege zu größeren Einheiten zusammenzufassen.

```

1 <relation id="1161610" version="1" changeset="5672498" uid="178396" user="the_smile"
  timestamp="2010-09-03T17:28:24Z">
2   <member type="way" ref="75767992" role="" />
3   <member type="way" ref="75767993" role="" />
4   <member type="way" ref="75767994" role="" />
5   <tag k="admin_level" v="8" />
6   <tag k="boundary" v="administrative" />
7   <tag k="name" v="Merten" />
8   <tag k="ref:INSEE" v="57460" />
9   <tag k="type" v="boundary" />
10 </relation>

```

Relationen stellen insoweit eine Sonderform dar, als dass sie aus Wegen und Knoten in beliebiger Vermischung bestehen können. Sie werden verwendet, um beispielsweise Länder, Bundesländer und Landkreise, aber auch um Routen des öffentlichen Nahverkehrs zu definieren. Sie haben für das physische Straßennetz nur in so weit eine Bedeutung, als dass über sie in manchen Fällen Abbiegerestriktionen definiert werden. Diese Fälle werden jedoch auf Grund der Seltenheit in dieser Arbeit ignoriert.

2.3. Transformationsverfahren

Um die XML Struktur in das im Abschnitt 2.1 definierte Graphformat zu übersetzen und dennoch alle Informationen schnell im Zugriff zu haben, wurden alle Knoten und alle Wege in eine MySQLTM Datenbank importiert. Für die reine Transformation in das Straßenformat wäre dies nicht nötig, jedoch bietet eine SQL Datenbank die Möglichkeit, auf nahezu alle Informationen des OSM Graph zuzugreifen, ohne spezielle Programme zu schreiben oder anzupassen. Der Export von Points of Interests oder unterschiedlichen Wegetypen ist direkt durch eine Abfrage der Datenbank oder durch die Änderung eines SQL-Befehls möglich.

2.3.1. Import

Auf Grund der Größe der Importdateien wurde beim Import ein SAX Parser verwendet. Dieser XML Parser baut keinen Objektbaum auf, sondern parst Elemente sequentiell nacheinander. Durch den bedeutend geringeren Arbeitsspeicherverbrauch im Vergleich zu einem DOM Parser ist ein Import auch mit normalen PCs möglich.

Der Import findet in eine MySQLTM Datenbank statt, die nicht auf dem Rechner laufen muss, der das Import Programm ausführt. Als Programmiersprache wurde Java gewählt, der Datenbankszugriff findet über die von Oracle bereitgestellten Bibliotheken statt. Durch die Zusammenfassung von 10.000 INSERT Statements zu einem Statement wurde eine ausreichende Performance sichergestellt. Außerdem wurden alle Indizes beim Importvorgang deaktiviert, da die einmalige Berechnung am Ende bedeutend schneller ist.

Knoten

Knoten können ohne weitere Verarbeitung in die Datenbank übernommen werden. Die Knotenelemente werden direkt in die Tabelle "nodes" importiert, ihre zugehörigen Tags werden, abgesehen der "created_by" Tags, in die Tabelle "node_tags" importiert.

Tabelle "nodes"

Name	Datentyp	Kommentar
osmID	bigint(20)	Eindeutige OSM ID des Knoten
lat	Double	Breitengrad in Dezimalgrad
lon	Double	Längengradgrad in Dezimalgrad

Tabelle "node_tags"

Name	Datentyp	Kommentar
nodeID	bigint(20)	Eindeutige OSM ID des zugehörigen Knoten
k	varchar(64)	Attribut
v	varchar(256)	Attributswert

Wege

Wege werden vor dem Import direkt in die entsprechenden Kanten zerlegt. Dazu ist die Unterscheidung in gerichtete und ungerichtete Wege wichtig. Gerichtete Wege enthalten entweder einen Tag mit den Attributen k="oneway" und v="yes" bzw v="true" oder sind Bestandteil eines Kreisverkehrs. Dies wird durch den Tag k="highway" und v="turning_circle" definiert. Wie bereits im Abschnitt 2.2 beschrieben, enthält jeder Weg eine Liste aus Knoten, die den Verlauf des Weges definieren. Bei einem gerichteten Weg wird eine Kante von Vorgänger zu Nachfolger angelegt, bei einem ungerichteten Weg noch eine Kante von Nachfolger zu Vorgänger. Zu beachten ist, dass in diesem Schritt in der Tabelle "way_tags" und "way_splits" noch Wege enthalten sind, die keiner realen Straße entsprechen.

Tabelle "way_splits"

Name	Datentyp	Kommentar
wayID	bigint(20)	Eindeutige OSM ID des Weges
sourceID	bigint(20)	OSM ID des Startknotens
destinationID	bigint(20)	OSM ID des Zielknotens

Tabelle "way_tags"

Name	Datentyp	Kommentar
wayID	bigint(20)	Eindeutige OSM ID des zugehörigen Weges
k	varchar(64)	Attribut
v	varchar(256)	Attributswert

Um aus diesen Tabellen alle befahrbaren Straßen zu exportieren, ist besonders der Tag mit dem Attribut `k="highway"` interessant. Wege mit einem Tag aus dieser Kategorie entsprechen einem physisch existenten Straßenabschnitt.

Um die Befahrbarkeit einer Straße genauer zu definieren, können über den Tag `k="access"` und `v="yes" / "no" / "agricultural" etc. . .` weitere Einschränkungen vorgenommen werden. Zur Vorbereitung des Exportes werden alle befahrbaren, physischen Straßen in eine extra Tabelle übernommen, um bei Abfragen und dem Export nur mit den realen Straßen arbeiten zu müssen. Der Export erfolgt über `"highway"` Eigenschaft, die eine notwendige Bedingung für die Befahrbarkeit einer Straße darstellt. Dies verringert die Menge der Kanten im Deutschlandgraph bereits von über 200 Millionen zu 72 Millionen Kanten.

Tabelle: "h_ways"

Name	Datentyp	Kommentar
sourceOSM	bingint(20)	Eindeutige OSM ID des Startpunktes
destinationOSM	bingint(20)	Eindeutige OSM ID des Zielpunktes
slat	Double	Latitude des Startpunktes
slon	Double	Longitude des Startpunktes
dlat	Double	Latitude des Zielpunktes
dlon	Double	Longitude des Zielpunktes
type	String(32)	Straßentyp

2.3.2. Export

Die "h_ways" Tabelle enthält momentan alle vorhandenen Straßentypen inklusive Radwege, Wanderwege und andere, nicht befahrbare, Straßen. Durch eine Liste von zu exportierenden Typen kann genau gesteuert werden, welche Wege exportiert werden sollen. Im Normalfall möchte man alle von Kraftfahrzeugen befahrbare Straßen exportieren, jedoch sind auch Wanderkarten und Radkarten mit geringen Anpassungen exportierbar. Die folgende Tabelle enthält alle von Kraftfahrzeugen befahrbare Wegtypen mit ihren implizierten Richtungen.

v	oneway	Kommentar
motorway	Ja	Autobahn
motorway_link	Ja	Autobahnauf- oder abfahrt
trunk	Ja	Autobahnähnliche Straße
trunk_link	ja	Auf/Abfahrt zu einer autobahnähnlichen Straße
primary	Nein	Bundesstraße
primary_link	Nein	Bundesstraßenanschlussstelle
secondary	Nein	Land oder Kreisstraße
secondary_link	Nein	Auf/Abfahrt zu einer Land oder guten Kreisstraße
tertiary	Nein	Kreisstraße oder innerstädtische Vorfahrtsstraße
tertiary_link	Nein	Verbindungsstücke zu Kreisstraßen
unclassified	Nein	Diverse, z.B. schlecht ausgebaute Gemeindestraßen
residential	Nein	Straße in Wohngebieten
living_street	Nein	Verkehrsberuhigter Bereich
road	Nein	Straße unbekannter Klassifikation
service	Nein	Erschließungsweg
turning_circle	Ja	Kreisverkehr

Sobald die Straßentypen definiert sind, werden entweder alle entsprechenden Kanten oder nur die Kanten eines bestimmten Bereiches in die Tabelle "e_ways" kopiert, welche von der Struktur identisch zu "h_ways" ist. Ist dieses abgeschlossen, werden alle Knoten, die entweder als Start oder Ziel in der Tabelle vorhanden sind, in die Tabelle "e_nodes" kopiert und neu indiziert, um einen durchgehenden Zahlenraum der Knoten sicherzustellen. Wie auch die Tabelle "h_ways" enthält diese Tabelle die GPS Koordinaten des Start- und Zielpunktes, um eine direkte Distanzberechnung zu ermöglichen.

Tabelle "e_nodes"

Name	Datentyp	Kommentar
ID	int(11)	Eindeutige neue ID
osmID	bingint(20)	Eindeutige OSM ID
lat	Double	Latitude des punktes
lon	Double	Longitude des Punktes

Sind beide Tabellen fertig aufgebaut, werden sie mittels eines hierfür entwickelten Javaprogramms exportiert. Da die OSM Daten keine Höhendaten erhalten, diese aber oft ebenfalls benötigt werden, werden die Daten noch durch Fernerkundungsdaten der "Shuttle Radar Topography Mission (SRTM)" der NASA aufbereitet.[NASoo]

Die OSM Daten beinhalten ebenfalls keine Distanz, sondern nur GPS Koordinaten, welche noch zu Distanzen umgerechnet werden müssen.

Auf Grund der kurzen Distanzen und der Diskretisierung der Weglängen ist eine Berechnung des Distanz mit der Formel zur Entfernungsberechnung auf Kugeloberflächen ausreichend. Werden exaktere Ergebnisse benötigt, muss eine Entfernungsberechnung für Ellipsoide mit einem entsprechenden Referenzellipsoid gewählt werden.

2.4. Bereinigung

Der exportierte Graph enthält jetzt immer noch Knoten und Kanten, die zwar physisch vorhanden sind, jedoch keine Verbindung zur größten Zusammenhangskomponente besitzen (Deutschen Inseln, Rennstrecken, etc. . .). Je nach exportierten Wegtypen kann der Graph aber auch Abschnitte enthalten, dessen Anbindung an das Straßennetz nicht zum Export ausgewählt wurden (Parkplätze im Wald, Firmengebiete, . . .).

Um Algorithmen ohne die Betrachtung von Spezialfällen verwenden zu können, ist oftmals ein Graph, der nur aus einer starken Zusammenhangskomponente besteht, wünschenswert. Um dies zu erreichen, wird mit Hilfe des Algorithmus von Tarjan [Tar72] die größte Zusammenhangskomponente bestimmt und alle enthaltenen Knoten und Kanten in den finalen Graphen exportiert. Um entstehende Lücken zu füllen, werden alle Knoten neu indexiert und die Kanten entsprechend angepasst.

Kapitel 3 beschäftigt sich weitergehend mit den SZHKs des Graphen.

2.5. Laufzeit

Die komplette Transformation des Deutschlandgraphen dauerte auf dem Server der Abteilung "Algorithmik" zwei Stunden und 22 Minuten. Ähnliche Zeiten sollten jedoch auch auf anderen Computern zu erreichen sein, da die Programme maximal zwei Kerne auslasten können und auch mit 8 Gigabyte Arbeitsspeicher auskommen. Die Laufzeit verteilt sich wie folgt auf die einzelnen Arbeitsschritte:

Programm	Dauer
Import	109 Minuten
Export	33 Minuten
Gesamt	142 Minuten

Da man den Import in die Datenbank nur vornimmt, um den Datenbestand anzulegen oder zu aktualisieren, ist die relevante Größe die Dauer des Exportes. Die Aktualisierung des Datenbestandes lässt sich ohne Probleme automatisiert durchführen, so dass man jederzeit aus einem aktuellen Datenbestand schöpfen kann. Durch die geringe Exportzeit von 33 Minuten kann man sehr kurzfristig auf geänderte Anforderungen eingehen, das Erstellen oder Anpassen von weiteren Umwandlungsprogrammen für neue Graphen kann so vermieden werden.

2.6. Ergebnis

Abschließend kann man sagen, dass das Transformationsverfahren einen guten Kompromiss zwischen Rechenzeit und Wartungsaufwand darstellt. Durch den Import in die SQL

Datenbank erkaufte man sich durch einen langsamen, aber selten auszuführenden Import auf Kosten von Rechenzeit die Möglichkeit, ohne größeren Arbeitsaufwand verschiedenste Graphen schnell zu exportieren und auch auf alle anderen Informationen der OSM Karte effizient zuzugreifen. Beispielsweise kann der Export der Position aller Supermärkte und Bäckereien durch eine einfache SQL Abfrage durchgeführt werden und benötigt nur wenige Sekunden.

Durch den Export von Ausschnitten lassen sich auch für Tests und Laufzeitanalysen von Algorithmen maßgeschneiderte, aber dennoch realistische Graphen ohne größeren Zeitaufwand erzeugen.

3. Bestimmung der starken Zusammenhangskomponenten

In einem optimalen, fehlerfreien Straßengraph würde man für jeden geographisch zusammenhängenden Bereich exakt eine starke Zusammenhangskomponente(SZHK) erwarten. Durch die Entstehungsweise der OpenStreetMap Karten sind jedoch Fehler zu erwarten. Außerdem könnten durch Fehler in der Transformation Gebiete entstehen, die keinen Anschluss zum Straßennetz haben.

Die SZHK Analyse dient so zum einen als ein Gütekriterium des Graphen, zum anderen auch zur Kontrolle des Transformationsverfahrens.

Sollte der Graph durch einen Export mit eingeschränkten Wegtypen zersplittert worden sein, könnten die exportierten Kantentypen nicht ausreichen, um einen Zusammenhang sicherzustellen; wir betrachten hier jedoch einen Graphen mit allen in 2.3.2 angegebenen Kantentypen.

3.1. Erwartete Struktur

In dem exportierten Graph erwarten wir, gegeben durch die geographische Lage der Landflächen, folgende Vorkommnisse von SZHKs: Die ersten drei Gruppen entsprechen den in einer Straßenkarte gewünschten starken Zusammenhangskomponenten, die vierte Gruppe von Zusammenhangskomponenten werden dem Entstehungsverfahren der OpenStreetMap geschuldet sein, da hier praktisch keine Kontrolle der Graphdaten stattfindet. Diese Zusammenhangskomponenten sollten als Fehler betrachtet und korrigiert werden.

1. Eine große Zusammenhangskomponente, welche nahezu alle Punkte auf dem Kontinentalgebiet von Deutschland enthält.
2. Einige kleinere Zusammenhangskomponenten(1.000 bis 10.000 Knoten), die die Knoten der deutschen Inseln enthalten (500 bis 10).
3. Kleinere Zusammenhangskomponenten(100 bis 1.000 Knoten) auf dem Kontinentalgebiet, die autonome Gebiete wie z.B. Rennstrecken oder Firmengebiete enthalten.
4. Kleinere und kleinste Zusammenhangskomponenten, die durch Fehler in der Ursprungskarte entstanden sind.

3.2. Vorgehensweise zur Bestimmung der SZHK

Zur Bestimmung der SZHK wird der Deutschlandgraph nach der Transformation vor der Bereinigung verwendet. Mit Hilfe des Algorithmus von Tarjan[Tar72] wird die Zählung der SZHKs durchgeführt.

Dazu wird eine Menge M gebildet, die am Anfang jeden Knoten des Graphen enthält. Durch die Anwendung des Algorithmus auf einen beliebigen Knoten wird die starke Zusammenhangskomponenten des Knoten ermittelt. Diese wird gespeichert und alle enthaltenen Knoten werden aus M entfernt. Dies wird so lange wiederholt, bis M keine Knoten mehr enthält.

Durch den geringen Aufwand von $O(|V| + |E|)$ ist die Laufzeit auch bei extrem großen Graphen unproblematisch.

Das folgende Listing enthält den wesentlichen Teil der erstellten Javaimplementierung des Algorithmus von Tarjan:

```

1 private void tarjan(int id)
2 {
3     int[][] edges;
4     int v;
5
6     dfs[id] = maxDfs;
7     lowlink[id] = maxDfs;
8     maxDfs++;
9     st.push(id);
10    inStack[id] = true;
11    visited[id] = true;
12
13    edges = this.ao.getSuccessors(id);
14
15    for(int i = 0; i < edges[0].length; i++)
16    {
17        if(!visited[edges[0][i]])
18        {
19            tarjan(edges[0][i]);
20            lowlink[id] = Math.min(lowlink[id], lowlink[edges[0][i]]);
21        }
22        else if(inStack[edges[0][i]])
23        {
24            lowlink[id] = Math.min(lowlink[id], dfs[edges[0][i]]);
25        }
26    }
27
28    if(lowlink[id] == dfs[id])
29    {
30        szhk++;
31        v = st.pop();
32        inStack[v] = false;
33        cur.push(v);
34    }

```

```

35         while(v != id)
36         {
37             v = st.pop();
38             inStack[v] = false;
39             cur.push(v);
40         }
41         szkhList.add(cur);
42         cur = new Stack<Integer>();
43     }
44 }

```

3.3. Analyse

Wie erwartet, gibt es eine sehr große starke Zusammenhangskomponente, die fast alle Knoten enthält. Zur übersichtlicheren Darstellung wurden die Zusammenhangskomponenten nach ihrer Größe zusammengefasst. Gruppe 1 enthält alle SZHKs mit der Größe 1, Gruppe 2 enthält alle SZHKs mit der Größe 2 bis 10 und so weiter. Die letzte Gruppe enthält die größte Zusammenhangskomponente.

Größe der SZHK	Anzahl SZHK	Summe enthaltener Knoten
1	14.989	14.989
2-10	13.816	57.784
11-100	2.859	63.756
101-1.000	78	17.372
1.001-10.000	8	16.961
15.959.387	1	15.959.387

Interessanter als die Anzahl der starken Zusammenhangskomponenten ist jedoch die Verteilung der Knoten, welche in dem folgenden Diagramm dargestellt wird. Deutlich ist erkennbar, dass sich fast alle Knoten in der größten starken Zusammenhangskomponente konzentrieren und sich das restliche 1% auf die verbleibenden starken Zusammenhangskomponenten verteilt.

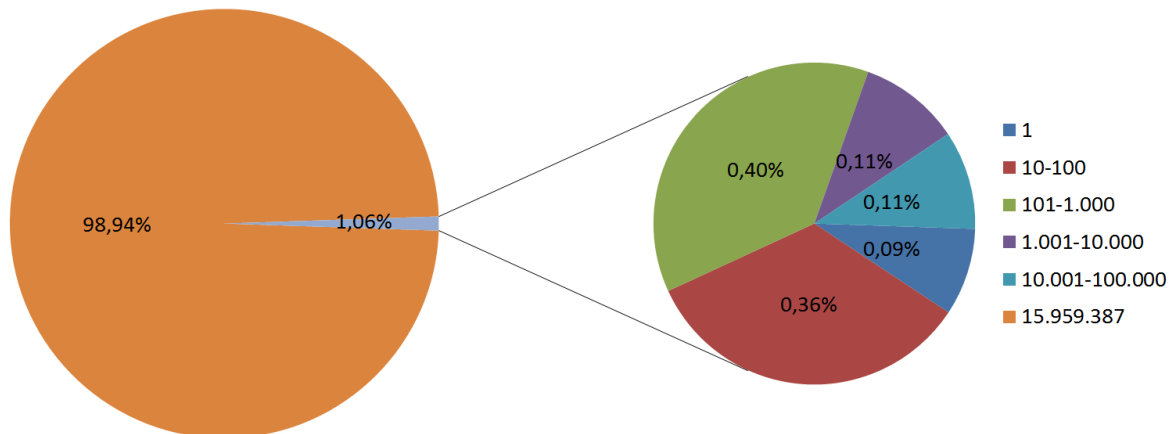


Abbildung 3.1.: Verteilung der Knoten in SZHKs

Um die Lage der Zusammenhangskomponenten in der Größe zwischen 101 und 10.000 Knoten zu überprüfen, wurden die enthaltenen Knoten mit einem erstellten Programm auf der Deutschlandkarte visualisiert. Dadurch soll sichergestellt werden, dass die SZHKs in dem erwarteten Bereich liegen. Exemplarisch werden hier fünf Zusammenhangskomponenten aus dem Bereich 1.001 bis 10.000 Knoten dargestellt; abgesehen von der SZHK in Pforzheim liegen alle im erwarteten Bereich.

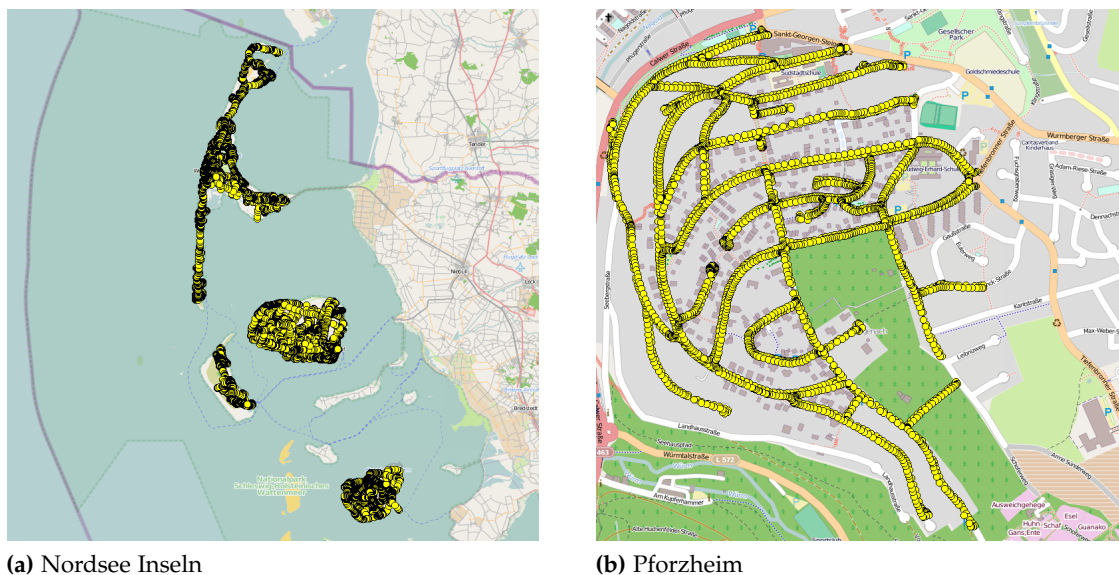


Abbildung 3.2.: SZHKs zwischen 1.000 und 10.000 Knoten

Abbildung 3.2 a) Die Inseln um Sylt enthalten 4 von 8 starken Zusammenhangskomponenten aus dem Bereich von 1.000 bis 10.000 Knoten, weitere 3 SZHK aus diesem Bereich befinden

sich auf weiteren Inseln und sind ebenfalls nicht durch Straßen zu erreichen. Abbildung 3.2 b) zeigt die einzige größere SZHK, die sich auf dem Festland befindet und die Anschluss an das Straßennetz haben müsste. Diese Zusammenhangskomponente ist anscheinend durch das doppelte Einpflegen von Straßen entstanden, da sich unter dieser SZHK noch Knoten und Kanten befinden, die Bestandteil der größten SZHK sind und die gleichen Straßen zu repräsentieren scheinen.

3.4. Ergebnis

Da die Verteilung der starken Zusammenhangskomponenten die geographischen Rahmenbedingungen widerspiegelt und der Graph zu 99% zusammenhängend ist, ist sichergestellt, dass der Graph für eine vertiefte Untersuchung geeignet ist. Das Kapitel 4 befasst sich weitergehend mit der Kartenqualität.

Durch die Verteilung der starken Zusammenhangskomponenten in Deutschland kann ebenfalls sichergestellt werden, dass die exportierten Wegtypen ausreichen, um einen Zusammenhang fast aller Knoten sicherzustellen. Durch den Export von Feldwegen und andere kleiner Straßen könnte vermutlich die Zahl der SZHK weiter verringert werden.

4. Geometrische Spannereigenschaften

4.1. Erläuterung des Konzepts

Unter einem geometrischen Spanner-Graph oder "k-spanner-graph" versteht man einen Graphen $G(V, E)$, für den gilt: $k * d_g(v, w) \geq d_w(v, w)$ für alle $v, w \in V$. Die Metrik $d_g(v, w)$ ist definiert als die geometrische Distanz zwischen dem Knoten v und w . Die Wegdistanz $d_w(v, w)$ ist definiert als der kürzeste Weg in G zwischen den Knoten v und w . Die Verzerrung k ist also der Quotient von Weglänge und Distanz.

In der vorliegenden Arbeit wird das dem Spanner-Graphen zugrundeliegende Konzept der "Verzerrung" (Verhältnis von Wegstrecke zu geometrische Distanz) zur Analyse von Anbindungsproblemen in Straßen-Graphen verallgemeinert.

Verbunden mit dem Konzept der Verzerrung ist ein intuitives Verständnis der Güte eines Weges: Ein "guter Weg" hat eine geringe Verzerrung.

Ziel des Kapitels ist es, Knoten zu identifizieren, deren Anbindung an das Straßennetz schlecht ist. Diese Knoten können viele andere Knoten nur über substanzielle Umwege erreichen, sie sind dementsprechend Startpunkt überproportional vieler "schlechter" Wege. Es wird erwartet, dass diese Knoten Hinweise auf fehlerhafte Kartendaten liefern oder ungünstigen aber realen Wegeführungen entsprechen.

4.2. Adaption

Um die Idee der Verzerrung jedoch verwenden zu können, muss die Funktion zur Verzerrungsberechnung der Spanner-Graphen an die sphärische Distanzberechnung angepasst werden.

Als Metrik wird anstatt der geometrischen Distanz die Entfernung auf der Kugeloberfläche verwendet.

$$d_g(v, w) = D_E * \cos^{-1}(\sin(v_{lat}) * \sin(w_{lat}) + \cos(v_{lat}) * \cos(w_{lat}) * \cos(w_{lon} - v_{lon}))$$

Mit dem Erdradius $D_E = 6.378.137m$.

Weiter wird erwartet, dass sich der erwartete Verzerrungsfaktor abhängig von der Distanz von Start und Zielpunkt ändert. Ein Verzerrungsfaktor von 1,5 mag bei einer Entfernung von 500 Metern noch im erwarteten Bereich liegen, bei einer Entfernung von 500 Kilometern wird die erwartete Verzerrung jedoch bedeutend geringer sein.

Dieses Problem lässt sich in 2 Unterprobleme zerlegen:

1. Bei sehr kurzen Distanzen (unter 300 Meter) führt ein hoher Quotient nicht automatisch zu einem problematischen Weg
2. Die Erwartung an den Verzerrungsfaktor k für "gute Wege" ändert sich bei zunehmender Entfernung

Das erste Problem wird durch eine modifizierte Definition des Verzerrungsfaktors k gelöst.

$$k = \max\left(\frac{d_w - 100}{\max(d_g, 1)}, 1\right)$$

Durch die Subtraktion von 100 Metern¹ von der Weglänge wird bei kurzen Distanzen eine große Menge von "Irrelevant Positives", also Wegen mit großer Verzerrung aber kleiner Weglänge, eliminiert. Da die Verringerung der Weglänge klein und konstant ist, wird sie bei größeren Entfernungen kaum noch ins Gewicht fallen.

Das zweite Problem wird durch die Unterteilung der Wege in verschiedene Bereiche gelöst. Durch diese Bereiche ist es möglich, entfernungsabhängige Grenzen für schlechte Wege zu definieren. In der vorliegenden Arbeit wurden folgende Grenzen verwendet (in Kilometer) : $[0..2^0],]2^0..2^1],]2^1..2^2], \dots$

4.3. Naiver Berechnungsansatz

Die einfachste Lösung wäre, die Distanz aller möglichen Wege zu bestimmen und die ermittelten Daten zur Analyse zu speichern. Dies führt jedoch dazu, dass die Laufzeit wie auch der Platzbedarf der Ergebnisse jeden Rahmen sprengt.

Im ungünstigsten Fall benötigt man zum Speichern der Daten pro Weg etwa 100 Byte. Bei 16 Millionen Knoten hat man $16.000.000^2$ mögliche Wege und benötigt ungefähr 230 Terrabyte zum Speichern aller Informationen.

Der Zeitbedarf zum Berechnen aller Informationen wäre ebenfalls kritisch. Die Berechnung der Weglänge von einem Knoten zu allen restlichen Knoten des Graphen dauert ungefähr 7 Sekunden inklusive der Berechnungen der Distanzen auf der Kugeloberfläche.

Auf dem Server der Algorithmik mit 24 Kernen würde man bei 20 Threads mehr als 2 Monate benötigen, um alle Berechnungen durchzuführen.

4.4. Verbesserungsideen

Da wir uns hauptsächlich für Knoten mit vielen Wegen mit hoher Verzerrung interessieren, ist es nicht notwendig, alle Wege zu speichern. Zu diesem Zweck reicht eine quantitative Analyse der Wege eines Knoten.

Durch das Zählen von guten und schlechten Wegen kann ein Verhältnis bestimmt werden,

¹Die Relevanzgrenze von 100 Metern wurde willkürlich gewählt, sie entspricht in etwa der Distanz, unter der selbst ein typischer IT Student ein Auto nicht mehr verwenden würde

welches Aufschluss über die Anbindung des Startknoten gibt. Die Klassifizierung von guten und schlechten Wegen wird im Abschnitt 4.5.3 vorgenommen.

Durch die Reduktion auf die rein quantitativen Daten wurde eine Reduktion des Platzbedarfes um 7 Größenordnungen auf ungefähr 500 Megabyte erreicht.

Eine weitere Reduktion von Laufzeit wie auch vom Platzbedarf erreichen wir durch die Einschränkung der Knotenmenge für Start- und Zielknoten. In der Realität sind Punkte auf der Autobahn oder Landstraße keine Startpunkte oder Ziele für eine Fahrt, sondern man möchte beispielsweise von seinem Wohnort in ein Einkaufszentrum fahren. Dementsprechend sind Punkte auf Autobahnen, Landstraßen und ähnlichen Orten als Start und Zielpunkt nicht interessant und werden nicht untersucht.

Eine weitere Möglichkeit zur Laufzeitreduktion bietet das Clustern von Knoten, siehe Abschnitt 4.6.3.

4.5. Schätzung ausgewählter Kennzahlen des Graphen

Um sich einen Überblick über die zu erwartenden Quotienten und der Wegverteilung zu machen, wurde eine Untersuchung anhand einer zufälligen Stichprobe von 25.000 Knoten durchgeführt.

4.5.1. Mittelwert und Nachbarknoten

Das Verhältnis der Summe aller Weglängen und der Summe aller Distanzen (Mittelwert) gibt Hinweise, welche Verzerrung in den Bereichen zu erwarten ist. Die Mittelwerte bestätigen die Erwartung, dass die Verzerrung abnimmt, je weiter entfernt das Ziel ist. Um jedoch eine Grenze für gute oder schlechte Wege abzuleiten, reicht die Betrachtung der Mittelwerte noch nicht. Die durchschnittliche Anzahl von Knoten im Umkreis ermöglichen eine Abschätzung, wie sich bei Änderung des untersuchten Entfernungsbereichs die Laufzeit verändert.

Verzerrungsmittelwerte und Zielknoten im Bereich:

Minimale Distanz	Maximale Distanz	Mittelwert	Knoten im Bereich
0	1.000	1,64	840
1.001	2.000	1,55	727
2.001	4.000	1,48	2.109
4.001	8.000	1,40	6.639
8.001	16.000	1,33	22.200
16.001	32.000	1,28	76.158
32.001	64.000	1,24	259.935
64.001	128.000	1,21	828.666
128.001	256.000	1,20	2.395.567
256.001	512.000	1,18	4.282.968
512.001	1.024.000	1,17	856.166

4.5.2. Verteilung der Wege

Um eine bessere Vorstellung zu erhalten, wie sich die Wege auf die verschiedenen Verzerrungsfaktoren verteilen, wurde mit einer Diskretisierung des Verzerrungsfaktors eine Zählung der entsprechenden Wege durchgeführt. Den Ausschnitt zwischen dem Faktor 1 und 5 der gewonnenen Daten findet man im Anhang im Abschnitt A.3.

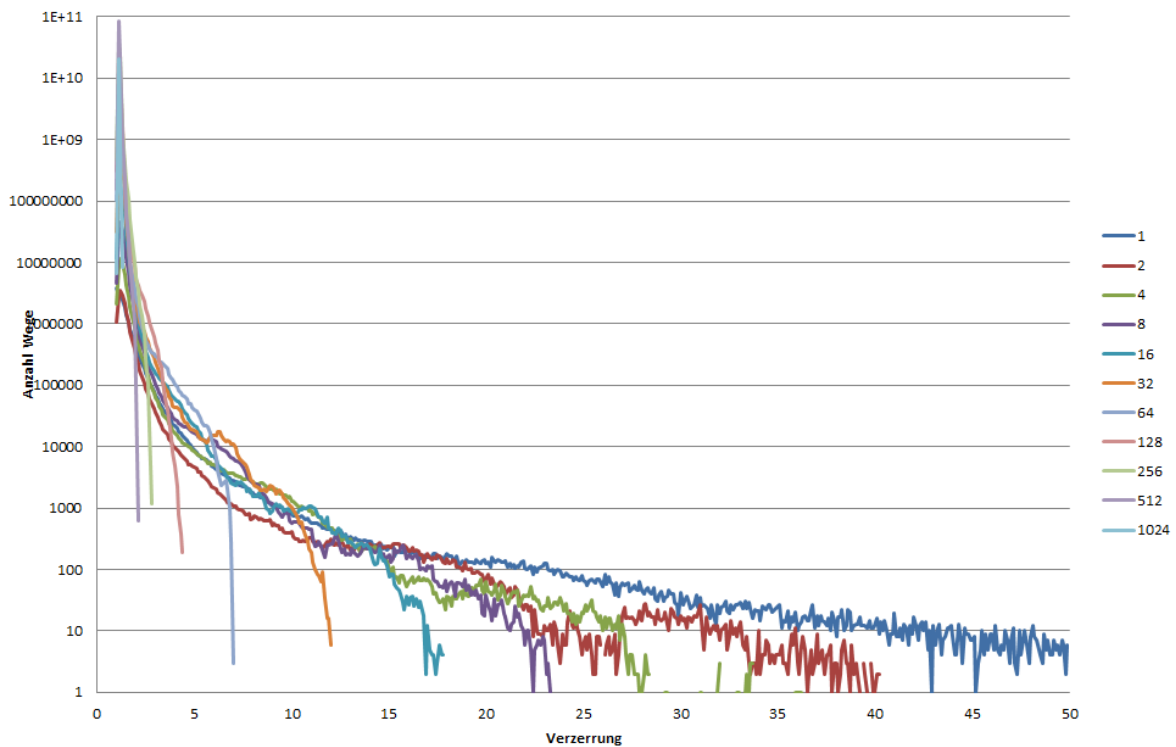


Abbildung 4.1.: Verteilung der Wege nach Verzerrung

Deutlich ist erkennbar, dass gerade in kurzen Distanzen die Verzerrung recht hohe Werte annehmen kann. Zu beachten ist die logarithmische Skalierung, die Anzahl der Wege mit großer Verzerrung ist also gering.

4.5.3. Klassifikation der Wege

Die Daten der Stichprobe ermöglichen eine genauere Festlegung, welche Verzerrung wir in den einzelnen Bereichen als "schlecht" betrachten wollen. Im weiteren betrachten wir Wege als schlecht, die zu den 10% der Wege mit größter Verzerrung aus ihrem Bereich gehören. Je nach Anwendungszweck lässt sich diese Grenze verschieben, möchte man beispielsweise nur die Wege mit größter Verzerrung darstellen. Durch eine kumulative Darstellung der in 4.5.2 verwendeten Daten erhält man eine klare Vorstellung, ab welcher Verzerrung hier ein Weg als schlecht betrachtet wird.

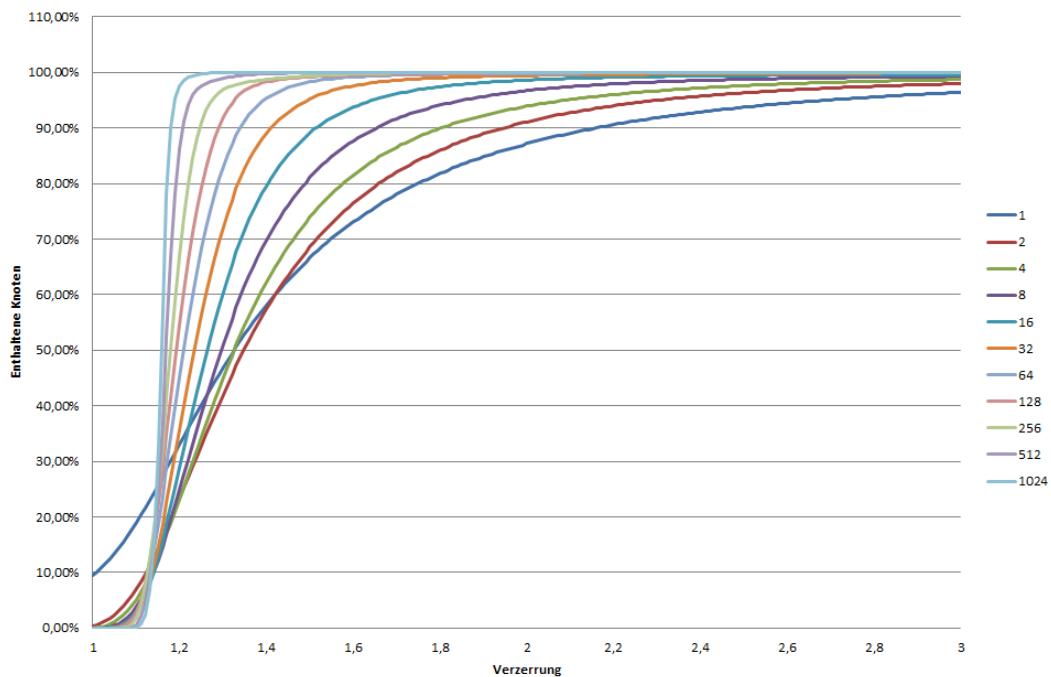


Abbildung 4.2.: Kumulative Verteilung der Wege

4.5. Schätzung ausgewählter Kennzahlen des Graphen

Die genauen Werte der Verzerrungsgrenze der 10 beziehungsweise 1 Prozent der Wege mit höchstem Verzerrungsfaktor findet man in der folgenden Tabelle:

Minimale Distanz	Maximale Distanz	10%	1%
0	1.000	2,16	4,89
1.001	2.000	1,94	3,9
2.001	4.000	1,81	3,29
4.001	8.000	1,66	2,7
8.001	16.000	1,52	2,16
16.001	32.000	1,42	1,8
32.001	64.000	1,34	1,58
64.001	128.000	1,29	1,47
128.001	256.000	1,25	1,44
256.001	512.000	1,21	1,31
512.001	1.024.000	1,19	1,22

4.5.4. Verteilung der Knoten

Zur Kontrolle der Grenzen wurde bei einer erneuten Stichprobe mit 25.000 Knoten die Verteilung von guten und schlechten Wegen ermittelt. Diese Analyse dient der Sicherstellung der Aussagekraft der Grenzen als Maß der Anbindung eines Knoten.

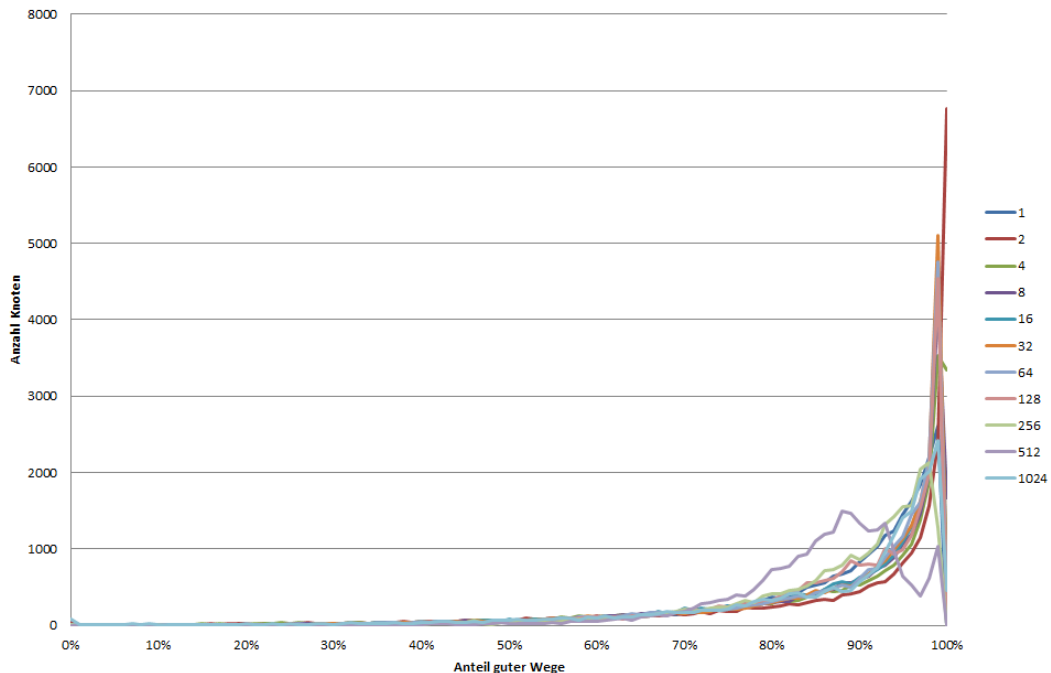


Abbildung 4.3.: Zählung der Knoten nach ihrer prozentualen Verteilung von guten und schlechten Wegen

Interessant ist auf den ersten Blick das Verhalten des Bereichs zwischen 256 und 512 km. Ein ähnliches Verhalten gibt es, wenn auch nicht so ausgeprägt, auch für den Bereich zwischen 64 und 128 km. Da das Verhalten jedoch in einem hier nicht interessanten Bereich liegt, wird auf eine weitere Untersuchung verzichtet.

Wichtig zur Analyse ist hauptsächlich der Bereich zwischen 0% und 60%, da sich hier die schlecht angebotenen Knoten befinden. Besonderheiten einzelner Bereiche, Häufungen oder stark unterschiedliche Verläufe in den Bereichen ließe auf eine schlecht gewählte Schranke schließen.

Bei den gewählten Schranken bietet sich jedoch in allen Bereichen das Bild einer sich der X-Achse asymptotisch annähernden Funktion.

4.6. Ermittlung der Anbindung

4.6.1. Algorithmus

Zum Finden der kürzesten Wege wurde Dijkstras Algorithmus verwendet, der mit einem Fibonacci-Heaps eine Laufzeit von $O(n * \log(n) + m)$ besitzt. In der vorliegenden Arbeit wurde ein Binary-Heap verwendet, jedoch wurde auf die Operation "decrease-key" verzichtet, indem Knoten mehrfach eingefügt werden konnten und die Duplikate ignoriert wurden. Dies ist besser, weil die geringe Menge an mehrfachen Knoten im Heap im Vergleich zu den höheren Konstanten des Fibonacci-Heaps bei allen Operationen zu vernachlässigen ist. Die maximale Größe des Heaps von 25.000 Elementen wurde zu keinem Zeitpunkt überschritten.

Zum Berechnen der Güte der Knoten wurde für jeden Bereich und jeden Knoten ein eigener Durchlauf gestartet. In den Gruppen unter 32 Kilometern wurde abgebrochen, sobald ein Knoten aus dem Heap entfernt wurde, dessen Distanz größer war als die maximale Distanz multipliziert mit der Schranke des Bereichs. Die Summe der guten Wege ist definiert durch die Anzahl der Knoten, die aus dem Heap entfernt wurden und deren k aus der Formel in 4.2 kleiner oder gleich der Grenze ist. Die Summe der schlechten Wege ist definiert als die Anzahl der Knoten, deren Verzerrung größer ist als k oder die sich im Bereich befinden, jedoch noch nicht aus dem Heap gezogen wurden. Zur genauen Ermittlung der Menge der Knoten, die sich in dem entsprechenden Bereich um den Startpunkt befinden, wurde ein k -d Baum verwendet.

Leider ist auf Grund der Anzahl der Knoten in Bereichen über 32 Kilometer das Verwenden eines k -d Baums nicht mehr sinnvoll möglich, da das Suchen einer großen Menge an Knoten zu zeitintensiv ist. Hier wird abgebrochen, sobald die Distanz eines aus dem Heap gezogenen Knoten das Zehnfache der oberen Grenze des Bereichs überschritten hatte. Die erwartete Menge der zu diesem Zeitpunkt nicht besuchten Knoten liegt im Bereich von 32 bis 64 Kilometern unter 0,2%, bei größeren Bereichen sind praktisch alle Knoten des Bereichs aus dem Heap entfernt worden.

Die Auswahl der Startknoten erfolgt durch die zufällige Auswahl von Knoten, in dem am Anfang alle Knoten einer Liste hinzugefügt und danach zufällig angeordnet werden. Danach wird immer das erste Element aus der Liste entfernt.

4.6.2. Knotenreduktion

Bei der Untersuchung des Graphen interessierten uns lediglich Punkte als Start- und Zielknoten, die auch für Menschen üblicherweise als Start- oder Endpunkt in Frage kommen. Die Verzerrung von einer Landstraße zu einem Punkt auf der Autobahn ist uninteressant, da diese Punkte keine üblichen Start- oder Zielpunkte sind.

Als übliche Startknoten werden alle Knoten betrachtet, die sich an einem Straßentyp befinden,

wie sie beispielsweise in Städten, Wohngebieten oder auch Industriegebieten vorkommen. Folgende Straßentypen liegen mindestens teilweise auch in diesen Gebieten:

highway	Beschreibung
tertiary	Kreisstraße oder innerstädtische Vorfahrtsstraße
tertiary_link	Verbindungsstücke zu Kreisstraßen
residential	Straße in Wohngebieten
living_street	Verkehrsberuhigter Bereich
road	Straße unbekannter Klassifikation
turning_circle	Kreisverkehr

Durch diese Einschränkung der Start- und Zielknoten konnte die Anzahl an gültigen Knoten von über 16 Millionen auf 8,7 Millionen reduziert werden. Dies reduziert sowohl die Rechenzeit wie auch den Platzbedarf um weite knappe 50%.

4.6.3. Clustering

Kleine Bereiche lassen sich noch ohne Probleme für jeden Startknoten berechnen, da die zu behandelnde Knotenmenge sehr gering ist. Bei der Stichprobe in 4.5.1 wurde auch die erwartete Anzahl der Knoten in jedem Bereich ermittelt. Multipliziert man jetzt noch den Faktor oder gar 10 auf die Schranke des Bereichs, wird man feststellen, dass die zu behandelnde Anzahl an Knoten bei großen Bereichen stark steigt. Um dennoch Rechenzeiten von weniger als einem Tag pro Bereich zu erhalten, repräsentiert jeder Startknoten in den Bereichen mit einem Radius von mehr als 8 km eine Menge von Knoten, deren Wegdistanz vom und zum Startknoten im Verhältnis zum Radius klein ist.

Die Radius $MaxDist$ lässt sich durch die Schranke q des Bereichs, die akzeptable Schranke q_{new} der Clusterknoten und der äußeren Grenze d des Bereichs berechnen.

$$MaxDist = \frac{d * (q_{new} - q)}{q_{new} + 1}$$

Alle Knoten, die innerhalb von $MaxDist$ den Startknoten erreichen und von ihm erreicht werden können, werden durch die Güte des Startknoten repräsentiert.

Ist der Startknoten gut an das Straßennetz angebunden und kann der Startknoten mit kurzer Distanz erreicht werden, sind seine nahen Knoten mindestens durch den Startknoten gut an das Straßennetz angebunden. Ist der Startknoten schlecht an das Straßennetz angebunden und erreicht er mit kurzer Distanz andere Knoten, müssen diese ebenfalls schlecht angebunden sein, da sonst der Startknoten über diese Knoten gut an das Straßennetz angebunden wäre.

Durch die Forderung, dass der Startknoten die durch ihn repräsentierten Knoten schnell erreicht, wie auch von ihnen schnell erreicht wird, können wir auf die Fallunterscheidung von Knoten mit hoher und niedriger Güte verzichten.

In der vorliegenden Arbeit wurde bei Clusterknoten eine maximale Abweichung der Verzerrung von 0,05 akzeptiert. Setzt man den maximalen Radius eines Bereichs und die Schranke

ein, erhält man die maximalen Clusterentfernungen aus der folgenden Tabelle:

Clustergrößen und maximale Abweichung der Verzerrung

Min Distanz	Max Distanz	Grenze	Abweichung	Max Clusterdistanz
0	1.000	2,16	0	0
1.001	2.000	1,94	0	0
2.001	4.000	1,81	0	0
4.001	8.000	1,66	0	0
8.001	16.000	1,52	0,05	155
16.001	32.000	1,42	0,05	323
32.001	64.000	1,34	0,05	669
64.001	128.000	1,29	0,05	1367
128.001	256.000	1,25	0,05	2782
256.001	512.000	1,21	0,05	5663
512.001	1.024.000	1,19	0,05	11428

4.6.4. Laufzeit

Wie erhofft, ist es gelungen, die Laufzeit in einem akzeptablen Bereich zu halten. Durch die Clustering ist es möglich, die Verzerrung in Bereichen mit großem Radius in einem akzeptablen Bereich zu halten.

Minimale Distanz	Maximale Distanz	Dauer	Anzahl Knoten	Dauer(s) pro Knoten
0	1.000	11656	8731618	0
1.001	2.000	12193	8731618	0
2.001	4.000	15831	8731618	0
4.001	8.000	30202	8731618	0
8.001	16.000	21350	1722426	155
16.001	32.000	43475	853497	323
32.001	64.000	78649	349358	669
64.001	128.000	64593	128990	1367
128.001	256.000	26710	48015	2782
256.001	512.000	8874	15574	5663
512.001	1.024.000	2462	4091	11428

Dass die meiste Rechenzeit im Bereich von 32km bis 64km benötigt wird, lässt sich durch die Größe des Bereichs erklären, der berechnet werden muss. Die Berechnung wird abgebrochen, sobald ein Knoten aus dem Heap entfernt wird, dessen Distanz vom Startknoten das Zehnfache der maximalen Distanz des Bereichs überschreitet. Innerhalb von 640 km befindet sich aber bereits der Großteil aller Knoten der Karte. In den späteren Bereichen ist die Reduktion der Knoten bereits stärker als der Zuwachs an Knoten.

4.7. Ergebnis

Zur Visualisierung der Ergebnisse wurde ein Java Programm geschrieben, das die Knoten entsprechend ihrem Verhältnis von Wegen mit guter und schlechter Verzerrung anzeigt. Wie erhofft, erhält man so Hinweise auf Gebiete in der Karte, die problematisch sind.

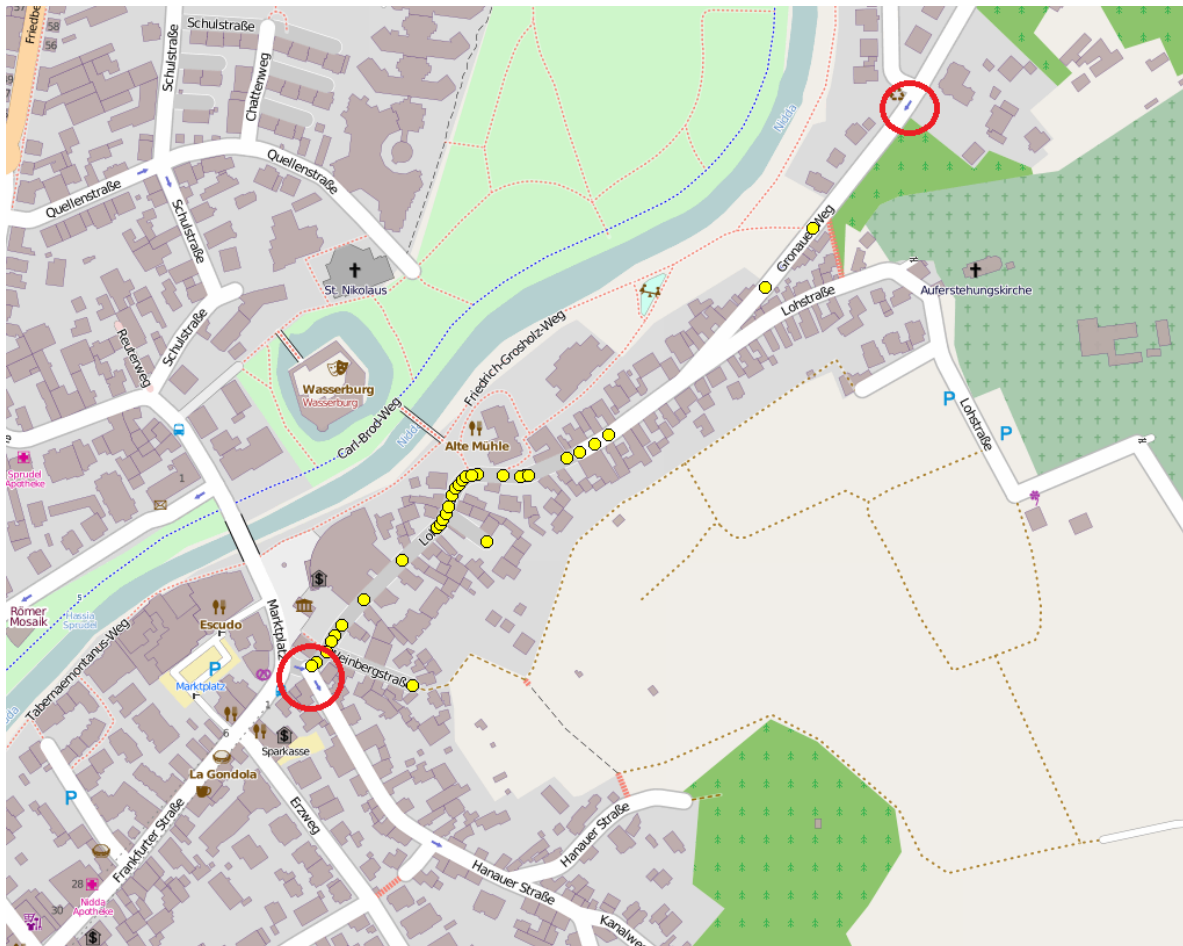


Abbildung 4.4.: Ungünstige Einbahnstraßen

In Abbildung 4.4 wird zum Beispiel durch eine Einbahnstraße ein großer Umweg für die Bewohner der Lohstraße (gelbe Punkte) erzwungen, wenn sie in den Bereich der anderen Mainseite gelangen wollen. Durch die Entfernung der Einbahnstraße an einem kleinen Straßenabschnitt könnte die Anbindung des Gebietes an den nördlichen Bereich bedeutend verbessert werden.



Abbildung 4.5.: Schlechte Anbindung durch geographische Barrieren

Da der Graph keine Information zu Fähren enthält, sind diese Knoten ebenfalls schlecht angebunden. Die Distanz zu den Knoten im Süden der Karte ist klein, jedoch ist die Fahrtstrecke über die nächst gelegene Brücke bedeutend länger.

Auf Grund der hohen Anzahl von problematischen Gebieten sind die Ergebnisse am besten im erstellten Programm zu betrachten, da hier die geographische Einordnung besser erfolgen kann.

Wie in den obigen Kartenausschnitten erkennbar, ist das Verfahren geeignet, um in einer Karte schlecht angebundene Knoten zu finden. Durch die Variierung der Grenzen kann auf verschiedene Probleme eingegangen werden. Leider ist die Unterscheidung zwischen Anbindungsproblemen, die aus geographischen Gegebenheiten entstanden sind, Anbindungsproblemen, die durch Straßenbau behoben werden könnten und Kartenfehlern hier nicht möglich. Durch das Verwenden von Informationen über Höhenverläufe, Flüsse oder anderer geographischer Barrieren könnte hier eventuell noch eine Verbesserung erreicht werden.

5. Zusammenfassung & Ausblick

Durch das Transformationsprogramm wurde eine Möglichkeit geschaffen, innerhalb von weniger als 3 Stunden einen komplett neuen und benutzerdefinierten Straßengraph aus den OpenStreetMap Rohdaten zu generieren. Durch die fakultative Bereinigung von starken Zusammenhangskomponenten ist dieser Graph einfach und ohne die Betrachtung von Sonderfällen für algorithmische Analysen verwendbar. Mit einer Automatisierung ist es möglich, einem größeren Publikum aktuelle Graphen zur Verfügung zu stellen ohne Zeit in die Wartung zu investieren.

Durch die Datenbank, die alle Informationen der OSM Rohdaten enthält, können auch spezielle Anfragen und Anforderungen unkompliziert beantwortet werden.

Die Analyse der starken Zusammenhangskomponenten hat, mit kleinen Abstrichen, die Erwartung erfüllt. 99% der Knoten befinden sich in der größten Zusammenhangskomponente. Es sind jedoch auch gröbere Fehler in den Rohdaten sichtbar geworden, die zur Verbesserung der Kartenqualität genutzt werden können.

Die Analyse der Verzerrungen hat Hinweise auf Knoten ergeben, die schlecht an das Straßennetz angebunden sind. Die Untersuchung einer Stichprobe dieser Stellen hat ergeben, dass die gefunden Stellen in der Tat aus verschiedensten Gründen schlecht an das Straßennetz angebunden sind. Durch das Programm zur Darstellung dieser Gebiete können die Problemzonen einfach untersucht werden.

Weitere Analysen könnten durch den Vergleich von Kartendaten aus verschiedenen Quellen genaue Hinweise auf Fehler in den Kartendaten geben, indem die Verzerrung der Wege von Punkten in unterschiedlichen Karten betrachtet wird. So könnte die Qualität von unterschiedlichem Kartenmaterial lokal und global verglichen werden.

Gebiete, die in allen Karten eine hohe Verzerrung besitzen, könnten durch die Betrachtung von Flüssen, Gebirgen und anderer Hindernisse nach straßenbaulichen Problemzonen und geographischen Problemzonen unterschieden werden.

A. Anhang

A.1. Datenbankdesign

A.1.1. Tabelle:nodes

```
1 DROP TABLE IF EXISTS 'nodes';
2 CREATE TABLE IF NOT EXISTS 'nodes' (
3   'osmID' bigint(20) NOT NULL,
4   'lat' double NOT NULL,
5   'lon' double NOT NULL,
6   KEY 'lat' ('lat','lon'),
7   KEY 'osmID' ('osmID')
8 ) ENGINE=MyISAM DEFAULT CHARSET=latin1;
```

A.1.2. Tabelle:node_tags

```
1 DROP TABLE IF EXISTS 'node_tags';
2 CREATE TABLE IF NOT EXISTS 'node_tags' (
3   'nodeID' bigint(20) NOT NULL,
4   'k' varchar(64) NOT NULL,
5   'v' varchar(256) NOT NULL,
6   KEY 'nodeID' ('nodeID'),
7   KEY 'k' ('k','v')
8 ) ENGINE=MyISAM DEFAULT CHARSET=latin1;
```

A.1.3. Tabelle:way_tags

```
1 DROP TABLE IF EXISTS 'way_tags';
2 CREATE TABLE IF NOT EXISTS 'way_tags' (
3   'wayID' bigint(20) NOT NULL,
4   'k' varchar(64) NOT NULL,
5   'v' varchar(256) NOT NULL,
6   KEY 'wayID' ('wayID'),
7   KEY 'k' ('k','v')
8 ) ENGINE=MyISAM DEFAULT CHARSET=latin1;
```

A.1.4. Tabelle:way_splits

```
1 DROP TABLE IF EXISTS 'way_splits';
2 CREATE TABLE IF NOT EXISTS 'way_splits' (
3   'wayID' bigint(20) NOT NULL,
4   'sourceID' bigint(20) NOT NULL,
5   'destinationID' bigint(20) NOT NULL,
6   KEY 'wayID' ('wayID')
7 ) ENGINE=MyISAM DEFAULT CHARSET=latin1;
```


A.2. Tabelle der Anzahl und Größe der starken Zusammenhangskomponenten

A.1.5. Tabelle:h_ways

```
1 DROP TABLE IF EXISTS 'h_ways';
2 CREATE TABLE IF NOT EXISTS 'h_ways' (
3   'sourceOSM' bigint(20) NOT NULL,
4   'destinationOSM' bigint(20) NOT NULL,
5   'slat' double NOT NULL,
6   'slon' double NOT NULL,
7   'dlat' double NOT NULL,
8   'dlon' double NOT NULL,
9   'type' varchar(32) NOT NULL,
10  PRIMARY KEY ('sourceOSM','destinationOSM','type'),
11  KEY 'type' ('type')
12 ) ENGINE=MyISAM DEFAULT CHARSET=latin1 DELAY_KEY_WRITE=1 ROW_FORMAT=FIXED;
```

A.2. Tabelle der Anzahl und Größe der starken Zusammenhangskomponenten

Größe der SZHK	Anzahl
1	14989
2	4968
3	2546
4	1966
5	1339
6	1061
7	765
8	659
9	512
10	405
11	343
12	291
13	244
14	199
15	183
16	156
17	141
18	95
19	108
20	85
21	83
22	70
23	81
24	43
25	52
26	51
27	39
28	38
29	33
30	35
31	24
32	24
33	24
34	23
35	20
36	17
37	14
38	12
39	16
40	14
41	20
42	11
43	13
44	21
45	12

A.2. Tabelle der Anzahl und Größe der starken Zusammenhangskomponenten

46	8
47	7
48	9
49	6
50	4
51	4
52	10
53	5
54	13
55	5
56	8
57	6
58	5
59	6
60	9
61	11
62	4
63	4
64	3
65	2
66	5
67	4
68	4
69	2
70	2
71	3
72	5
73	6
74	3
75	7
76	2
77	1
78	5
79	6
80	4
81	1
82	1
83	3
84	2
85	1
86	2
88	1
89	3
90	3
92	3
93	2
94	3
95	4
96	1
97	3
98	2
100	1
102	1
103	1
104	3
105	1
109	1
111	1
114	2
115	1
116	1
117	1
120	2
121	2
122	1
125	2
126	1
127	1
132	1
134	1
135	2

A.3. Verteilung der Wege aus der Stichprobe

137	1
139	1
141	1
145	1
148	1
150	1
153	1
154	1
167	1
170	1
183	1
187	2
189	1
193	3
195	1
197	1
198	1
199	1
202	1
205	1
219	1
224	1
226	1
229	1
230	1
231	1
246	1
247	1
266	1
273	1
283	1
287	1
289	1
296	1
328	1
337	1
338	2
348	1
353	1
377	1
392	1
403	1
418	1
522	1
535	1
594	1
674	1
683	1
1052	1
1452	1
1548	1
1608	1
2995	1
3102	1
5204	1
15959387	1

A.3. Verteilung der Wege aus der Stichprobe

	1	2	4	8	16	32	64	128	256	512	1024
1	1975700	71457	33580	7884	1803	214	1	0	0	0	0
1,01	147113	40868	38631	21827	11074	2763	227	0	0	0	0
1,02	157988	55928	65823	56481	45907	24788	7671	547	0	0	0
1,03	170784	72510	104107	120399	129013	118800	88429	101664	67	0	0
1,04	183931	90371	153502	220855	314236	409786	512478	1438893	156716	0	0
1,05	197716	109799	212424	372125	644920	1130635	1914685	5328914	1907385	0	0
1,06	210242	129266	277686	566028	1169872	2540758	5659060	14041926	8624460	455034	0

A.3. Verteilung der Wege aus der Stichprobe

1,07	223307	148582	346254	797392	1913968	4861289	13923342	35620938	32823142	6769032	7200
1,08	236271	169100	423471	1074604	2902121	8423341	28186482	83293173	112096623	36456999	671264
1,09	244179	188705	500554	1393375	4127732	13434857	50401451	170978272	286329000	124749334	8434574
1,1	258292	209025	581245	1725090	5599012	19797452	81034132	291187109	680742566	487368758	43154074
1,11	266332	227951	663540	2097434	7241459	27722409	119510545	434115506	1384863720	1615317952	122866801
1,12	276734	248882	744616	2477322	9040401	36638273	162899251	631433978	2149559772	2863406643	327638715
1,13	284335	266944	819461	2844337	10882352	46160159	208652749	838267846	2757378798	4637782929	962009693
1,14	291815	281794	884605	3192083	12655435	55666393	254333771	1009498970	3387051033	6970651684	1887446095
1,15	296765	297318	948744	3504927	14298726	64178655	293536455	1137261019	4086192092	980080456	3645658885
1,16	299956	309965	1008419	3793991	15770008	71170451	323391449	1236782821	4676371581	13491084551	5186803601
1,17	305163	318939	1061916	4014741	16988054	76567807	341009739	1316114107	5110204972	16115914142	4514689331
1,18	306827	329039	1101995	4208162	17892383	80278988	348552320	1368008192	5338833099	15529735047	2592352146
1,19	312014	340121	1133555	4354335	18610002	82744298	349504542	1377379021	5199502798	12306731882	1144628632
1,2	307350	341662	1152931	4445862	19007248	83465390	345549480	1335204909	4799297044	8498455709	484557223
1,21	308527	344876	1173039	4517673	19201199	83105016	334994485	1252407600	4175755709	5335069837	222047013
1,22	305410	346086	1187851	4556858	19147560	81746895	320166410	1143358119	3494581623	3181310568	113144569
1,23	303339	346594	1192647	4560640	18909819	79274600	303167437	1024288915	2838694899	1901820415	61722938
1,24	297262	345258	1181603	4515138	18511592	75760307	282944519	907085093	2208700658	1111232787	37270478
1,25	301042	349031	1183745	4474488	18043218	72285101	260152661	795829401	1662117165	688460827	26346014
1,26	292884	343669	1165313	4377807	17479249	68563743	236531116	695722698	1213907120	457580659	22260031
1,27	287721	339548	1148114	4289112	16840319	64211127	213047041	601332558	865807997	329919868	12660003
1,28	280999	334923	1127976	4180115	16177601	59766618	191041085	511084127	622126074	266170635	9098155
1,29	272939	329115	1104892	4042782	15434488	55487231	170728490	426714652	445299117	219899182	5796451
1,3	270525	322961	1081528	3908179	14656959	51091014	151754154	353233093	320243476	183970075	3262561
1,31	260975	314660	1049288	3752665	13874723	46856399	134227063	288970695	232720241	153543093	2208040
1,32	256780	308326	1017048	3607946	13091826	42916695	118201571	233890218	174682941	131062089	1295455
1,33	248782	299648	982215	3451526	12280154	39122911	103523311	189075956	135943609	111904630	559869
1,34	243384	290928	950443	3309406	11510168	35574967	90439769	151868902	110402654	95423566	129095
1,35	237301	283106	921221	3165485	10779719	32250069	78925489	121281368	92796826	81410169	23668
1,36	230333	276862	889741	3007613	10093426	29232946	68592836	96591777	80791896	71683407	6378
1,37	223485	267682	861019	2863153	9410026	26427137	59330575	77303093	72192581	65015622	915
1,38	217884	260176	826759	2724788	8760931	23920016	51497424	62080462	64972790	59238021	491
1,39	215705	251162	799075	2600396	8154477	21624581	44728987	50209683	59668546	50155668	27
1,4	204661	240583	767008	2468746	7575553	19546668	38579443	40753454	54643597	40543773	0
1,41	202491	236366	736938	2341304	7062034	17712214	33559537	33661655	50359531	34942372	0
1,42	197909	227205	707324	2218329	6599536	16096623	29186256	28290232	47377485	30099980	0
1,43	192693	218283	679123	2103618	6132794	14598369	25390673	24187942	44294490	28561116	0
1,44	187095	210979	651500	1991509	5727172	13209241	22148905	20551117	41222648	21863438	0
1,45	180998	203344	622397	1894867	5328358	11930040	19265027	17539128	37716817	18642792	0
1,46	177503	198253	596988	1794236	4952766	10766221	16844589	15102130	34774445	16134713	0
1,47	172805	190629	574248	1701164	4599071	9728245	14671065	13197859	32255157	14401573	0
1,48	167743	185180	547723	1613994	4293731	8834919	12797729	11734822	30245282	12626587	0
1,49	159070	176412	523522	1522403	3989728	7996439	11252522	10661340	28687506	11256323	0
1,5	161719	173436	506143	1452409	3730030	7268459	9913092	9761789	27191062	9722070	0
1,51	153431	167008	483801	1371451	3449514	6576441	8753645	8999571	25719903	8604344	0
1,52	150193	162566	463775	1301314	3213516	5933290	7732225	8236667	25139523	7457257	0
1,53	145583	156680	444111	1237102	2999573	5371160	6864133	7471905	24394492	6770745	0
1,54	141053	150269	425222	1172116	2799670	4900435	6077177	6942719	23427581	6073551	0
1,55	137524	145611	408350	1107528	2608814	4446966	5458749	6453406	22132421	5421936	0
1,56	133486	141796	390818	1049049	2428586	4063483	4906111	6047105	20351905	4821036	0
1,57	129370	136326	373489	994585	2263026	3688409	4409456	5730251	18774377	4543809	0
1,58	126261	131091	359971	947159	2120123	3351172	3983160	5491659	17309427	4293597	0
1,59	121453	127176	341846	898783	1970761	3059708	3563978	5262240	15865391	3992977	0
1,6	121248	124162	330092	856798	1842493	2802982	3232551	5069408	14555862	3698282	0
1,61	116717	118047	315438	813639	1717118	2574529	29252131	4786049	13380921	3318517	0
1,62	113558	113804	302156	774424	1608972	2372657	2714682	4455160	12363084	3002196	0
1,63	111527	109626	289542	735731	1508837	2181356	2479824	4182942	11458592	2698775	0
1,64	107889	105878	278440	699688	1413516	2012800	2278026	3955048	10715165	2440197	0
1,65	104819	101731	267264	666940	1323844	1859402	2084310	3722219	9998752	2215628	0
1,66	101909	98980	258390	634201	1241678	1707059	1918581	3494465	9284333	2017089	0
1,67	99027	95165	246996	601532	1168039	1563439	1759545	3295281	8696747	1852799	0
1,68	96623	91828	236427	573715	1097002	1435937	1626280	3098495	8047488	1680568	0
1,69	93870	89052	227648	544976	1031054	1345273	1523164	2927288	7439898	1514364	0
1,7	92273	86384	219375	521708	968028	1251217	1422765	2780623	6905166	1374291	0
1,71	89341	83922	211394	498327	914843	1151651	1319403	2671892	6543375	1219157	0
1,72	86934	79845	203314	475208	860712	1066977	1228077	2579567	5983628	1094464	0
1,73	84114	77275	195727	453548	808873	997846	1147706	2437509	5289658	1013289	0
1,74	81047	75255	187709	433535	752720	936978	1096997	2336466	4747252	940461	0
1,75	81136	73637	181747	412847	713807	873191	1031271	2231450	4408898	875032	0
1,76	78829	71089	176332	390764	670032	821221	967282	2133509	4097142	818757	0
1,77	77281	68182	169504	376487	632603	773289	910340	2061323	3739902	766885	0

A.3. Verteilung der Wege aus der Stichprobe

1,78	75067	66566	162003	360079	599283	716940	865069	1954083	3419869	680542	0
1,79	71891	64232	155726	342285	562334	673291	809067	1875185	3125775	607263	0
1,8	71786	62559	150462	328430	538452	634891	764008	1782379	2853273	524800	0
1,81	69551	60454	145192	316131	507999	604235	732273	1718419	2597265	458753	0
1,82	68221	58591	139333	300509	480159	563563	691658	1647354	2353927	404337	0
1,83	66451	56971	133526	288409	459205	527426	663936	1578725	2174527	363757	0
1,84	64444	54923	129026	274782	433203	495783	630613	1517956	2015174	321899	0
1,85	63793	53855	124211	263231	409556	468474	600656	1431399	1875009	286872	0
1,86	62162	51925	120251	251357	388258	446225	573838	1336339	1751677	244321	0
1,87	60585	50788	115639	241982	366810	430383	552282	1278957	1627445	229082	0
1,88	58759	49199	110774	232446	348220	411264	523443	1216371	1505373	200551	0
1,89	57634	47723	107705	221652	331848	388800	504953	1169243	1410388	182194	0
1,9	55894	46133	104281	212834	315670	368464	487985	1138218	1279780	155053	0
1,91	55076	44736	100389	204972	301571	348180	474184	1076754	1198193	141994	0
1,92	54181	43984	96820	197742	287597	332400	456980	1018766	1156397	127226	0
1,93	52252	42674	93955	188842	273049	321417	439529	955607	1060303	107392	0
1,94	51033	41203	90860	179989	261920	300973	429551	900898	969751	94591	0
1,95	50354	39702	87913	173324	249950	287632	411647	844001	928749	90932	0
1,96	48410	38708	84373	166941	239271	274862	392588	791709	893169	75280	0
1,97	47578	37819	80935	161312	225967	268161	376202	751424	835335	63208	0
1,98	46609	36636	78582	155491	217199	258832	357352	708869	797510	52047	0
1,99	43061	34838	75682	148924	208110	247461	337802	667375	782823	39397	0
2	47393	35152	74949	144404	200304	235678	331766	625823	751118	31369	0
2,01	43388	33695	71678	137952	193108	226047	317873	614894	708366	23312	0
2,02	43153	32612	69653	133068	186434	220072	311480	578730	656344	17523	0
2,03	41831	32046	66607	127180	177856	212911	303957	550955	631480	13639	0
2,04	41393	30978	65627	123153	169647	206685	294021	530403	609974	7202	0
2,05	39990	30075	63417	117613	161632	201273	292176	509151	586342	4778	0
2,06	39651	29213	62002	115161	156432	192407	280923	491206	545681	3137	0
2,07	38354	28370	59878	110493	151241	185876	270209	472014	517806	1327	0
2,08	37793	27568	58115	106947	144459	180380	264293	463491	478371	605	0
2,09	36611	26922	55781	102206	139160	172674	255734	456980	448702	865	0
2,1	36418	26203	54572	98822	133505	169317	250872	442764	420868	239	0
2,11	35385	25576	53331	94917	128938	161238	247549	430442	408605	48	0
2,12	34615	25213	51576	91607	124727	155819	236997	416346	397054	14	0
2,13	33761	24271	49920	86962	119306	154349	231013	404154	379403	0	0
2,14	33232	23592	48642	84913	113921	148614	223275	394165	359814	0	0
2,15	32559	23136	47270	81761	109208	144041	214624	392562	350636	0	0
2,16	31944	22151	45489	79805	106368	140888	208366	386402	329624	0	0
2,17	30889	21778	44014	76602	103191	137306	196620	374267	314632	0	0
2,18	30450	21140	43251	74296	98742	131811	187748	372490	285670	0	0
2,19	29414	20687	41769	71896	95014	128297	182170	361537	267860	0	0
2,2	29551	20131	40895	69985	92527	126165	172330	351074	245898	0	0
2,21	28519	19728	39925	66808	87694	124364	171133	344196	234284	0	0
2,22	28092	19097	38251	64880	85787	119908	166253	335199	234613	0	0
2,23	27483	18706	36892	64197	82878	118176	161447	335864	216306	0	0
2,24	26690	18038	35670	61950	80586	115470	157446	327937	208918	0	0
2,25	26757	17829	35770	59918	77597	115050	149657	316320	201715	0	0
2,26	26011	17399	34898	58457	73857	112820	148259	314195	204527	0	0
2,27	25757	17121	33924	56560	73632	110475	143924	313058	192884	0	0
2,28	24974	16398	33006	53973	71160	108606	140391	307615	192985	0	0
2,29	24657	16316	32060	52671	70149	108324	136241	294762	185255	0	0
2,3	23991	15639	31446	50686	68207	104933	130491	289134	176440	0	0
2,31	23856	15106	30714	49342	65830	103347	123449	290682	163465	0	0
2,32	23026	14828	29835	47714	63157	101302	118587	291739	154267	0	0
2,33	22326	14513	29190	46553	61490	99392	110307	281003	139499	0	0
2,34	22101	14233	28983	45393	59728	97312	105814	270334	128746	0	0
2,35	21712	13896	27888	43650	58365	95048	103452	264524	124521	0	0
2,36	21564	13345	27274	42615	57083	92315	101621	256423	125073	0	0
2,37	20948	13173	26823	41487	55069	88883	99335	251825	120633	0	0
2,38	20510	12965	26160	39657	51753	87426	97552	249105	107774	0	0
2,39	19769	12503	25555	38869	50361	86982	94147	246740	103216	0	0
2,4	20066	12563	24454	37703	48202	86942	94445	244930	95798	0	0
2,41	19688	12261	23935	36912	47938	85839	91316	240229	84191	0	0
2,42	19068	11689	23488	36232	47300	85266	89521	232183	74245	0	0
2,43	18652	11446	23201	35141	45766	83790	88049	226939	62856	0	0
2,44	18656	11296	22409	34704	44195	81656	86432	225967	55327	0	0
2,45	18421	10555	22223	33511	43732	81652	83166	214879	52047	0	0
2,46	17649	10236	21626	32712	42634	81244	80720	213691	47860	0	0
2,47	17318	10273	20992	31660	41831	77122	77245	211503	44741	0	0
2,48	16973	10152	20507	30578	40968	76111	77238	209695	42030	0	0

A.3. Verteilung der Wege aus der Stichprobe

2,49	16023	9683	19989	30334	39227	74963	75557	203369	36783	0	0
2,5	17047	9382	19594	29719	38873	73766	74322	195844	33324	0	0
2,51	16322	9334	19346	28468	37867	70345	70898	191433	33415	0	0
2,52	15978	9387	18600	28032	37244	67945	71335	177127	31025	0	0
2,53	15904	9027	18413	27133	36544	67767	70820	172146	27891	0	0
2,54	15565	8815	17871	26332	35404	66489	69300	165928	23276	0	0
2,55	15210	8603	17600	25652	34279	64435	67927	161230	21022	0	0
2,56	14926	8653	17090	25259	33352	62912	65973	153433	19684	0	0
2,57	14693	8354	16928	24498	32325	61529	64174	149762	17560	0	0
2,58	14668	7933	16390	23990	31706	61207	60663	141302	15794	0	0
2,59	13810	7822	15961	24010	30867	60316	58344	136232	15378	0	0
2,6	14100	7738	15560	23695	31080	60179	55587	133672	12684	0	0
2,61	13582	7559	14886	22730	29900	59748	54081	128905	9927	0	0
2,62	13432	7400	14666	22493	29753	58977	52902	126214	8553	0	0
2,63	13437	7385	14318	21725	28349	60322	52233	121131	6805	0	0
2,64	13090	7274	14291	21119	28219	59188	52651	117515	6726	0	0
2,65	12901	6954	13872	21158	27574	57815	51779	118827	7193	0	0
2,66	12565	6849	13575	20541	26603	58009	51112	116539	6305	0	0
2,67	12375	6641	13386	20430	26653	56650	48555	116096	4790	0	0
2,68	11789	6600	13187	20053	25865	56537	48502	111944	5817	0	0
2,69	12022	6469	12768	19581	25246	55575	45917	109842	4571	0	0
2,7	12166	6294	12496	19458	25071	54880	46438	105776	3725	0	0
2,71	11580	6078	12193	18852	24324	52927	45918	103102	2353	0	0
2,72	11670	6008	11786	18565	23464	51194	47299	103752	2024	0	0
2,73	11166	5936	11527	17942	22930	50125	48459	105197	1766	0	0
2,74	10835	5852	11224	16911	21997	48626	45971	99720	1528	0	0
2,75	11185	5681	11266	17052	22382	48058	46311	97239	1702	0	0
2,76	10791	5582	11119	16889	22084	46461	45771	94128	2211	0	0
2,77	10369	5673	10728	16535	22134	44602	44688	92756	1342	0	0
2,78	10381	5383	10233	16294	21937	43926	42491	90585	969	0	0
2,79	10539	5348	10038	16411	21875	44332	41653	88099	832	0	0
2,8	10019	5174	9775	15772	20988	43767	42045	86082	477	0	0
2,81	9874	4991	9575	15805	20530	42713	39832	83453	160	0	0
2,82	9741	4875	9388	15082	20504	42280	39227	78391	147	0	0
2,83	9662	4967	9183	14789	20108	42407	37523	77270	203	0	0
2,84	9545	4713	9204	14847	19713	41854	35817	70859	133	0	0
2,85	9370	4684	9090	14787	19551	41251	35374	68251	179	0	0
2,86	9304	4715	9020	14354	19043	39559	35473	68476	163	0	0
2,87	8901	4653	8567	13973	19333	38417	37444	70489	37	0	0
2,88	8875	4636	8610	13689	18920	37842	36126	69760	0	0	0
2,89	8805	4415	8285	13658	18538	37079	35248	65824	0	0	0
2,9	8589	4548	8406	13199	19168	36821	35976	63036	0	0	0
2,91	8397	4389	8000	12937	19429	36452	34933	61188	0	0	0
2,92	8319	4236	7903	12680	18715	35542	33019	61274	0	0	0
2,93	8360	4264	7629	12636	18174	35484	33260	60287	0	0	0
2,94	8163	4245	7474	12378	18348	34371	32639	55874	0	0	0
2,95	8147	4171	7305	12053	17937	33173	30691	53608	0	0	0
2,96	7920	4076	7316	11827	18273	32946	30644	52649	0	0	0
2,97	7714	3972	7187	11564	17592	32683	30553	52546	0	0	0
2,98	7689	3880	7033	11281	17480	32325	30421	50241	0	0	0
2,99	7081	3833	6926	10790	17183	32089	29907	48625	0	0	0
3	8042	3868	6889	11167	16817	31755	29249	45016	0	0	0
3,01	7364	3775	6689	11128	16881	30863	28170	42737	0	0	0
3,02	7157	3699	6689	10808	16561	31100	29057	42202	0	0	0
3,03	7042	3700	6390	10540	16553	29990	29051	42375	0	0	0
3,04	7198	3680	6416	10143	16238	28558	28322	42803	0	0	0
3,05	6698	3539	6510	10088	16267	28877	29192	44379	0	0	0
3,06	6679	3419	6351	10266	15651	28490	28890	43735	0	0	0
3,07	6769	3420	6076	10073	15631	28387	27687	42220	0	0	0
3,08	6639	3384	5991	9684	15391	28214	26666	37876	0	0	0
3,09	6527	3236	5911	9758	15090	27530	26750	36215	0	0	0
3,1	6588	3160	5771	9599	15270	26500	26520	36213	0	0	0
3,11	6400	3139	5776	9526	15424	26427	26233	36269	0	0	0
3,12	6269	3050	5539	9418	15841	25847	26153	36071	0	0	0
3,13	6284	3089	5566	9219	15289	25217	26968	36236	0	0	0
3,14	6193	2841	5421	8675	15400	25643	27561	34635	0	0	0
3,15	6025	2799	5385	8583	15525	25360	26881	34237	0	0	0
3,16	6029	2823	5406	8746	15064	24863	26397	32637	0	0	0
3,17	5932	2782	5237	8456	15197	24126	26461	30590	0	0	0
3,18	5796	2862	5179	8378	15193	24093	27075	30507	0	0	0
3,19	5551	2872	5078	8168	14944	22980	26498	28812	0	0	0

A.3. Verteilung der Wege aus der Stichprobe

3,2	5629	2713	5004	8354	15010	22727	26271	27245	0	0	0
3,21	5605	2630	4983	7850	15366	22714	26490	25021	0	0	0
3,22	5477	2590	5020	7814	15185	21392	26632	23580	0	0	0
3,23	5483	2587	4938	7891	14742	22052	26717	23689	0	0	0
3,24	5275	2504	4722	7764	14749	21245	25800	22561	0	0	0
3,25	5427	2497	4698	7605	14507	21173	25486	22025	0	0	0
3,26	5222	2456	4597	7520	14429	21004	24334	20358	0	0	0
3,27	5245	2379	4423	7359	14148	20581	24736	19702	0	0	0
3,28	5157	2381	4551	7258	14171	20316	24492	19367	0	0	0
3,29	5107	2320	4511	7224	14414	19875	24999	18925	0	0	0
3,3	4962	2395	4291	7137	14178	19680	23906	18440	0	0	0
3,31	4937	2258	4368	6917	13828	19141	23986	17024	0	0	0
3,32	4864	2322	4220	6853	13521	18989	24065	17449	0	0	0
3,33	4760	2259	4197	6757	13645	18998	23663	17190	0	0	0
3,34	4899	2249	4036	6639	13678	18798	23070	16516	0	0	0
3,35	4670	2300	4094	6651	13887	18762	23510	16644	0	0	0
3,36	4590	2188	3964	6373	13506	18416	22820	15566	0	0	0
3,37	4642	2128	3842	6436	13184	18255	22830	15777	0	0	0
3,38	4588	2084	3771	6243	13470	17901	23677	15673	0	0	0
3,39	4269	1993	3761	6288	13162	17622	23600	14913	0	0	0
3,4	4474	2111	3739	6262	13094	17263	22632	13536	0	0	0
3,41	4204	2053	3736	6140	13127	16853	23065	12969	0	0	0
3,42	4194	2058	3729	6010	12675	17391	23226	11411	0	0	0
3,43	4227	2100	3459	6030	12690	16830	23568	9873	0	0	0
3,44	4147	2019	3439	5931	12572	15840	24151	9257	0	0	0
3,45	4157	1945	3531	5893	12355	15576	23223	8935	0	0	0
3,46	4025	1886	3541	5602	12534	14967	22192	8350	0	0	0
3,47	3956	1833	3350	5585	12281	14844	21744	8293	0	0	0
3,48	3993	1840	3373	5514	12105	14309	21437	8012	0	0	0
3,49	3715	1846	3251	5569	12235	14430	20917	7465	0	0	0
3,5	4008	1830	3287	5337	12489	14262	20473	6706	0	0	0
3,51	3915	1827	3092	5315	12286	13724	20981	6082	0	0	0
3,52	3811	1815	3033	5290	12141	13301	20376	6249	0	0	0
3,53	3654	1817	3060	5215	11787	13271	20205	5751	0	0	0
3,54	3701	1835	3081	5181	11780	12831	21479	5554	0	0	0
3,55	3514	1894	3094	4923	11616	12909	20165	4912	0	0	0
3,56	3551	1759	3005	4977	11454	12389	19773	4436	0	0	0
3,57	3758	1874	2994	4825	11317	11765	20078	4199	0	0	0
3,58	3486	1770	2940	4639	11071	11821	20450	4071	0	0	0
3,59	3454	1688	2914	4543	11083	11673	20128	3939	0	0	0
3,6	3466	1663	2683	4563	10694	11307	20065	3894	0	0	0
3,61	3384	1589	2730	4385	11207	11045	19810	4025	0	0	0
3,62	3350	1586	2618	4350	10882	10798	19806	3865	0	0	0
3,63	3336	1546	2609	4463	10808	10825	19717	3582	0	0	0
3,64	3332	1492	2699	4326	10532	10484	20365	3145	0	0	0
3,65	3220	1528	2668	4233	10821	10582	19828	3043	0	0	0
3,66	3274	1497	2606	4137	10300	10371	19262	2945	0	0	0
3,67	3203	1464	2458	4173	10337	10113	18473	2921	0	0	0
3,68	3061	1394	2473	4137	10327	9427	18204	2720	0	0	0
3,69	3108	1481	2470	4060	9955	9480	17316	2902	0	0	0
3,7	3068	1434	2490	4259	10026	9453	16528	2951	0	0	0
3,71	2998	1388	2389	4029	10221	9718	16127	2646	0	0	0
3,72	2981	1415	2280	4205	9700	9416	15769	2531	0	0	0
3,73	2923	1396	2443	4083	9849	9733	15633	2284	0	0	0
3,74	2859	1376	2351	4196	9595	8991	14950	2107	0	0	0
3,75	2990	1403	2463	4122	9634	9019	14583	2088	0	0	0
3,76	2910	1296	2289	4083	9569	8770	14705	2163	0	0	0
3,77	2817	1328	2468	4158	9552	8666	14930	2069	0	0	0
3,78	2744	1292	2360	3954	9275	8531	15138	1869	0	0	0
3,79	2704	1341	2356	3936	9255	8113	14644	1604	0	0	0
3,8	2956	1253	2284	3921	9034	8085	14446	1395	0	0	0
3,81	2748	1305	2260	3917	8772	7860	13764	1287	0	0	0
3,82	2748	1204	2398	3871	8989	7956	12868	1011	0	0	0
3,83	2598	1143	2218	3776	8581	7824	13057	812	0	0	0
3,84	2649	1104	2298	3780	8577	8166	12470	784	0	0	0
3,85	2638	1214	2289	3679	8574	7946	12690	812	0	0	0
3,86	2516	1132	2179	3659	8476	7607	13202	757	0	0	0
3,87	2549	1240	2123	3688	8228	7707	12304	620	0	0	0
3,88	2506	1138	2171	3717	8277	7301	11843	573	0	0	0
3,89	2521	1103	2087	3608	8290	6831	12303	618	0	0	0
3,9	2586	1150	2040	3567	8319	6837	12338	593	0	0	0

A.3. Verteilung der Wege aus der Stichprobe

3,91	2447	1081	2104	3463	8335	6718	11708	465	0	0	0
3,92	2475	1121	1998	3466	8210	6467	11501	471	0	0	0
3,93	2406	1117	1966	3159	8054	6333	11550	417	0	0	0
3,94	2284	1102	1909	3138	8235	6474	11817	447	0	0	0
3,95	2308	1078	1998	3253	7976	6538	11555	417	0	0	0
3,96	2381	1040	1917	3298	7817	6089	11028	385	0	0	0
3,97	2341	1022	1850	3223	7849	6231	10915	309	0	0	0
3,98	2175	1049	1783	2977	7790	6380	10758	249	0	0	0
3,99	1950	1020	1837	3018	7687	6358	10525	407	0	0	0
4	2509	1174	1797	3058	7889	6101	10524	350	0	0	0
4,01	2218	1042	1867	2981	7634	6009	9993	408	0	0	0
4,02	2252	1048	1807	2995	7354	5959	9344	433	0	0	0
4,03	2120	1012	1705	2911	7261	6152	9605	474	0	0	0
4,04	2161	1008	1734	2842	7395	6121	9345	558	0	0	0
4,05	2180	951	1656	2895	7484	5886	9511	412	0	0	0
4,06	2101	969	1668	2722	7333	5948	9339	223	0	0	0
4,07	2152	958	1616	2763	6985	5803	9109	203	0	0	0
4,08	2080	900	1624	2650	6947	5822	9180	189	0	0	0
4,09	2120	919	1626	2655	6906	5851	8794	308	0	0	0
4,1	2023	914	1752	2665	6788	5950	8622	373	0	0	0
4,11	1939	922	1675	2744	6869	5869	8563	506	0	0	0
4,12	2078	879	1665	2663	6718	5418	8625	366	0	0	0
4,13	1904	931	1631	2571	6525	5382	8924	333	0	0	0
4,14	1882	850	1609	2563	6528	5218	8651	223	0	0	0
4,15	1967	821	1612	2515	6450	5334	8441	345	0	0	0
4,16	1898	851	1676	2541	6246	5403	8425	186	0	0	0
4,17	1880	862	1617	2484	6126	5171	8617	221	0	0	0
4,18	1879	833	1590	2545	6264	5121	8605	266	0	0	0
4,19	1797	841	1543	2485	6131	4940	8485	125	0	0	0
4,2	1896	850	1535	2468	6296	5094	8253	24	0	0	0
4,21	1797	822	1565	2288	6123	4631	7999	34	0	0	0
4,22	1769	835	1574	2280	6268	4624	8379	28	0	0	0
4,23	1792	814	1476	2297	6130	4684	8385	27	0	0	0
4,24	1656	812	1472	2242	6250	4499	8201	17	0	0	0
4,25	1782	823	1512	2411	6204	4485	8250	15	0	0	0
4,26	1757	754	1438	2376	6020	4611	8067	14	0	0	0
4,27	1638	795	1423	2458	5816	4343	7979	18	0	0	0
4,28	1790	759	1449	2335	5780	4367	8339	33	0	0	0
4,29	1639	764	1302	2411	5748	4350	8351	19	0	0	0
4,3	1707	747	1295	2352	5877	4367	7971	18	0	0	0
4,31	1681	765	1378	2268	5691	4411	7783	22	0	0	0
4,32	1626	798	1337	2261	5525	4256	7666	28	0	0	0
4,33	1536	727	1376	2200	5670	4342	7583	43	0	0	0
4,34	1595	761	1263	2281	5535	4402	7169	52	0	0	0
4,35	1591	731	1309	2209	5374	4568	7407	18	0	0	0
4,36	1614	723	1250	2249	5193	4295	7455	23	0	0	0
4,37	1548	695	1258	2322	5362	4293	7211	26	0	0	0
4,38	1548	704	1299	2200	5253	4152	7127	9	0	0	0
4,39	1504	723	1219	2305	5147	3870	6958	25	0	0	0
4,4	1557	702	1147	2274	5191	3697	7014	12	0	0	0
4,41	1491	703	1227	2235	5222	3725	7016	22	0	0	0
4,42	1539	701	1236	2185	5203	3692	7043	20	0	0	0
4,43	1447	682	1220	2243	5076	3709	6817	18	0	0	0
4,44	1502	681	1126	2223	5185	3696	6613	11	0	0	0
4,45	1471	696	1176	2319	5101	3666	6784	10	0	0	0
4,46	1376	710	1203	2227	5165	3610	6556	19	0	0	0
4,47	1442	684	1166	2279	5129	3522	6612	1	0	0	0
4,48	1421	710	1130	2291	4921	3521	6867	0	0	0	0
4,49	1381	685	1165	2277	5125	3527	6406	0	0	0	0
4,5	1568	662	1133	2090	5002	3320	6292	0	0	0	0
4,51	1389	639	1151	2080	4912	3335	6296	0	0	0	0
4,52	1380	620	1154	2224	4951	3299	5932	0	0	0	0
4,53	1405	620	1106	2147	4915	3248	5905	0	0	0	0
4,54	1343	600	1179	2226	4891	3187	5949	0	0	0	0
4,55	1298	643	1066	2191	4801	3168	6071	0	0	0	0
4,56	1310	612	1160	2183	4872	3025	5986	0	0	0	0
4,57	1309	604	1087	2218	4623	2977	6091	0	0	0	0
4,58	1321	557	1074	2086	4644	3134	6013	0	0	0	0
4,59	1368	543	1107	2143	4605	2948	6084	0	0	0	0
4,6	1264	581	1056	2058	4474	2863	5877	0	0	0	0
4,61	1140	520	1081	2215	4381	2794	5844	0	0	0	0

A.3. Verteilung der Wege aus der Stichprobe

4,62	1282	590	1073	2095	4510	2760	5920	0	0	0	0
4,63	1370	597	1039	2038	4321	2703	6062	0	0	0	0
4,64	1132	584	1076	2122	4157	2681	5946	0	0	0	0
4,65	1263	609	1006	2050	4254	2657	5899	0	0	0	0
4,66	1206	633	1063	2061	4038	2658	5887	0	0	0	0
4,67	1201	582	1051	1965	4010	2548	5727	0	0	0	0
4,68	1213	604	1014	2094	3862	2518	5831	0	0	0	0
4,69	1180	614	1056	2000	3954	2523	5801	0	0	0	0
4,7	1200	593	1041	1904	3843	2399	5237	0	0	0	0
4,71	1185	589	970	2001	3958	2390	5208	0	0	0	0
4,72	1142	574	1041	2070	3959	2297	4850	0	0	0	0
4,73	1151	542	981	2018	3940	2208	4689	0	0	0	0
4,74	1140	571	1047	2118	3826	2160	4473	0	0	0	0
4,75	1095	598	1032	2016	3792	2243	4501	0	0	0	0
4,76	1097	549	947	2059	3689	2201	4739	0	0	0	0
4,77	1134	551	955	1900	3892	2254	4797	0	0	0	0
4,78	1107	571	941	1889	3756	2127	4409	0	0	0	0
4,79	1058	549	870	1854	3757	2058	4364	0	0	0	0
4,8	1082	542	925	1915	3842	2008	4363	0	0	0	0
4,81	1033	528	855	1859	3866	2032	4334	0	0	0	0
4,82	1074	525	869	1932	3770	2082	4358	0	0	0	0
4,83	1008	547	899	1860	3813	1878	4338	0	0	0	0
4,84	1037	546	902	1857	3906	2012	4283	0	0	0	0
4,85	1053	525	804	1846	3829	1848	4289	0	0	0	0
4,86	1080	558	847	1881	3637	1820	4453	0	0	0	0
4,87	1085	507	864	1910	3596	1799	4487	0	0	0	0
4,88	1012	514	855	1863	3657	1611	4344	0	0	0	0
4,89	974	505	841	1801	3573	1628	4223	0	0	0	0
4,9	1029	510	866	1826	3658	1690	4533	0	0	0	0
4,91	1019	485	891	1830	3500	1605	4266	0	0	0	0
4,92	973	531	868	1743	3535	1783	4285	0	0	0	0
4,93	1003	511	886	1778	3351	1730	4143	0	0	0	0
4,94	893	470	870	1845	3284	1757	4141	0	0	0	0
4,95	1049	510	811	1822	3277	1753	3986	0	0	0	0
4,96	1005	498	868	1861	3267	1695	4077	0	0	0	0
4,97	968	496	854	1754	3127	1688	4447	0	0	0	0
4,98	987	448	821	1764	3197	1655	4220	0	0	0	0
4,99	817	456	770	1839	3067	1747	4025	0	0	0	0
5	1088	478	794	1842	3173	1724	4100	0	0	0	0

Literaturverzeichnis

- [DZ10] A. Z. Dennis Zielstra. A Comparative Study of Proprietary Geodata and Volunteered Geographic Information for Germany. Technischer Bericht, University of Heidelberg, Department of Geography, 2010. (Zitiert auf Seite 4)
- [NAS00] Shuttle Radar Topography Mission, 2000. URL <http://www2.jpl.nasa.gov/srtm/>. (Zitiert auf Seite 10)
- [OSM12a] OSM XML, 2012. URL http://wiki.openstreetmap.org/wiki/OSM_XML. (Zitiert auf Seite 6)
- [OSM12b] Planet.osm, 2012. URL <http://wiki.openstreetmap.org/wiki/Planet.osm>. (Zitiert auf Seite 6)
- [Tar72] R. E. Tarjan. Depth-First Search and Linear Graph Algorithms. *SIAM J. Comput.*, 1(2):146–160, 1972. (Zitiert auf den Seiten 11 und 14)

Alle URLs wurden zuletzt am 01.05.2012 geprüft.

Erklärung

Hiermit versichere ich, diese Arbeit selbständig verfasst und nur die angegebenen Quellen benutzt zu haben.

(Frederik Hartmann)