

Optimierung und Evaluierung einer Routenabgleichkomponente in einer Ad-Hoc-Mitfahrerzentrale



Bachelorarbeit, vorgelegt von Martin Gorges

Aufgabenstellung

Ziel dieser Arbeit ist das Testen und Überarbeiten eines bestehenden Mitfahrersuchalgorithmus, der vom Fraunhofer Institut für offene Kommunikationssysteme entwickelt wird.

Mit dem Algorithmus werden passende Mitfahrergesuche für einen Fahrer ermittelt, wobei die Lösung bei Kurzstrecken zum Einsatz kommen soll. Es gilt herauszufinden, unter welchen Bedingungen, wie viele Mitfahrer, in welcher Zeit, für den Fahrer gefunden werden. Weiterhin werden alternative Ansätze entwickelt und miteinander verglichen.

Größtes Problem ist der Vergleich der verschiedenen Ansätze, weil sie sich in unterschiedlichen Situationen als vorteilhaft bzw. nachteilig erweisen. Es könne deswegen Empfehlungen zu mehreren Ansätzen geben.

Testszenerien

Als Testumgebung dient die Stadtkarte von Berlin. Die Daten der Karte stammen von OpenStreetMap und werden in eine Datenbank geschrieben. Aus diesen Daten werden Testszenerien erstellt.

Für jeden Ansatz werden zwei Szenarien getestet. Das erste Szenario ist eine zwölf Kilometer lange Strecke, die ein gedachter Fahrer zurücklegen will. Um diese Strecke werden in einem bestimmten Bereich Mitfahrergesuche erzeugt. Bei einer zweiten identischen Strecke werden zusätzlich Mitfahrergesuche über ganz Berlin verteilt. Ziel dieses Szenarios ist es die Vollständigkeit zu ermitteln, indem gemessen wird, wie viele Gesuche gefunden wurden, sowie die Richtigkeit, durch Messen wie viele Gesuche bei der zweiten Strecke richtig gefunden wurden, also auch in der ersten Strecke vorkamen.

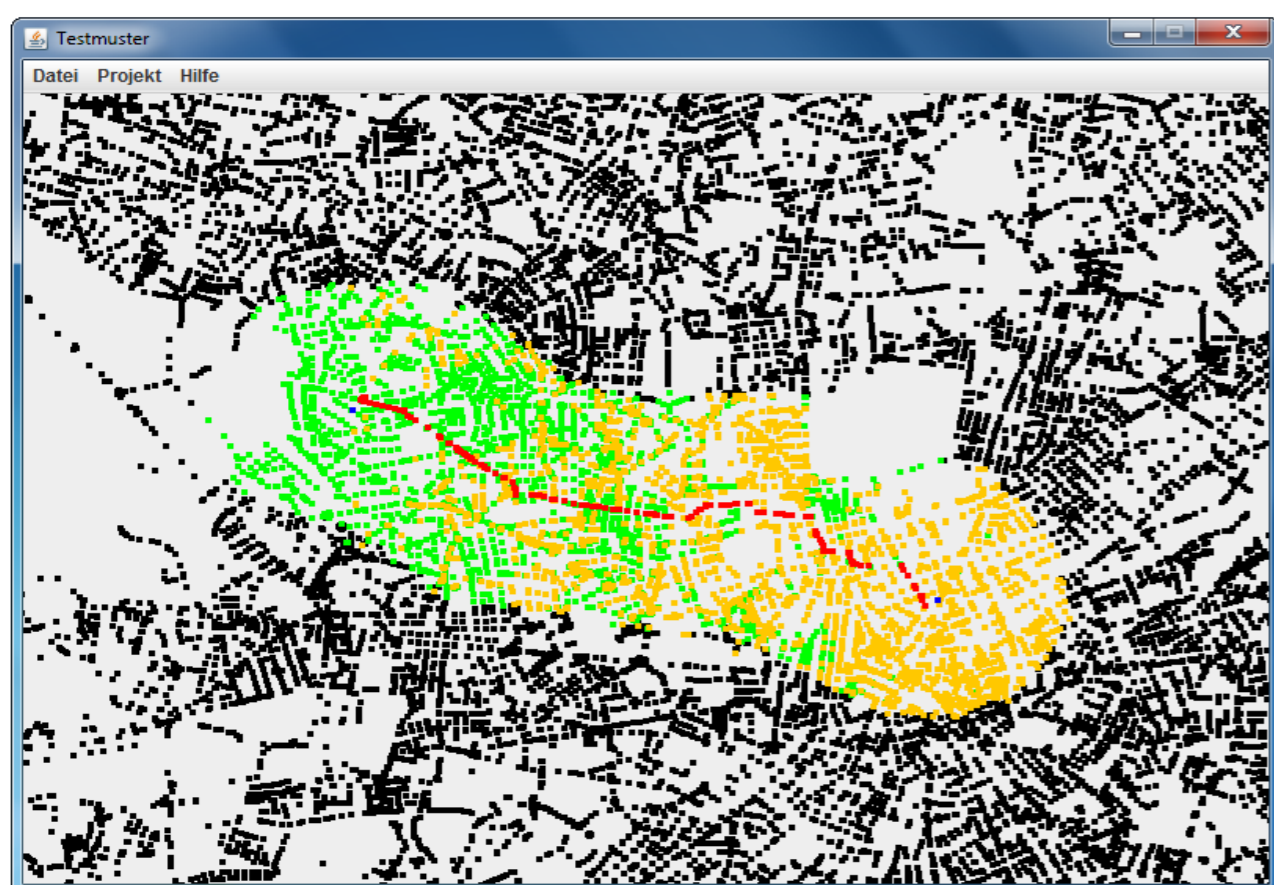


Abb.1 Strecke von Szenario Eins. (Rote und blaue Punkte bilden die Strecke, grüne und orange die Mitfahrergesuche)

Das zweite Szenario dient als Belastungstest. Es werden zwei Fahrerstrecken erzeugt mit jeweils größtmöglicher Anzahl an Gesuchen. Die erste Strecke ist sieben Kilometer lang und damit kürzer als jene aus Szenario Eins. Die zweite Strecke ist mit 24 Kilometern deutlich länger.

Das Szenario soll ein Bild erzeugen, in dem abzulesen ist, wie sich ein Ansatz bei verschiedenen Streckenlängen verhält. Nebenbei dient es als Vergleich zwischen den verschiedenen Ansätzen.

Veranschaulichung

Zur Erzeugung der Testszenarien und zur Nachvollziehbarkeit des Algorithmus wurde im Laufe dieser Arbeit ein Programm geschrieben. Das Programm lädt aus der Datenbank die Wege einer Stadt (hier Berlin) und stellt deren Eckpunkte grafisch dar. Durch Wahl einiger Punkte kann eine Fahrerstrecke erzeugt werden, für die automatisch Mitfahrergesuche gewählt werden. Die Strecke kann in verschiedenen Einstellungen abgespeichert werden, um dann für den zu testenden Algorithmus zur Verfügung zu stehen.

Der ursprüngliche Algorithmus ist in zwei Versionen im Programm eingebunden und kann für eine Teststrecke Schritt für Schritt durchgegangen werden.

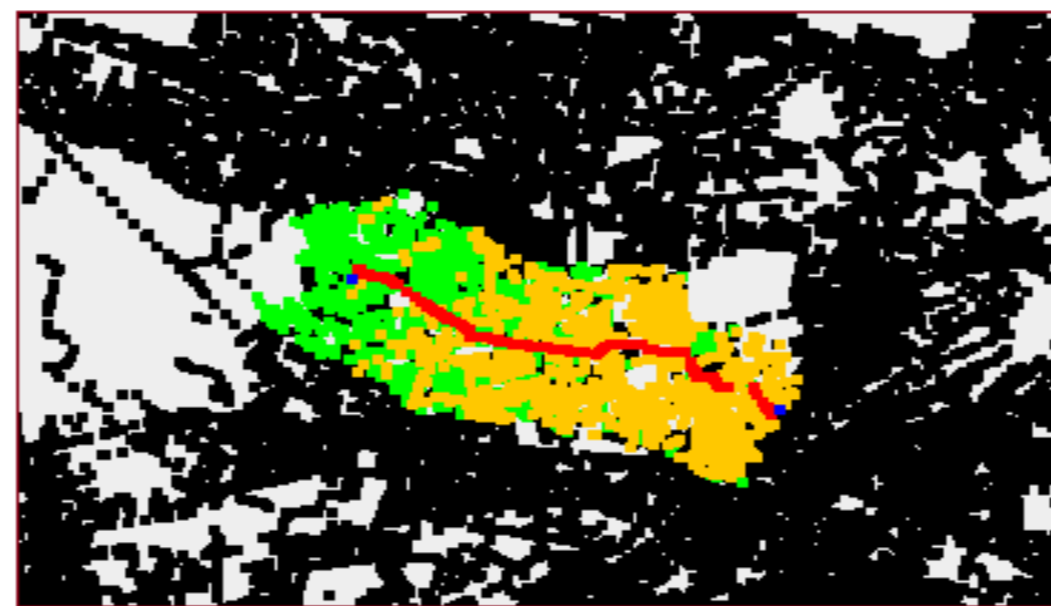
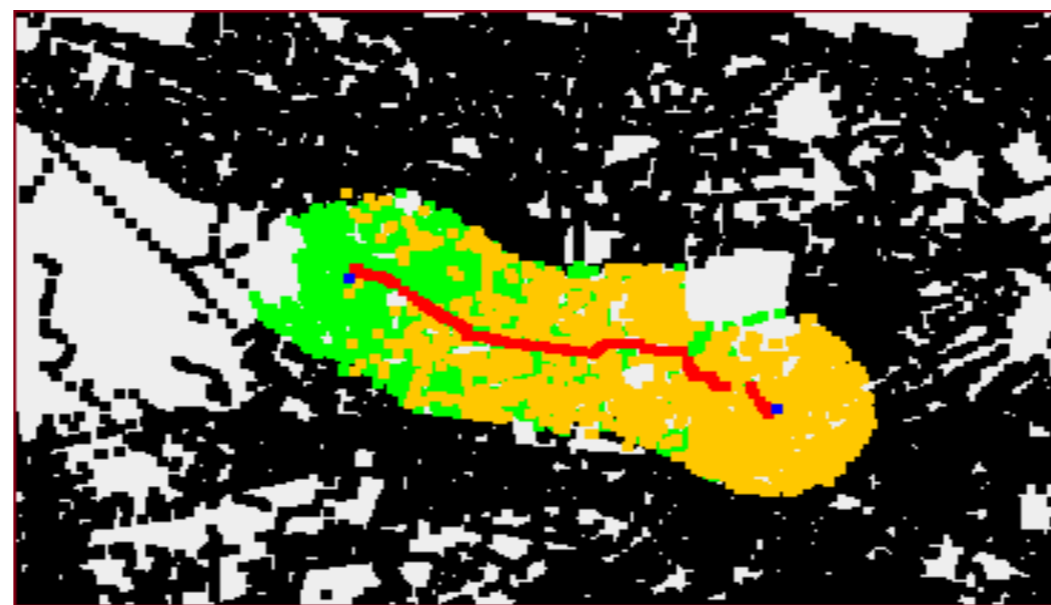


Abb.2 Beispiel eines Schritt für Schritt-Durchgangs

Ansätze

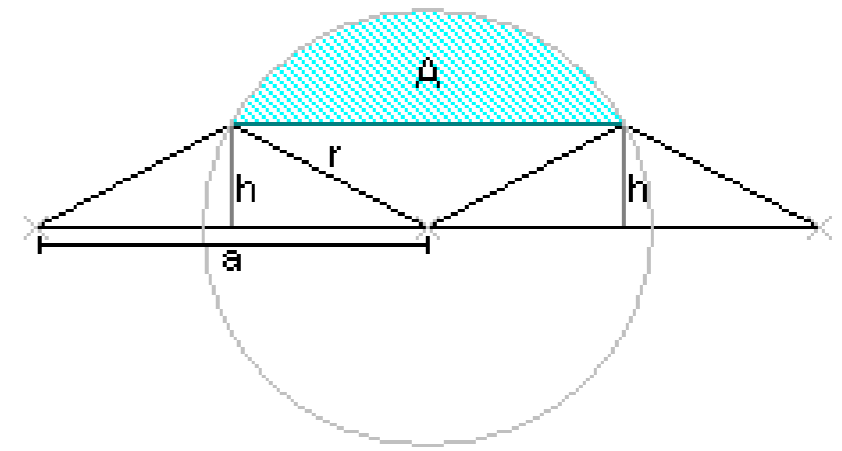
Wie zuvor erwähnt, wurden verschiedene Ansätze getestet.

Ein Ansatz muss möglichst schnell sein und eine große Anzahl an verfügbaren Mitfahrergesuchen finden, die zusätzlich noch richtig, also für den Fahrer passend sein sollen. In einem weiteren Schritt, der nicht in der Arbeit bearbeitet wurde, werden falsch gefundene Mitfahrergesuche erkannt und gestrichen. Falsche Gesuche sind jene, die einen bestimmten Umweg für den Fahrer überschreiten, den er zuvor festlegen kann.

Jeder Ansatz ermittelt mit Hilfe einer GIS-fähigen Datenbank die Mitfahrergesuche in einer bestimmten Umgebung in der Nähe der Strecke. Die Abkürzung GIS steht für Geoinformationssystem und ermöglicht einer Datenbank, Objekte im zwei- und dreidimensionalen Raum zu verwalten und dadurch Berechnungen wie in der Geometrie durchzuführen.

Bei den geometrischen Berechnungen gibt es einige Dinge zu beachten, die sich auf die benötigte Zeit und die gefundenen Mitfahrergesuche auswirken. Berechnungen mit einfacheren Flächen benötigen weniger Zeit als mit komplexeren. In größeren Flächen finden sich mehr Mitfahrergesuche als in kleineren. In größeren Flächen finden sich mehr falsche Gesuche (vgl. Abb.3, Formel für überflüssiges Kreissegment). In kleineren Flächen finden sich weniger richtige Gesuche. Flächen können sich überschneiden und somit Gesuche unnötigerweise zweimal herausgesucht werden. Bei längeren Strecken kann der Filtereffekt mitwachsender Flächen verloren gehen.

Sie filtern weniger falsche Gesuche heraus und die dafür benötigte Zeit ist höher als die Ersparnis, die man erhält, wenn die falschen Gesuche bei weiteren Abfragen nicht mehr betrachtet werden müssen.



$$A = \int_{-a/2}^{a/2} \sqrt{(r^2 - x^2)} dx - a * h$$

$$A = \left[x/2 * \sqrt{(r^2 - x^2)} + \frac{r^2}{2} \arcsin(x/r) \right]_{-a/2}^{a/2} - a * h$$

Abb.3 Beispiel von wachsender, überschüssiger Fläche (Hellblau) eines Kreises über die Strecke, bei größer werdendem Radius.

Weiterhin stellt sich die Aufgabe nach einer Abfrage, die nur jene Gesuche findet, die in die gleiche Richtung wollen, in die auch der Fahrer fährt. Unter diese Thematik fallen auch andere, grundsätzliche Probleme wie das Kreuzen der Mitfahrerstrecke mit der Fahrerstrecke oder die Frage, ob Mitfahrer am Streckenende abgesetzt oder gar nicht erst mitgenommen werden sollen (vgl. Abb.2, oranger Halbkreis hinter rechten blauen Punkt).

Auswertung

Zwei Ansätze konnten für die weitere Verwendung empfohlen werden. Zum Einen der ursprüngliche Ansatz mit den im Laufe der Arbeit ermittelten Parametern. Problem ist bei diesem die Richtigkeit, die nicht bei nahe 100% liegt. Ein zweiter Ansatz gewährt diese Richtigkeit, kann aber unter Umständen langsamer sein. Ein letztendlich hybrider Ansatz scheint eine gute Lösung oder auch eine bedingte Auswahl, entschieden je nach Streckenlänge oder anderen wichtigen Kriterien.

Die Wirklichkeit lässt sich schwer mit einigen Testszenerien nachstellen. Gegebenenfalls wurden maßlos viele oder auch wenige Mitfahrergesuche für die Tests verwendet. Auch andere Faktoren wie zum Beispiel Verkehrsstoßzeiten werden nicht berücksichtigt. Die Auswahl eines Ansatzes kann erst durch einen Praxistest gelöst werden.

Zusammenfassung

Bei der Bearbeitung des Algorithmus traten zuerst grundlegende Fragen auf. Die durch die Klärung gewonnenen Erkenntnisse konnten für den weiteren Arbeitsverlauf verwendet werden.

Der Hauptteil der Arbeit am Algorithmus widmete sich den Anfragen an die GIS-Datenbank. Das ursprüngliche Konzept erwies sich als vorteilhaft und wurde in verschiedenen Varianten getestet. Darauf aufbauend wurde ein zweiter Ansatz entwickelt, der in verschiedenen Varianten getestet wurde.

Weiterhin wurde eine Anfrage, die in allen Varianten auftritt, auf ihre Schwachstelle untersucht.

Am Ende konnte kein Ansatz vollständig überzeugen und es wurden lediglich Empfehlungen abgegeben.

Die weitere Verwendung des korrigierten Quelltextes und der Ansätze liegt beim Fraunhofer Institut.