

Entwicklung von GIS-Funktionen zur Analyse und Visualisierung GPS-basierter Mobilitätsdaten

Uta Samaga

Diplomarbeit am Institut für Kartographie,
TU Dresden

1. Dezember 2003

Entwicklung von GIS-Funktionen zur Analyse und Visualisierung GPS-basierter Mobilitätsdaten

Uta Samaga,

Diplomarbeit am Institut für Kartographie (IfK), TU Dresden,

in Zusammenarbeit mit dem Institut für Kartographie (IKA) und dem Institut für Verkehrsplanung und Transportsysteme (IVT), ETH Zürich

Telefon: +49-(0)351-8022553

eMail: uta.samaga@web.de

Dezember 2003

Kurzfassung

Für diese Arbeit lagen Mobilitätsdaten aus der schwedischen Studie „Rätt fart i Borlänge“ vor. Die Besonderheiten dieser Daten sind einerseits eine für Mobilitätsuntersuchungen ungewöhnlich lange Beobachtungsdauer der Verkehrsteilnehmer (bis zu zwei Jahren) mit exakten Zeit- und Ortsangaben aus GPS-Messungen und andererseits die durch die passive GPS-Beobachtung fehlenden Informationen zu ihren Wegzielen, die bei traditionellen Verkehrsverhaltensstudien immer erhoben werden und zur Erforschung des Verkehrsverhaltens nötig sind.

Diese Informationslücke musste geschlossen werden. Das Ziel dieser Arbeit war deshalb die Entwicklung eines Algorithmus, durch den die plausible Rekonstruktion der fehlenden Wegzwecke automatisiert abläuft.

Dazu wurden verschiedene Module in der ArcMacroLanguage (AML) programmiert. Das dabei entstandene Makropaket untersucht die Mobilitätsdaten auf grobe Fehler, sortiert Fahrten nach dem Wochentag, clustert Wegziele mit vermutlich identischen Wegzwecken, vergleicht die Positionen der Wegziele mit Informationen zur Flächennutzung und statistischen Angaben aus einer schwedischen Volksbefragung und ordnet schließlich jedem Wegziel einen wahrscheinlichen Zweck zu. Zusätzlich wird eine Einschätzung der Sicherheit dieser Zuordnung angegeben.

Eine kartographische Umsetzung der Ergebnisse, die Berechnungen mit dem Makropaket für 39 ausgewählte Studienteilnehmer ergaben, wird im letzten Teil der Arbeit gezeigt und erläutert.

Danksagung

Professor Buchroithner vom Institut für Kartographie (IfK) der TU Dresden half mir sehr, diese Arbeit auf unkomplizierte Weise in Kooperation mit der ETH Zürich durchzuführen. Dafür und für die vielen fachübergreifenden Anregungen bin ich ihm dankbar.

Professor Hurni nahm mich wieder freundlich im Institut für Kartographie (IKA) der ETH Zürich auf. Er unterstützte mich sowohl fachlich in den Fragen der automatisierten GIS-Kartographie als auch durch das technische Umfeld des Instituts. Vielen Dank!

Die höchst interessante Aufgabe von Professor Axhausen aus dem Institut für Verkehrsplanung und Transportsysteme (IVT) der ETH Zürich hat mir Einblicke in die Verkehrsverhaltensforschung, einem spannenden Anwendungsfeld meines theoretischen Fachwissens, verschafft. Darüber habe ich mich sehr gefreut. Außerdem hat mir die Organisation durch Professor Axhausen mit vielen für mich sehr hilfreichen Besprechungen an Eckpunkten der Arbeit geholfen.

Ganz besonders danken möchte ich den Diplom-Ingenieuren Stefan Schönfelder vom IVT und Sven Etzold vom IfK für ihre große Unterstützung in den praktischen Fragen – Stefan Schönfelder für die verkehrswissenschaftlichen und Sven Etzold für die GIS-bezogenen Hilfestellungen. Sie haben mir beide unermüdlich entscheidende Tipps und Verbesserungsvorschläge gegeben.

Selbständigkeitserklärung

Hiermit erkläre ich, dass ich die von mir am heutigen Tage der Diplomkommission der Fachrichtung Geowissenschaften eingereichte Diplomarbeit zum Thema

„Entwicklung von GIS-Funktionen zur Analyse und
Visualisierung GPS-basierter Mobilitätsdaten“

vollkommen selbständig und nur unter Benutzung der in der Arbeit angegebenen Literatur angefertigt habe.

Dresden,

TECHNISCHE UNIVERSITÄT DRESDEN

Aufgabenstellung für die Diplomarbeit

Studiengang: Kartographie

Name des Diplomanden: Frau Uta Samaga

Thema: Entwicklung von GIS-Funktionen zur Analyse und Visualisierung
GPS-basierter Mobilitätsdaten

Zielsetzung:

Im Rahmen der Verkehrssicherheitsstudie „Rätt Fart“ (Richtige Geschwindigkeit) in der schwedischen Stadt Borlänge wurden ab dem Jahr 2000 Positionen von ca. 400 privat und gewerblich genutzten Fahrzeugen bis zu 2 Jahre lang mittels GPS-Technik aufgenommen. Teile dieses komplexen Datensatzes können nun für Studien im Bereich der Verkehrsverhaltensforschung genutzt werden, die über die Analyse der Verkehrssicherheit hinausgehen. Damit sind die Rätt Fart-Daten eine potentielle Ergänzung zu den üblichen Mobilitätsdaten aus Haushaltsbefragungen.

Die GPS-Daten aus Borlänge sind insbesondere deshalb von Interesse, weil erstmalig Informationen über die regionalen Wege von Verkehrsteilnehmern über einen so langen Zeitraum erhoben wurden – wenn auch nur solche, die mit dem Auto zurückgelegt wurden. Dabei erfordert jedoch die Menge und die Komplexität der GPS-Daten eine möglichst automatisierte Datenverarbeitung. Da die „Testfahrer“ keine weiteren Informationen zu ihren automobilen Wegen liefern mussten (rein „passives“ Beobachten), müssen übliche Attribute wie Zweck oder genaue Destination nachträglich zugespielt werden.

Ziel dieser Arbeit ist deshalb die Entwicklung verschiedener GIS-Funktionen, die die Voraussetzungen zur Analyse von Aktivitätenräumen der Studienteilnehmer schaffen. Die zu entwickelnden Module sollen den drei unten genannten Themenbereichen entsprechen. Dabei werden die Übergänge von einem zum nächsten Modul jedoch fließend sein.

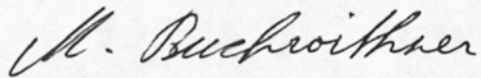
- (1) Bildung von Clustern „wahrscheinlich identischer“ Aufenthaltsorte, die durch die Nutzung verschiedener benachbarter Parkplätze bzw. Systemfehler bei der GPS-Aufzeichnung entstanden sind.
- (2) Verschneidung der Mobilitätsdaten mit dem lokalen Flächennutzungsplan und weiteren digital vorliegenden Informationen zur Siedlungsstruktur wie z.B. Standorte wichtiger Einkaufszentren etc., um die eigentlichen Wegzwecke zu identifizieren.
- (3) Ermittlung und Einbeziehung zeitlicher und räumlicher Regelmäßigkeiten der Mobilität: Bewegungsmuster einzelner Teilnehmer (z.B. 8 Std. arbeiten) bzw. Rhythmen typischer Aktivitäten wie „Einkaufen“ oder „Arbeiten“ sollen ermittelt und in die Bestimmung der Wegzwecke einbezogen werden.

Die Ergebnisse dieser Analyse sind anschließend anschaulich zu visualisieren. Dazu können bereits bestehende kartographische Visualisierungsmethoden von Mobilitätsmustern evaluiert und angewendet oder neue entworfen werden. Ein A1-Poster über die durchgeführten Bearbeitungsschritte ist zu gestalten.

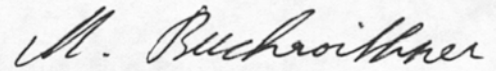
Betreuer: Prof. Dr. M. F. Buchroithner, Dipl. Ing. S. Etzold
(Institut für Kartographie, TU Dresden),
Prof. Dr. L. Hurni (Institut für Kartographie, ETH Zürich),
Prof. Dr. K. W. Axhausen, Dipl.-Ing. S. Schönfelder
(Institut für Verkehrsplanung, Transportsysteme, ETH Zürich)

Ausgehändigt am: 1. November 2002

Einzureichen am: 1. Mai 2003



Prof. Dr. Manfred Buchroithner
Prüfungsausschuß



Prof. Dr. Manfred Buchroithner
Betreuender Hochschullehrer

Inhaltsverzeichnis

Kurzfassung	I
Danksagung	II
Selbständigkeitserklärung	III
Aufgabenstellung	IV
Inhaltsverzeichnis	VI
1. Einleitung und Aufbau der Arbeit	1
2. Messung, Analyse und Modellierung räumlicher Aspekte der Mobilität	3
2.1 Ausprägungen der individuellen Mobilität im Raum	3
2.2 Individuelle Mobilität mit dem Automobil	5
2.3 „Traditionelle“ Mobilitätsstudien: Erhebungsmethodik und Datenverfügbarkeit	6
2.4 Studien mittels GPS-Technik.....	8
2.5 Geo-Informationssysteme und Mobilitätsstudien	13
3. Verwendete Software	15
3.1 Geographische Informationssysteme	15
3.2 Statistiksoftware.....	15
3.3 Ergänzende Software	15
4. Borlänge: Mobilitätsdaten aus GPS-Messungen und GIS-Bearbeitung	16
4.1 Die Studie „Rätt Fart i Borlänge“	16
4.2 Abgeleitete Mobilitätsdaten	17
4.3 Zusätzliches Datenmaterial	20
5. Automatisierte Auswertung.....	31
5.1 Anforderungen und Struktur des automatisierten Auswertungsprozesses	31
5.2 Vorverarbeitung der Daten	34
5.3 Clusterbildung – Stopp-Positionen	35
5.4 Clusteranalyse – wahrscheinliche Wegzwecke.....	41

6.	Ergebnisse	51
6.1	Visualisierung der Ergebnisse des Makropakets	51
6.2	Qualität der Ergebnisse aus verkehrswissenschaftlicher Sicht	58
7.	Ausblick	65
7.1	Verbesserungen bei den Studienbedingungen	65
7.2	Verbesserungen bei der Aufbereitung der Rohdaten	66
7.3	Verbesserungen im Analyseprozess	66
8.	Literatur	69
Anhang A:	Glossar	73
A 1:	Begriffe der Mobilitätsforschung	73
A 2:	Abkürzungsverzeichnis	75
Anhang B:	„Datengeschichte“	76
Anhang C:	Codierungen zwischen Flächennutzungen und Reisezwecken	78
C 1:	Einbeziehung des digitalen Flächennutzungsplans	78
C 2:	Einbeziehung verschiedener „Points of Interest“	80
Anhang D:	Quellcodes der Makros	82
Anhang E:	Kartenbeispiele	99

1. Einleitung und Aufbau der Arbeit

In der schwedischen Kleinstadt Borlänge wurden im Rahmen einer Verkehrssicherheitsstudie während zwei Jahren insgesamt ca. 400 Autos mit GPS-Geräten ausgestattet. Mit diesen wurden alle Bewegungen der Fahrzeuge aufgenommen und die jeweiligen Positionen und Zeiten gespeichert. Die so entstandenen Daten berichten nun, wann welches *Fahrzeug* an welchem Ort der Stadt war. Für die Verkehrsverhaltensforschung ist aber wichtig, was *eine Person* zu einem Zeitpunkt für eine gewisse Dauer an einem bestimmten Ort zu einem speziellen Zweck gemacht hat. Um die Daten aus Borlänge dafür nutzen zu können, müssen also aus den Informationen über die Fahrzeugpositionen Informationen über die Fahrer der Autos gewonnen werden. Was die Mobilitätsforscher dabei im Einzelnen interessiert, was Mobilität ist und wie man sie misst, wird der Schwerpunkt des zweiten Kapitels sein.

Die Software, die für die praktischen Teile der Arbeit zum Einsatz kommt, wird im dritten Kapitel genannt.

Um die gewünschten Informationen über die Studienteilnehmer zu extrahieren, stehen zusätzlich zu den GPS-Daten Flächennutzungsinformationen und Ergebnisse einer Verkehrsbefragung zur Verfügung. Die Flächennutzung beschreibt *die Orte* näher, an denen sich Personen aufhielten. Aus der Verkehrsbefragung ist bekannt, was typischerweise zu bestimmten *Zeiten* getan wird. Durch diese Zusatzinformationen sind Rückschlüsse darüber möglich, was jemand wo getan hat. Die einzelnen Datenquellen werden im vierten Kapitel vorgestellt.

Die Datenmengen sind allerdings so groß, dass es nicht möglich ist, manuell vorzugehen. Deshalb wird in dieser Arbeit ein Algorithmus entwickelt, der aus den numerischen Positions- und den weiteren Daten Informationen über das Mobilitätsverhalten der Personen, die Fahrzeuge gelenkt haben, gewinnen soll. Im fünften Kapitel wird der Aufbau dieses Algorithmus im Detail beschrieben.

Das sechste Kapitel legt einerseits dar, wie Resultate des Algorithmus kartographisch dargestellt werden sollten. Andererseits werden Plausibilitätsprüfungen der Ergebnisse aus verkehrswissenschaftlicher Sicht durchgeführt.

Wie in Zukunft verfahren werden könnte, um derartige Analyseaufgaben noch besser lösen zu können, und welche günstigeren Ausgangsbedingungen dafür nötig wären steht im abschließenden siebten Kapitel.

2. Messung, Analyse und Modellierung räumlicher Aspekte der Mobilität

2.1 Ausprägungen der individuellen Mobilität im Raum

Nach dem Fremdwörterbuch des Dudenverlags ist Mobilität „1. die (geistige) Beweglichkeit, 2. die Beweglichkeit von Individuen oder Gruppen innerhalb der Gesellschaft bzw. 3. die Häufigkeit des Wohnsitzwechsels einer Person“ (Drosdowski, Müller, Scholze-Stubenrecht und Wermke, 1990). Aus dem Blickwinkel der Verkehrsforschung und aktionsräumlichen Analyse von Mobilität kann diese Definition auf die *rekurrente Mobilität* beschränkt werden, d.h. „auf Bewegungen, durch die nicht, wie bei der Migration, der Wohnsitz verlegt wird, sondern die immer wieder von der Wohnung ausgehen und zu ihr zurückführen“ (Schwesig, 1988, S.29).

Diese Mobilität macht einen Großteil unseres alltäglichen Lebens aus, da sich die funktionale Differenzierung unserer Städte in Industriegebiete, Einkaufszonen, Wohngebiete etc. stets weiterentwickelt und auch durch Siedlungskonzepte wie beispielsweise das Punkt-Axiale-Raumkonzept gezielt gefördert wird. Dadurch sind die persönlichen Beziehungsnetzwerke ausgeweitet worden, d.h., jeder Bewohner einer Stadt hat eigene räumliche Bezugspunkte, an denen er (immer wieder) bestimmten Aufgaben oder Bedürfnissen nachgeht (Canzler, 1999).

Dabei ist die „Pendelwanderung zwischen Wohnstandort und Arbeitsplatz [...] die wichtigste Form rekurrenter Mobilität, nicht aber die einzige. Ebenso dazugehören unter anderem Einkaufsfahrten [...], Vergnügungs-, Besuchs- und Bildungsreisen und Erholungsfahrten am Wochenende und Ferienreisen. Die beiden letzten Kategorien werden in der Regel im Rahmen der Aktionsraumforschung nicht in die Analyse einbezogen“ (Schwesig, 1988, S.29).

Die o.g. Bezugspunkte, die eine Person immer wieder aufsucht, spannen den so genannten *Aktivitätenraum*¹ auf. Diese Aktivitätenräume können als geometrische Indikatoren für Mobilitätsverhalten definiert werden. Um ihre Größe und Struktur zu bestimmen stehen verschiede-

¹ englisch: activity space

ne Verfahren zur Verfügung. Durch sie wird bestimmt, wie wahrscheinlich es ist, dass sich eine Person an einem Ort aufhält, oder wie intensiv welche Regionen einer Stadt, eines Landkreises etc. von einer Person genutzt werden. Details dazu finden sich z.B. in Schönfelder und Axhausen (2003).

Die Ausdehnung der Aktivitätenräume wird durch drei Faktoren wesentlich beeinflusst (vgl. Golledge und Stimson, 1997):

- durch die Lage des Wohnortes,
- durch Positionen von Orten, an denen die Person regelmäßigen Aktivitäten nachgeht, und
- durch die Wege, die sie im Raum nutzt, um von einem Ziel zum anderen zu gelangen.

Inzwischen liegen erste Daten zum längerfristigen individuellen Mobilitätsverhalten vor. So ist ein aktueller Forschungsschwerpunkt in diesem Bereich, die Untersuchung von Räumen, die Verkehrsteilnehmer tatsächlich für ihre Aktivitäten aufsuchen (Schönfelder und Axhausen, 2002). Bisher konzentrierte sich ein Großteil der Studien eher auf die Potentiale der Mobilität, die sich aus physischen oder organisatorischen Möglichkeiten der Teilnehmer ergaben, da so detaillierte und große Datenmengen wie sie heute dank des fortgeschrittenen Stands der Erhebungstechnik zur Verfügung stehen nicht zu erfassen waren. Zu den bisherigen Themen gehörten also eher Fragen nach der Kenntnis von Räumen („mental maps“ etc., vgl. Downs und Stea, 1977) als nach der tatsächlichen Nutzung.

Die heutigen Studien zur Analyse der Aktivitätenmuster untersuchen also konkrete Mobilitätsnetzwerke – sowohl einzelner Stadtbewohner als auch bestimmter Bevölkerungsgruppen mit jeweils ähnlichen soziodemographischen Merkmalen. Kern dieser Forschungsarbeiten sind die Fragen nach der Art und teilweise auch nach dem Grund der Tätigkeit, welche die Personen zu einer bestimmten Zeit für eine bestimmte Dauer an einen bestimmten Ort ausführt. Es werden dabei verschiedenste Mobilitätscharakteristika evaluiert: z.B. die räumliche Ausdehnung der Aktivitätenräume (s.o.), die Anzahl der pro Tag unternommenen Fahrten und ihre Streckenlängen, signifikante Dauern für bestimmte Aktivitäten, die Verkehrsmittelwahl, typische Aktivitätsketten und ihre Häufigkeit (z.B. Wohnen – Arbeiten – Wohnen, Wohnen – Einkaufen – Freizeit – Wohnen) oder die Routenwahl, um zu einem bestimmten Ziel zu kommen (d.h., ob jeweils der kürzeste oder ein anderer Weg gewählt wurde).

Ein Beispiel für Ergebnisse solcher Mobilitätscharakteristika aus dem Bereich der Routenwahl aus der Lexington Area Travel Data Collection Studie von 1997 (Jan, Horowitz und Peng, 2000) ist, dass ein Fahrer in der Regel zwar jeweils identische Wege zu identischen Zielen nutzt, fahren aber verschiedene Fahrer häufig ein bestimmtes Ziel an, variieren die von ihnen gewählten Wege. Außerdem zeigte sich bei dieser Analyse, dass die Autofahrer zwar oft andere als die im Straßennetz metrisch kürzesten Wege wählen, die Zeiten, die sie dafür benötigen, aber den Zeiten für die im GIS ermittelten kürzesten Strecken entsprechen. (Jan *et al.*, 2000)

Die so gewonnenen Erkenntnisse über das Verkehrsverhalten der Bevölkerung und die Gründe für spezifisches Verhalten können dann z.B. dazu eingesetzt werden, um verkehrs- und siedlungsplanerische Maßnahmen zu ergreifen und auf diese Weise Verkehrsströme zu lenken oder einzudämmen.

2.2 Individuelle Mobilität mit dem Automobil

Lebten wir im Jahre 1880, so könnten wir im Brockhaus über das „Automobil“ lesen: „Name, der den manchmal sonderbaren, von einem Explosionsmotor bewegten Fahrzeugen gegeben wurde. Diese heute vergessene Erfindung hat nur Fehlschläge und die Missbilligung der wissenschaftlichen Autoritäten erlebt“. Und George Clemaseau würde zwei Jahre später erklären: „Der selbstfahrende Wagen, den in Deutschland Benz und Daimler soeben Kaiser Wilhelm vorgestellt haben, ist gefährlich, stinkt, ist unbequem, ganz sicherlich lächerlich, zu schnell und zum Vergessenwerden verurteilt“. (aus: Knoflacher, 1999, S.122)

Doch der Verlauf des 20. Jahrhunderts hat anderes gelehrt. Gut hundert Jahre später steht im Fremdwörterbuch des Dudenverlags noch lediglich, dass ein Automobil ein Kraftfahrzeug bzw. ein Kraftwagen ist und sich von „selbst-beweglich“ aus dem Griechischen und Lateinischen ableitet (Drosdowski *et al.*, 1990). Sämtliche Statistiken über den Automobilismus des letzten Jahrhunderts belegen Erich Kästners Aussage „die Zeit fährt Auto“ durch eindruckliche Steigerungen in der Automobilverbreitung etc. (Schmidt, 1990). Psychologische und physiologische Ursachen dafür beschreibt Knoflacher in seinem Artikel „Wo uns das Auto erwischt hat“ (Knoflacher, 1990).

Die in dieser Arbeit zu Vergleichszwecken herangezogene schwedische Verkehrserhebung (vgl. Kapitel 4.3.3) belegt den Trend ebenso. 2001 wurden ca. 60% aller Wege in Schweden

mit dem Auto zurückgelegt. In den USA waren es 1997 sogar 90% der Wege, die mit privaten Fahrzeugen gefahren wurden (Murakami, Wagner und Neumeister, 1997).

Doch auch wenn ein derart großer Teil der Wege mit dem privaten PKW erledigt wird, müssen bei Studien, in denen ausschließlich mit dem Automobil gefahrene Strecken analysiert werden (so wie bei dieser), stets zwei Tatsachen in Erinnerung behalten werden. Erstens fehlen immer sämtliche mit öffentlichen Verkehrsmitteln, dem Fahrrad und zu Fuß zurückgelegte Wege einer beobachteten Person. Und zweitens kann, wenn die Automobilbewegung automatisch beobachtet wird und die Fahrer sich vor Fahrtantritt nicht identifizieren müssen, nicht mit Sicherheit gesagt werden, wer das Auto gelenkt hat und wie viele und welche weiteren Personen noch im Fahrzeug saßen.

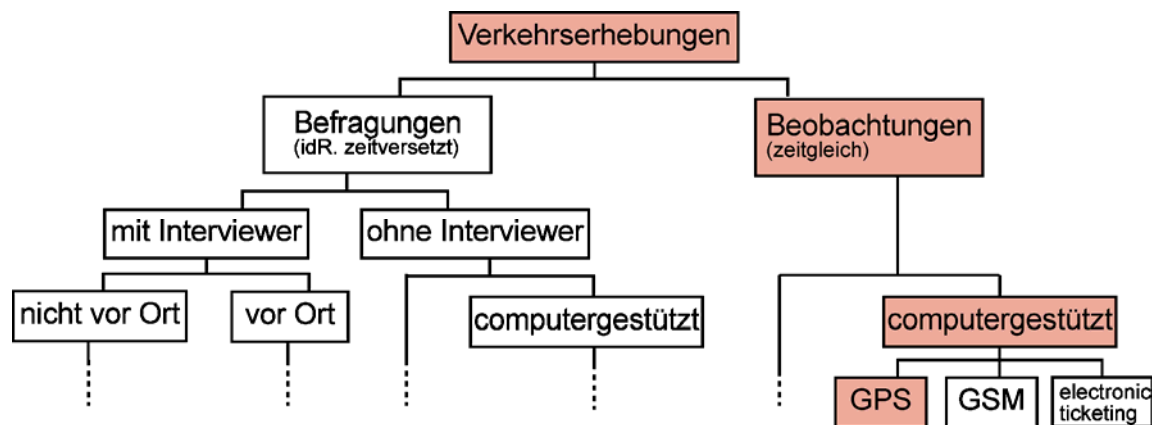
2.3 „Traditionelle“ Mobilitätsstudien: Erhebungsmethodik und Datenverfügbarkeit

Es existieren verschiedene Methoden, mit denen die Mobilität von Personen und Haushalten aufgezeichnet werden kann. Sie unterscheiden sich nach der Erhebungsmethode (Befragung oder Beobachtung), der Art der Aufzeichnung der Daten (computergestützt vs. „traditionelle“ analoge Verfahrensweisen), dem Zeitpunkt der Erhebung (zeitgleich mit dem Verkehrsverhalten oder zeitversetzt, d.h. nachträglich) und bei Befragungen danach, ob ein Interviewer bei einer Befragung anwesend ist oder nicht (vgl. Abbildung 1).

Bisher basierten empirische Studien zum individuellen Verkehrsverhalten in der Regel auf persönlichen *Befragungen*, bei denen die Verkehrsteilnehmer ihre Wege und Aktivitäten retrospektiv, d.h. zeitversetzt, berichteten. Üblicherweise geschah dies mittels eines Wegetagebuchs (ohne Anwesenheit eines Interviewers) oder in direktem Kontakt mit dem Befragungsunternehmen per Telefon (Anwesenheit eines Interviewers bzw. Unterstützung durch einen Interviewer). Um den Aufwand für die Befragten so gering wie möglich zu halten, wurden meist Stichtagsbefragungen durchgeführt, die hohe Rücklaufquoten und somit angemessene Datengenauigkeit für die Verkehrsplanung garantierten. Aus diesem Grund laufen auch z.B. in Deutschland (KONTIV – Kontinuierliche Erhebung zum Verkehrsverhalten) und Schweden (RES – Schwedische nationale Verkehrserhebung, vgl. Kapitel 4.3.3) nationale Verkehrserhebungen in Form von Befragungen zu jeweils einem Berichtstag ab (Infas und Diw, 2001; Sika Institut, 2001).

Anstelle von Befragungen können auch *Beobachtungen* durchgeführt werden. Diese haben den Vorteil, gar keine Auskunftsbereitschaft oder -fähigkeit der Studienteilnehmer zu erfordern, deshalb auch nicht in das Mobilitätsverhalten einzugreifen und somit „reales, weitgehend unverfälschtes Verkehrsverhalten“ messen zu können (Kreitz, 2002, S.31). Da bei diesen Verfahren die Belastung für die Studienteilnehmer sehr niedrig ist, können auch Daten über einen längeren Zeitraum leichter und präziser erfasst werden (Schönfelder, Axhausen, Antille und Bierlaire, 2002; Kreitz, 2002).

Abbildung 1 Schema zur Einteilung von Erhebungsmethoden zum Mobilitätsverhalten



Die eingefärbten Kästchen geben die für diese Arbeit zutreffenden Kategorien wieder. ²

Quelle: angelehnt an Kreitz (2002), S.30

Dass zeitgleiche Beobachtungen präzisere Ergebnisse als nachträgliche Befragungen liefern können, konnten beispielsweise Wolf, Loechl, Myers und Arce (2001b) in der California Statewide Household Travel Survey GPS Study zeigen. Dabei wurden Fahrten der Teilnehmer mittels GPS-Technik automatisiert aufgezeichnet. Die zurückgelegten Strecken wurden dann

² GSM steht für Global System for Mobile Communications und mit ihm werden Daten per Handy übertragen (weitere Informationen zu dieser Technologie unter GSM Association, 2003). Beim „electronic ticketing“ kaufen die Reisenden ihre Fahrscheine auf elektronischem Wege, so dass über diesen Weg Informationen zum Mobilitätsverhalten gewonnen werden können. (vgl. Kreitz, 2002)

kartographisch am Bildschirm aufbereitet und die Fahrer im Nachhinein nochmals zu ihren gefahrenen Strecken telefonisch befragt, sodass letztendlich Verkehrsbeobachtungen und Befragungen zum Verkehrsverhalten direkt miteinander verglichen werden konnten. Schon bei den ersten Vergleichen zeigte sich eine klare Tendenz, dass bei den per GPS beobachteten Wegen mehr Fahrten pro Tag aufgezeichnet als später am Telefon angegeben wurden. Bei einer ersten Stichprobe waren es ca. 20% der Wege, die beim Interview fehlten (Wolf *et al.*, 2001b). Dieser Wert lässt sich aber nicht verallgemeinern, denn bei ähnlichen Studienbedingungen in anderen Städten wich dieser Wert nach oben und nach unten ab. Als erwiesen gilt jedoch, dass bei Befragungen weniger Wege angegeben werden. (Wolf, 2003)

Allerdings liefern Verkehrsbeobachtungen meist erst dann aufschlussreiche Ergebnisse, wenn sie durch Befragungen ergänzt werden, in denen soziodemographische Informationen ermittelt werden. Außerdem fehlen bei Beobachtungen ohne ergänzende Befragungen (sog. ‚passive monitoring‘) Hinweise zu den Motivationen, Reisezielen etc. Diese fehlenden Informationen müssen, wenn sie nicht mit zusätzlichen Befragungen ermittelt werden können, durch aufwändige Verfahren so weit wie möglich rekonstruiert werden. Ein solches Verfahren wird in dieser Arbeit entwickelt und beschrieben.

2.4 Studien mittels GPS-Technik

2.4.1 Technik der Verkehrsbeobachtung mit GPS

Satellitengestützte Verkehrsbeobachtungen gehören gemäß oben beschriebener Einteilung der Erhebungsmethoden zu *zeitgleichen, computergestützten Beobachtungen*.

Mit einer geometrischen Genauigkeit von bis zu 3 m und sekundengenauen Zeiten liefern GPS-basierte Daten eine sehr präzise Basis für die weitere Auswertung. Ein großer Vorteil ist, dass auch die räumlichen Daten inklusive der Routenwahl, die bei Befragungen besonders umständlich zu ermitteln ist, bereits digital vorliegen. Die Studienteilnehmer sind während der Beobachtung insofern kaum belastet, als dass sie keine Erinnerungsleistung erbringen müssen, nicht durch die ständigen Interviews die Motivation zur Teilnahme verlieren und insgesamt weniger Zeit aufwenden müssen. Für Verkehrsanalysen, bei denen die Teilnehmer ihre Wege selbst retrospektiv berichten, geben Murakami *et al.* (1997) 20 Minuten Zeitaufwand pro Teilnehmer und Tag an. Erfolgt ein Interview am Telefon, müssen sogar bis zu 60 Minuten einkalkuliert werden. Bei den mit GPS erfassten Bewegungen fällt die personalintensive

Betreuung für Interviews weg. Der Zeitaufwand für die beobachteten Personen entfällt ebenso gänzlich oder verringert sich erheblich, denn selbst für die mit PDAs unterstützte Untersuchung in Lexington geben die Teilnehmer einen Aufwand pro unternommener Fahrt von lediglich einer Minute an. Die Wege und ihre Ziele werden, wie oben bereits beschrieben, in der Regel vollständiger, räumlich und vor allem zeitlich genauer erfasst als bei Befragungen. Gerade die Zeitangaben werden in Interviews meist auf 30, 15, 10 oder auch 5 Minuten gerundet, die Auswertung der mit GPS aufgezeichneten Reisedaten belegt aber, dass die tatsächlichen Abfahrtzeiten ohne jegliche Spitzen verteilt sind. Gleiches gilt für die gemessenen Wegstrecken im Vergleich zu den von Verkehrsteilnehmern im Interview angegebenen. Es können außerdem sehr einfach zusätzliche Größen wie z.B. Geschwindigkeiten gespeichert werden. (Murakami *et al.*, 1997; Murakami und Wagner, 1999; Kreitz, 2002; Wolf, Guensler und Bachman, 2001a; Schönfelder *et al.*, 2002)

Aufgrund der längeren Beobachtungszeiträume ist es zusätzlich möglich, nicht nur, wie in den bisherigen Stichtagsbefragungen die Unterschiede im Verkehrsverhalten zwischen Personen zu untersuchen (inter-personal level), sondern auch die Mobilitätsmuster von einzelnen Personen über einen bestimmten Zeitraum zu beschreiben (intra-personal level) (Schönfelder *et al.*, 2002).

Nachteile entstehen durch die aufwändige Technik, die für eine GPS-gestützte Beobachtung benötigt wird. Es werden zwar ständig neuere, kleinere und vor allem leistungsfähigere Modelle entwickelt, die eine flexiblere Nutzung zulassen (GeoStats, 2003). Doch derzeit ist die Stromversorgung noch ein großes Problem. In Fahrzeugen kann das GPS-Gerät über den Zigarettenanzünder versorgt werden, bei anderen Verkehrsmitteln, bei denen transportable Geräte zum Einsatz kommen müssen, sollten noch andere Lösungen gefunden werden (wie z.B. leistungsfähigere Batterien). Also ist man bei der Mobilitätsbeobachtung noch häufig auf fest in Fahrzeugen installierte GPS-Empfänger angewiesen, d.h. es werden auch nur mit diesem Verkehrsmittel zurückgelegte Wege ermittelt, oder die Aufzeichnungsdauer darf nicht zu lang sein. Zudem variiert die Positionsgenauigkeit und Signalqualität nach der Art des Verkehrsmittels, weil stets verkehrsmittelspezifische Störfaktoren auftreten. Beispielsweise können in öffentlichen Verkehrsmitteln nur Signale empfangen werden, wenn die Testpersonen nahe am Fenster sitzen. Ein weiteres Hindernis sind Oberleitungen der Straßenbahnen, die Teile der GPS-Signale abfangen und somit die Genauigkeit verringern. (Draijer, Kalfs und Perdok, 2000).

Generell können Positionen per Satellit nur dann aufgezeichnet werden, wenn der Empfänger eine ungestörte Verbindung zu mindestens 4 der 24 im Orbit befindlichen Satelliten hat³. Daraus resultieren Schwierigkeiten bei der Positionsüberwachung in engen Häuserschluchten, unter dichten Baumkronen, in Tunneln etc., da durch sie der örtliche Horizont eingeschränkt wird. (Murakami *et al.*, 1997, Murakami und Wagner, 1999)

Ein spezielles Problem tritt bisher noch bei den meisten GPS-Empfängern auf: Die GPS-Geräte brauchen einige Zeit nach dem Einschalten, um ihre erste Position zu ermitteln. Draijer *et al.* (2000) geben 15 bis 60 Sekunden an, Wolf, Guensler, Frank und Ogle (2000) 45 Sekunden. Das ist die Dauer, in der die zur Berechnung des Standortes auf der Erde nötigen Satelliten geortet werden und aus deren Signalen die erste Position berechnet wird. Bei der Aufzeichnung der Bewegungen fehlen dadurch in der Regel die Startpositionen, die jedoch durch die jeweils letzten Stopp-Positionen rekonstruiert werden können. Bei der Verarbeitung der Routenwahlinformationen und somit der Berechnung der zurückgelegten Strecken können die ersten fehlenden Sekunden nicht immer exakt rekonstruiert werden. Dies spielt bei der Ermittlung der Wegzwecke der angefahrenen Orte aber keine Rolle. Störender ist die ungenaue Information zur Abfahrtszeit. Vor allem bei der Frage danach, was noch als Fahrt zu werten ist und was als Datenfehler gilt, sind diese Sekunden entscheidend (vgl. Kapitel 4.2). Durch leistungsfähigere Batterien in den GPS-Empfängern könnte dieses Problem aber gelöst werden. Die Geräte könnten dann permanent angeschaltet bleiben und lückenlos Daten von den Satelliten empfangen.

Bei Mobilitätsbeobachtungen ohne direkte Zusatzinformation durch die Verkehrsteilnehmer muss im Nachhinein festgelegt werden, welche Stopps wirklich Stopps und nicht bloß Wartezeit an einer Ampel etc. waren. Einen Schwellwert von 120 Sekunden bestimmten z.B. Wolf *et al.* (2001a) in einer 2000 an der Georgia Tech University durchgeführten Studie dafür, dass „keine Bewegung“ ein Wegziel bedeutete. Mit diesem Wert erreichten sie die beste Deckung zwischen Interviewergebnissen und GPS-Daten-Auswertung. Derartige Werte bleiben jedoch immer nur Schätzungen (vgl. Kapitel 4.2).

Speziell bei der Aufzeichnung von Automobilbewegungen gibt es ohne zusätzliche Angaben der Studienteilnehmer fehlende Informationen zur Anzahl der mitgenommenen Personen, da-

³ Eine Abdeckung mit 8 oder noch besser 12 Satelliten ist aber wünschenswert, um die Genauigkeit zu verbessern. (Wolf, 2003)

zu, wer tatsächlich gefahren ist und wer eventuell an welchem Ort aus- bzw. zugestiegen ist. Es fehlen außerdem die Wegzwecke, die für Mobilitätsanalysen entscheidend sind (Murakami und Wagner, 1999).

Doherty, Noël, Lee-Gosselin, Sirois und Theberge (2001) gehen sogar so weit, dass sie sagen, ohne zusätzliche Informationen von den Verkehrsteilnehmern können keine Erklärungen für das Mobilitätsverhalten gewonnen werden.

2.4.2 Beispiele bisheriger Studien

Seit Mitte der 1990er Jahre wurden bereits verschiedene Untersuchungen mit GPS gestützten Verfahren durchgeführt (vgl. Tabelle 1). Die Motivationen dazu unterscheiden sich. Einige Studien verfolgen umweltrelevante Themen wie die Kopplung zwischen Geschwindigkeit der Autos und Abgasmengen. Andere zielen darauf, wie oben beschrieben, traditionelle Reisetagebücher zu ersetzen. Weitere haben als Studienziel die Verbesserung der GPS- oder zusätzlichen PDA-Technik. Oder die Arbeiten wollen die Verkehrssicherheit analysieren und verbessern, wie auch die „Rätt fart“-Studie in Borlänge (vgl. Kapitel 4.1). (Schönfelder *et al.*, 2002; Lee-Gosselin, 2002; FHWA, 2003a)

Die meisten Studien befassen sich bisher lediglich mit der Aufzeichnung von Automobilbewegungen, da die Größe und Stromversorgung der Empfänger bei anderen Verkehrsmitteln noch Probleme bereiten. Eine Ausnahme bildet die vom Verkehrsministerium der Niederlande durchgeführte Studie, die alle Reisemodi der Teilnehmer analysierte (Draijer *et al.*, 2000). In Zukunft wird diese Restriktion aber durch verbesserte Techniken aufgehoben werden (GeoStats, 2003; Wolf, 2003).

Studien zur Verkehrssicherheit zählen normalerweise zu den „rein passiven“ Beobachtungen im Sinne der in Kapitel 2.3 erläuterten Einteilung von Verkehrserhebungen. Ihnen fehlen also sämtliche zusätzliche Informationen wie die Wegzwecke, die in anderen das Verkehrsverhalten analysierenden Studien erhoben werden. Dieses Merkmal ist der wesentliche Unterschied dieser im Nachgang der Verkehrssicherheitsstudie „Rätt fart i Borlänge“ durchgeführten Verkehrsverhaltensstudie über die Testfahrer in Borlänge im Vergleich zu anderen bisher mit GPS evaluierten Verkehrsverhaltensdaten. Bisher wurden jeweils zusätzliche Informationen zu den Studienteilnehmern und vor allem zu den von ihnen zurückgelegten Wegen im studienbegleitenden oder -nachfolgenden Kontakt gewonnen. (Schönfelder *et al.*, 2002)

Tabelle 1 Beispiele bisheriger Studien zum Thema GPS-gestützte Verkehrsanalyse

Studienname/ -ort, Jahr	Veröffentlichungen	Ziel der Studie
Lexington Area Travel Data Collection Test, 1996	Murakami <i>et al.</i> (1997), Murakami und Wagner (1999), FHWA (2003a)	<ul style="list-style-type: none"> • Vergleich elektronisches Wegetagebuch (per PDA – personal digital assistant) und durch GPS aufgezeichnete Reisedaten mit „traditionellen“ CATI (computer assisted telephone interview) • Untersuchung zu GPS Datengenauigkeit und -vollständigkeit • Auswertung der Routenwahl (Jan et al., 2000)
Diverse Städte in den Nieder- landen, 1997	Draijer <i>et al.</i> (2000)	<ul style="list-style-type: none"> • Untersuchung, inwiefern mobile GPS-Geräte für alle Fortbewegungsarten herkömmliche Reisetagebücher ersetzen können
Quebec City	Doherty <i>et al.</i> (2001), Lee-Gosselin (2000), Lee-Gosselin (2002)	<ul style="list-style-type: none"> • Entwicklung GIS-Algorithmus zur Identifizierung von Wegzielen, Reisezeiten, Routenwahl und Geschwindigkeit • Entwicklung eines Verkehrsmodells zur Vorhersage des Verkehrsflusses
Georgia Tech Experiment 2000 in Atlanta, Georgia	Wolf <i>et al.</i> (2001a)	<ul style="list-style-type: none"> • Untersuchung zu den Möglichkeiten, traditionelle Reisetagebücher gänzlich durch GPS-Messungen zu ersetzen • Test verschiedener GPS-Geräte auf Genauigkeit in bezug auf GIS-Straßennetz
SMARTAQ und Teilprojekt Atlanta Metropolitan Regional Travel Survey 2000	Wolf <i>et al.</i> (2000), French, Bachman und Frank (2001)	<ul style="list-style-type: none"> • Aufbau einer Landnutzungsdatenbank für Abgleich mit Aktivitäten • Beobachtung von mehr als 1000 Fahrzeugen bis zu 2 Jahre lang • Vergleich GPS-Daten und Reisetagebuch, Ermittlung von Parametern zu Korrekturen von in Tagebüchern vergessenen Wegen etc. • Verkehrssicherheit und -verhalten • Aktivitätenanalyse der Teilnehmer • Untersuchung zur Luftverschmutzung

Quelle: nach Schönfelder *et al.* (2002) S. 6 und Wolf *et al.* (2001a) S. 3

2.5 Geo-Informationssysteme und Mobilitätsstudien

Geographische Informationssysteme (GIS), auch als Geo-Informationssysteme bezeichnet, sind computergestützte Werkzeuge und Methoden, die in der Lage sind, raumbezogene Daten zu erheben, zu verwalten, abzuändern und auszuwerten (Buhmann, Bachhuber und Schaller, 1996). Diese Definition deutet schon das 4-Komponenten-Modell an, aus dem GIS aufgebaut sind: Sie integrieren Funktionen zur **Erfassung**, **Verwaltung** (bestehend aus Datenmodellierung, -strukturierung und -speicherung), **Analyse** und **Präsentation** von Geodaten (EVAP) (Bill und Fritsch, 1991).

Für den Einsatz von GIS gibt es in Mobilitätsstudien mehrere Einsatzfelder:

Ein großes Problem bisheriger Verkehrsbefragungen war die Lokalisierung von Wegstarts und -zielen der Verkehrsteilnehmer. In der Regel wurden sie über Adressangaben der Teilnehmer realisiert. Doch allzu leicht traten Fehler durch falsche Buchstabierungen oder fehlerhafte Einträge in den ‚address-matching‘-Datenbanken auf bzw. kannten die Teilnehmer nicht immer die Adressen der von ihnen aufgesuchten Orte. Um diese Probleme zu beseitigen, konnten erste GIS-basierte Tools geschaffen werden, mit denen Studienteilnehmer die von ihnen aufgesuchten Orte über graphische Bedienoberflächen auf dem eigenen PC, übers Internet oder auf Laptops mit Touchscreens berichten können. Darin wurden neben den bekannten ‚address-matching‘-Funktionen zusätzlich interaktive Karten implementiert. (Adler, Rimmer, Bandy und Schellinger, 2000; Kreitz, 2002).

Ein weiteres Anwendungsfeld für GIS in Mobilitätsuntersuchungen ist die Analyse. Auch hier gibt es bereits einige Beispiele. Dazu gehören die Algorithmen zur Erkennung von Stopppositionen aus GPS-Rohdaten, die in Quebec entwickelt wurden (Doherty *et al.*, 2001), und die verschiedenen Methoden zur Messung und Visualisierung von Aktivitätenräumen von der ETH Zürich (Schönfelder und Axhausen, 2002; Schwarze und Schönfelder, 2001). Auch die in dieser Arbeit entwickelten Methoden gehören in das Umfeld der im GIS automatisierten Analyse von Mobilitätsdaten.

Und schließlich werden, da Analyseergebnisse graphisch ausgegeben werden können, auch Mobilitätsdaten im GIS visualisiert. Der letzte Teil dieser Arbeit zählt dazu sowie zahlreiche bestehende Arbeiten z.B. die oben schon erwähnten Methoden zur Visualisierung von Aktivitätenräumen (Schönfelder und Axhausen, 2002; Schwarze und Schönfelder, 2001).

Da es also in dieser Arbeit vornehmlich um die Realisierung eines Analysewerkzeugs im GIS geht, fiel die Wahl des dafür verwendeten GIS auf ArcInfo mit seiner Makrosprache ArcMacroLanguage (AML). Auf dieser Basis stehen bereits eine Vielzahl von GIS-Funktionen zur Verfügung und können durch die Programmierung in AML auf das spezielle Problem angepasst und miteinander verknüpft werden.

Für den Präsentationsteil wurde ArcMap von ESRI benutzt, da es umfangreichere Funktionen zur Kartengestaltung anbietet als ArcInfo.

3. Verwendete Software

3.1 Geographische Informationssysteme

Es wurden ausschließlich GIS-Pakete der Firma ESRI verwendet, um die Kompatibilität der Daten zu gewährleisten. Wie bereits oben angedeutet, kam für die Automatisierung der Auswertung ArcInfo mit seiner Makrosprache AML zum Einsatz. Zur Datenkonvertierung von den aus Schweden gelieferten Shape-Files (vgl. Kapitel 4.3.2) und teilweise zur Digitalisierung wurde ArcView verwendet. Für die Datenorganisation und -verwaltung bot sich ArcCatalog des neuen ArcGIS-Systems an, eine Art Explorer für GIS-Datensätze mit recht komfortabler Datenvorschau, und für die Visualisierung das ArcMap des gleichen Systems.

3.2 Statistiksoftware

Sowohl zur Bewältigung der sehr großen Datenmengen als auch zur Ermittlung verschiedener statistischer Werte mussten Statistikprogramme genutzt werden. Um die Ausgangsdaten zu verwalten, diese in durch ArcInfo importierbare Dateien umzuwandeln und um Referenzwerte aus der schwedischen Volksbefragung zu ermitteln wurde SPSS von der SPSS AG verwendet. Mit dem SAS-Paket des SAS Instituts wurden Look-Up-Tabellen für das im Rahmen dieser Arbeit entwickelte Makropaket generiert und diverse Plausibilitätsprüfungen in den Daten vorgenommen.

3.3 Ergänzende Software

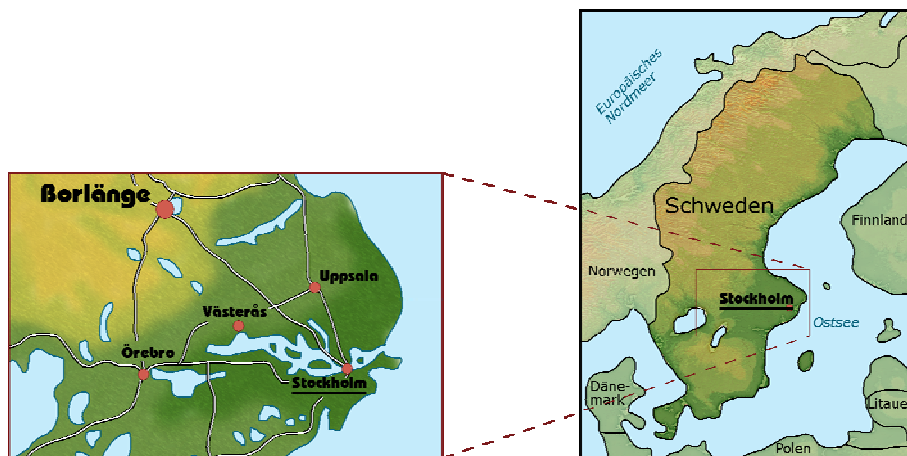
Um Graphiken anzufertigen, wurden die Bildbearbeitungssoftware Photoshop (Version 6) von Adobe und die Vektorgraphiksoftware Freehand (Versionen 9 und 10) der Firma Macromedia genutzt.

Für weitere Aufgaben wie Textverarbeitung und einfache Tabellenkalkulationen kamen die Microsoftprodukte MS Word und MS Excel zum Einsatz.

4. Borlänge: Mobilitätsdaten aus GPS-Messungen und GIS-Bearbeitung

Das Testgebiet dieser Arbeit, die Stadtregion von Borlänge, liegt ca. 200 km nordwestlich von Stockholm in Schweden (vgl. Abbildung 2 und Anhang E).

Abbildung 2 Lage der Stadt Borlänge

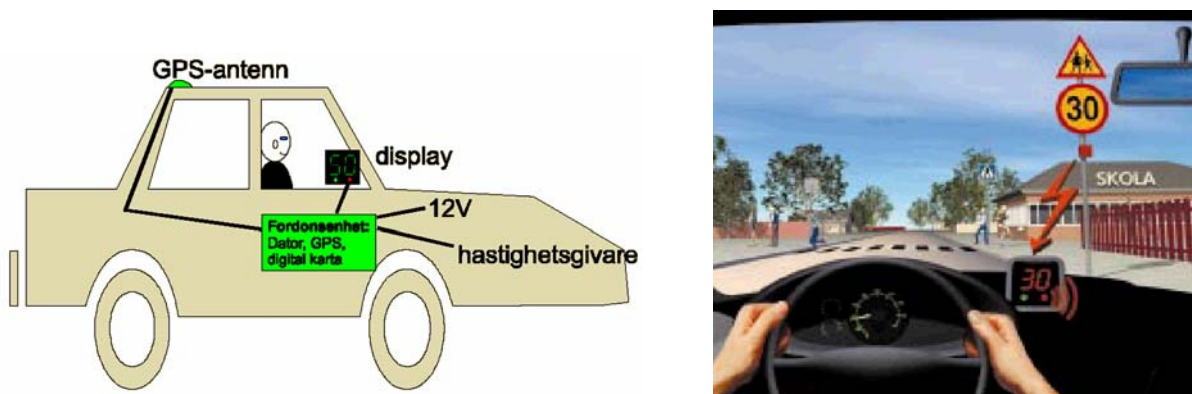


4.1 Die Studie „Rätt Fart i Borlänge“

Die Verkehrssicherheitsstudie „Rätt Fart i Borlänge“ wurde von der schwedischen Straßen-Administration (Statens Vägverket – Swedish National Road Administration) im Rahmen der Initiative ISA (Intelligent Speed Adaption) durchgeführt. Dabei wurde ein Warnsystem getestet, das Autofahrer alarmiert, wenn ihr Wagen die zugelassene Höchstgeschwindigkeit der gerade befahrenen Straße überschreitet. Die Autos waren dazu mit einem GPS-Gerät ausgestattet, das mit dem digitalen, mit zulässigen Höchstgeschwindigkeiten attribuierten Straßennetz von Borlänge verknüpft war. Zusätzlich befand sich ein 6*4 cm² großes Display im Wagen, auf dem die aktuelle zugelassene Höchstgeschwindigkeit abgelesen werden konnte und das ein Warnsignal aussendete, wenn die Geschwindigkeit überschritten wurde (siehe Abbildung

3). Einige Testfahrzeuge wurden zusätzlich mit der Funktion „aktives Gaspedal“ ausgestattet. Bei dieser Funktion wird das Gaspedal blockiert, wenn die maximal zulässige Geschwindigkeit erreicht ist. (Rätt fart i Borlänge, 2003a, Rätt fart i Borlänge, 2003b)

Abbildung 3 Die GPS-Technik der Verkehrssicherheitsstudie „Rätt Fart i Borlänge“



Quelle: linke Abbildung aus Rätt fart i Borlänge, 2003b und rechte Abbildung aus Rätt fart i Borlänge, 2003c

4.2 Abgeleitete Mobilitätsdaten

Die in der Verkehrssicherheitsstudie gewonnenen Positionsdaten der Fahrer wurden von der schwedischen Straßenadministration für die vom Institut für Verkehrsplanung und Transportsysteme (IVT) der ETH Zürich durchgeführte Studie „Structure and use of human activity spaces“ zur Verfügung gestellt.

Gemäß der in Kapitel 2.3 und 2.4 erläuterten Möglichkeiten zur Erhebung von Verkehrsdaten gehören die für diese Studie bereitgestellten Datensätze zu rein passiven, computergestützten, mit GPS-Technik gewonnenen Beobachtungen. Sie stehen somit im Gegensatz zu den in Kapitel 2.4.2 erwähnten bisherigen Studien, in denen zu den passiven Beobachtungen zusätzliche Befragungen durchgeführt wurden. Im Nachgang wurden zwar auch für die Borlänge-Testpersonen noch einige wenige soziodemographische Merkmale erhoben, Informationen über die von ihnen zurückgelegten Wege, die Wegziele und dazugehörigen Wegzwecke feh-

len aber in dieser Befragung. Bei einer Beobachtungsdauer von durchschnittlich mindestens einem halben Jahr pro Teilnehmer wäre dies im Nachhinein auch kaum möglich gewesen. Es stehen somit erstmalig Daten zum Mobilitätsverhalten von Individuen über einen so langen Zeitraum zur Verfügung, wenn auch ohne die sonst übliche Information zum Wegzweck etc. (Axhausen und Bierlaire, 2001; Schönfelder *et al.*, 2002; Schönfelder und Samaga, 2003)

Die einzelnen Positionen wurden bei der Verkehrssicherheitsstudie jeweils vom Fahrzeugcomputer auf die nächstgelegene Straße (Link) und die Position darauf umgerechnet, um die aktuelle Höchstgeschwindigkeit abzulesen. Um Speicherplatz zu sparen, wurde darauf verzichtet, die ursprünglichen xy-Positionen zu speichern (Auskunft von der Erhebungsfirma Columna vom 15.10.2002). Der zur Umrechnung von GPS- in Straßenlink-Positionen verwendete Algorithmus ist leider nicht bekannt. Die x- und y-Positionen der gelieferten Stopps können zwar im GIS problemlos berechnet werden, insgesamt wäre es jedoch wünschenswert, auch die Originalpositionen zu kennen, um den Snapping-Prozess, mit dem die Positionen auf den Straßen bestimmt wurden, nachvollziehen zu können und eventuell Fehlerquellen zu erkennen.

Ohne also die Fehlerquellen im Detail ermitteln zu können, mussten einige Stopp-Positionen eliminiert werden, bevor die Auswertung beginnen konnte, denn teilweise fehlten Angaben oder sie waren offensichtlich fehlerhaft (z.B. negative Anzahl der Fahrten pro Tag). Des Weiteren wurden Fahrten aussortiert, die eine zu kurze Reisezeit oder Aktivitätendauer hatten oder weitere nachfolgend erläuterte Kriterien erfüllten (eine stichwortartige Übersicht dieser Kategorien findet sich nochmals in Kapitel 5.2).

Der in dieser Arbeit übernommene Schwellwert für minimale Aktivitätendauern von 120 Sekunden ist in empirischen Untersuchungen in Atlanta ermittelt worden (Wolf, 2000). Doherty *et al.* (2001) gingen bei ihren Untersuchungen sogar von einem Wert von ca. 5 Minuten aus, in Austin (Texas) ermittelten Wolf *et al.* (2003b) einen günstigen Schwellwert von 45 Sekunden. Es gibt also kein allgemeingültiges Maß dafür, welche Zeitschwelle am günstigsten gewählt wird, um tatsächliche Wegziele, an denen einer Aktivität nachgegangen wird (die allerdings auch nur kurze Zeit in Anspruch nehmen kann wie z.B. jemanden absetzen oder einen Brief einwerfen), aus kurzzeitigen Fahrtunterbrechungen durch Staus, rote Ampeln u.ä. herauszufiltern. Da der Wert für minimale Aktivitätendauern im Quellcode lediglich einer Variablen entspricht, kann er aber jederzeit variiert werden, sodass die vorläufige Festlegung auf 120 Sekunden nicht überbewertet werden muss.

Erschwerend kommt bei dieser Problematik allerdings hinzu, dass durch die Verzögerung beim Starten der GPS-Empfänger (während der ersten Positionsbestimmung) Ungenauigkeiten bei den Zeiten für die Stillstände der Autos und die nachfolgenden Fahrzeugbewegungen entstehen (vgl. Kapitel 2.4.1 und Anhang A/ Abbildung 22).

Deshalb gestaltete sich auch die Bestimmung der minimalen Reisedauer schwierig. Es ist unrealistisch, dass jemand für weniger als ein bis zwei Minuten sein Auto nutzt und nicht zu Fuß geht bzw. das Fahrrad nimmt. Wenn aber für die erste Positionsbestimmung durch das Satellitensystem schon bis zu 60 Sekunden benötigt werden können (Draijer *et al.*, 2000), ergibt das natürlich eine große Schwankungsbreite. Für die minimal in Frage kommende Reisezeit wurde deswegen für diese Arbeit ein Schwellwert von 30 Sekunden eingeführt, um durch Fehler bei der GPS-Aufzeichnung entstandene Wege bzw. Stopps zu ermitteln. Dieser 30-Sekunden-Schwellwert repräsentiert somit Fahrten von 0 Sekunden bis in Extremfällen zu 90 Sekunden.⁴

Außerdem wurden Stopps eliminiert, die auf den Ausfallstraßen von Bolänge⁵ lagen, da es sich bei ihnen mit sehr großer Wahrscheinlichkeit um Fahrten ins weitere Umland handelte und an diesen Positionen kein Wegziel erreicht wurde. Das gleiche Argument gilt für Fahrten von mehr als 25 km Länge und für durchschnittliche Reisegeschwindigkeiten über 50 km/h. Die einzelnen Kriterien für den Ausschluss sind im Code `short_trips.aml` in Anhang D aufgelistet. Um wie viele Stopps es sich pro Testfahrer handelt, ist Anhang B zu entnehmen.

Erschwerend kommt hinzu, dass sich die Fahrer - im Gegensatz zu anderen Verkehrsstudien – bei Fahrtantritt nicht identifizieren mussten. Es nahmen zwar nur Haushalte an der Studie teil, die jeweils die gleiche Anzahl Fahrzeuge und Testfahrer besaßen. Allerdings bringt allein diese Tatsache keine Sicherheit darüber, welche Person tatsächlich am Steuer saß. Die Autofahrer könnten eventuell über die Analyse ihrer Fahrstile nachträglich identifiziert werden. Dabei treten aber weitere Probleme wie eine mögliche Veränderung des Fahrstils, wenn das Signal

⁴ Langfristig müssten Tools geschaffen werden, die einen Abgleich in den Rohdaten vornehmen: Je nachdem wie stark die erste Position eines Fahrzeugs nach dem Start von der letzten Stopp-Position vor dem Ausschalten abweicht, müssten die Variablen für minimale Aktivitäten und Wegdauern variiert werden.

⁵ Das für die Untersuchung zur Verfügung stehende Wegenetz beschränkt sich auf die Gemeinde Borlänge, vgl. Kapitel 4.3.1.

für Geschwindigkeitsüberschreitung ertönt, zutage. Diese Analyse gehört aber nicht zur Aufgabenstellung der Diplomarbeit.

So wurden für die Entwicklung des Algorithmus vorerst lediglich Haushalte ausgewählt, in denen nur ein Autofahrer und ein Wagen vorhanden waren. Als weitere Einschränkung wurden erst mal lediglich vollzeitbeschäftigte Arbeitnehmer und Rentner als zwei charakteristische soziodemographische Gruppen ausgewählt. Es handelt sich hierbei um 28 Vollzeitbeschäftigte und 11 Rentner. Die Makros sind jedoch so geschrieben, dass sie erstens für jeden weiteren Testfahrer ebenso funktionieren und zweitens stets durch Spezifika für andere soziodemographische Bevölkerungsgruppen erweitert werden können. Rein technische Vorteile bot zusätzlich die Auswahl derjenigen Fahrer, die ihre Wohnadresse angegeben hatten. Der Algorithmus des Makros ist jedoch auch für alle anderen Studienteilnehmer kompatibel.

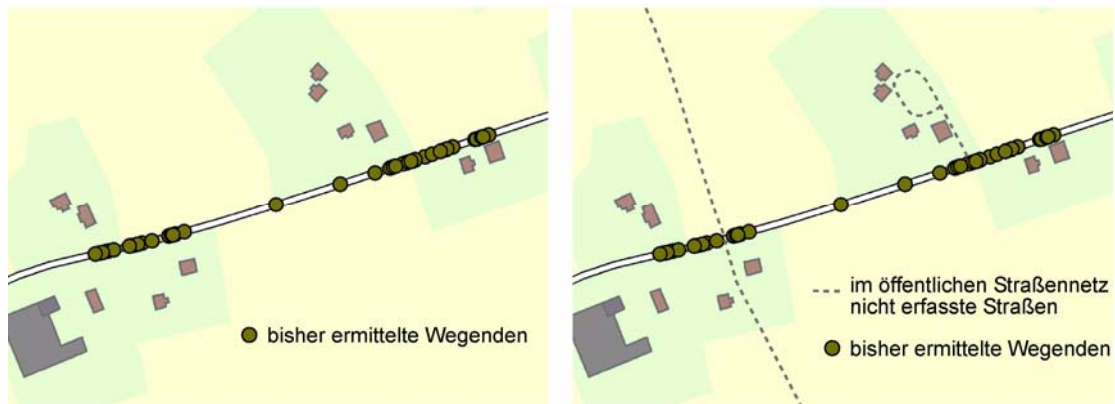
4.3 Zusätzliches Datenmaterial

4.3.1 Straßennetz

Das Straßennetz, auf das sich bereits die GPS-Messungen der Studie „Rätt Fart i Borlänge“ bezogen, umfasst lediglich öffentliche Wege (Schönfelder *et al.*, 2002). Da die Positionen der Testwagen während der Geschwindigkeitsstudie jeweils sofort in Straßenlink-Nummer und Position auf dem Link (in Metern) umgerechnet und auch nur so gespeichert wurden, ist es sinnvoll, genau dieses Straßennetz auch für diese nachfolgende Untersuchung zu übernehmen.

Allerdings fehlen sämtliche private Wege wie Hauszufahrten, Feldwege etc. Das hat den Nachteil, dass erstens Fahrten auf diesen Wegen nicht erfasst wurden und zweitens jeweils ein Stopp an den Stellen im Straßennetz ermittelt wurde, an denen die Wagen das öffentliche Straßennetz verließen (vgl. Abbildung 4). Das bedeutet, jedes Mal wenn ein Fahrer eine nicht im digitalen Netz erfasste Straße nutzte, wurde in der Nähe der Kreuzung zu dieser Straße die letzte Position des Wagens und somit eine Stopp-Position erfasst. Im weiteren Prozedere wird ihr dann eine Aktivität zugeordnet, ohne dass der Fahrer an dieser Stelle ein Wegziel hatte. Da es sein kann, dass er häufig eine bestimmte, im digitalen Netz nicht erfasste Straße nutzte, ist es auch möglich, dass scheinbar viele Wegziele an dieser auf den ersten Blick nicht sichtbaren Kreuzung liegen.

Abbildung 4 Mögliche fehlerhafte Stopp-Positionen durch fehlende private Wege



Das linke Bild zeigt die Situation, die derzeit aus den Daten hervorgeht. Es könnte sich hinter den Daten aber eine der rechten Abbildung ähnliche Situation verbergen.

Theoretisch wäre es möglich, diese fehlerhaften Stopps zu ermitteln, indem die nicht erfassten Wege nachträglich aufgenommen und entsprechend attribuiert werden. Dann müssten sämtliche Stopp-Positionen auf ihre Lage in Bezug auf diese ursprünglich nicht erfassten Wege überprüft und von Fall zu Fall bei der Wegzweck-Bestimmung speziell behandelt werden. Dieses Verfahren ist allerdings sehr aufwändig, vor allem die manuelle Digitalisierung der fehlenden Straßen, sodass im Rahmen dieser Arbeit darauf verzichtet wurde.

4.3.2 Flächennutzungsinformation

Es liegen für das Gebiet der Stadt Borlänge verschiedenste (leider teilweise widersprüchliche) Informationen zur Flächennutzung vor, die für die Ermittlung der Wegzwecke von Bedeutung waren und hier im Einzelnen vorgestellt werden.

Grundstückskarten, Übersichtspläne und Stadtpläne einschließlich ihrer Digitalisierung zum digitalen Flächennutzungsplan

Von der Stadt Borlänge lagen drei Arten von analogen Karten vor, aus denen ein digitaler Flächennutzungsplan abgeleitet wurde.

Die als erste einbezogene und wichtigste Quelle waren die *Grundstückskarten* (Lantmäteriet, 2002). Sie werden von der Vermessungsstelle in Gävle im Maßstab 1:12'500 herausgegeben und es lagen davon 4 Blätter⁶ mit Stand von 2002 vor. Die Grundstücksnutzung wird in diesen Karten nach 7 Klassen unterschieden. Aus ihnen wurden die Klassen ‚dichte Bebauung‘, ‚offene Bebauung‘, ‚Industriegebiet‘, ‚Landwirtschaftsgebiet‘ und ‚Wald und offenes Land‘ für den Flächennutzungsplan in Vektorform abgeleitet. Sie wurden für das gesamte Gebiet, für das die Grundstückskarten vorlagen, übernommen (Landmäteriet, 2002).

Die *Übersichtspläne* (ÖversiktsPlan –90) werden von der Stadt Borlänge und der Vermessungsstelle im Maßstab 1:50'000 herausgegeben. Sie lagen für diese Arbeit in drei Varianten vor: für die Stadt Borlänge, für die umliegenden Gemeinden und für die Gebiete zwischen den Gemeinden. Leider benutzen nicht alle die gleiche Legende. Außerdem widersprechen die Aussagen dieser Pläne teilweise denen der Grundstückskarten. Da die Übersichtspläne nur mit Stand von 1990 und in einem kleineren Maßstab vorlagen, wurden sie lediglich an Randbereichen einbezogen, wo die Grundstückskarten das Testgebiet nicht vollständig abdeckten, und in offenen Bereichen, in denen die Übersichtspläne eine größere Genauigkeit als die Grundstückskarten aufzeigten. Da die Klassifizierung der Übersichtspläne nicht direkt auf die der Grundstückskarten übertragen werden konnte, wurden Polygone mit der Nutzung ‚Ackerland‘ und ‚Forstwirtschaft‘ direkt übernommen (letztere wurde zu ‚Wald und offenes Land‘ hinzugezählt) und Flächen mit den Nutzungen ‚Wohnbebauung‘ und ‚Bebauung für Arbeitsplätze‘ als Flächennutzungskategorien zusätzlich eingeführt.

Die *Stadtpläne* von Borlänge und den umliegenden Gemeinden Romme, Torsång und Idkerberget werden vom städtischen „Ingenieursbüro Borlänge“ im Maßstab 1:10'000 herausgegeben. Ihnen ist jeweils ein Luftbild hinterlegt. Zusätzlich sind die Flächen nach den Nutzungsarten ‚Wohnviertel‘, ‚Industrie und Eisenbahngebiet‘, ‚öffentliche Gebäude‘, ‚Wald‘, ‚offenes Land‘ und ‚Park‘ eingefärbt. Die als ‚Park‘ ausgewiesenen Areale sind nicht plausibel, da nahezu alle nicht bebauten Flächen in der Stadt als Park gekennzeichnet sind. Die übrigen fünf Kategorien konnten allerdings als Ergänzung zu den Informationen aus den Grundstückskarten und Übersichtsplänen herangezogen werden.

Leider widersprechen sich die Karten immer wieder in ihren Angaben über die Flächennutzung. Insgesamt dürfte der vorliegende digitalisierte Plan trotzdem eine gute Annäherung an

⁶ Es sind die Kartenblätter mit den Bezeichnungen 13f0g, 13f1f, 13f1g und 13f2f.

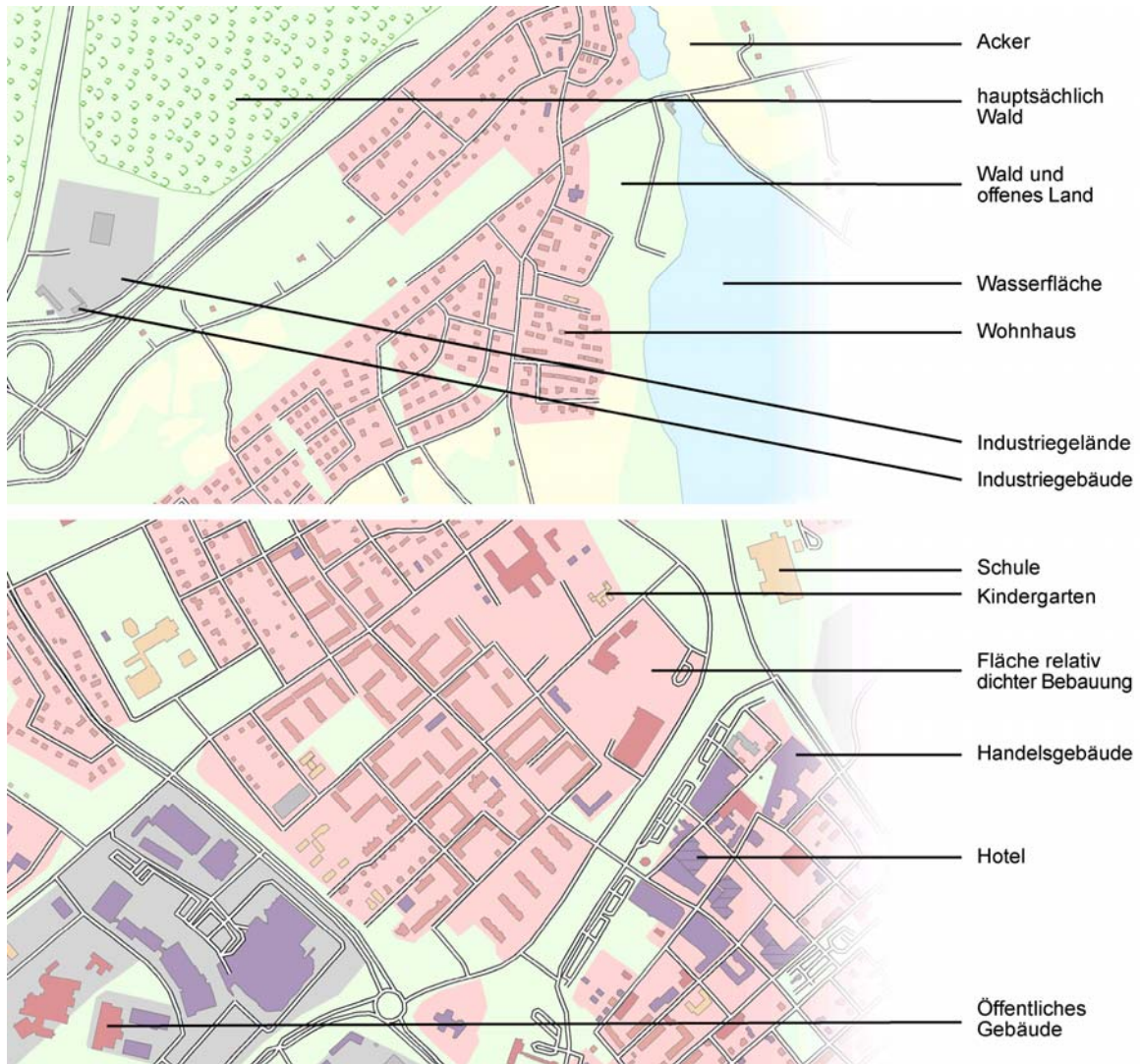
die tatsächlichen Flächennutzungen darstellen. Seine Aufgabe für diese Arbeit, als Indikator für potenzielle Wegzwecke zu dienen, kann er sicherlich erfüllen.

Einzelhäuser und ihre Einbeziehung in den Flächennutzungsplan

Die Stadt Borlänge stellte zusätzlich bereits digital vorliegende Flächeninformationen in Form von ArcView-Shape-Files zur Verfügung (© 2002 Stadsbyggnadskontoret, Borlänge kommun). Sie umfassen verschiedenste Informationen von öffentlichen Gebäuden, Einzelhäusern der Wohnbebauung und Schulen über Parks und Eisenbahnlinien bis hin zu Naturschutzgebieten und Wäldern. Leider waren die geometrische Genauigkeit und Plausibilität der Geometrien meist zu schlecht, als dass die Shapes in den digitalisierten Flächennutzungsplan hätten einbezogen werden können bzw. als dass sie sogar als Grundlage für die Digitalisierung hätten dienen können, wie es ursprünglich geplant war. Ein Grund könnte der Basismaßstab sein, denn sämtliche auf 1:10'000 basierenden Shapes passten schlecht, die auf 1:1'000 basierenden hingegen recht gut. Allerdings sollte 1:10'000 grundsätzlich kein zu kleiner Maßstab sein, betrachtet man die Maßstäbe der verwendeten analogen Karten (1:10'000, 1:12'500 und 1:50'000). Die auf dem Maßstab 1:1'000 basierenden Ebenen mit Nachführungsjahr 2001, d.h. die Ebenen mit Einzelgebäuden, konnten also übernommen und in den digitalen Flächennutzungsplan integriert werden. Es handelt sich hierbei um die Kategorien ‚Nachmittagsbetreuung‘, ‚Wohnhaus‘ (getrennt nach Wohnblocks und freistehenden Häusern), ‚kulturelles Gebäude‘, ‚Hotel‘, ‚Industriegebäude‘, ‚Kindergarten‘, ‚Handelsgebäude‘, ‚öffentliches Gebäude‘, ‚Schule‘ und ‚Sporthalle‘.

Die Flächennutzungsinformationen werden im AML-Teil `clus-analyse.aml` verarbeitet. Der Ablauf wird in Kapitel 5.4.1 genau beschrieben.

Abbildung 5 Ausschnitte des digitalisierten Flächennutzungsplans



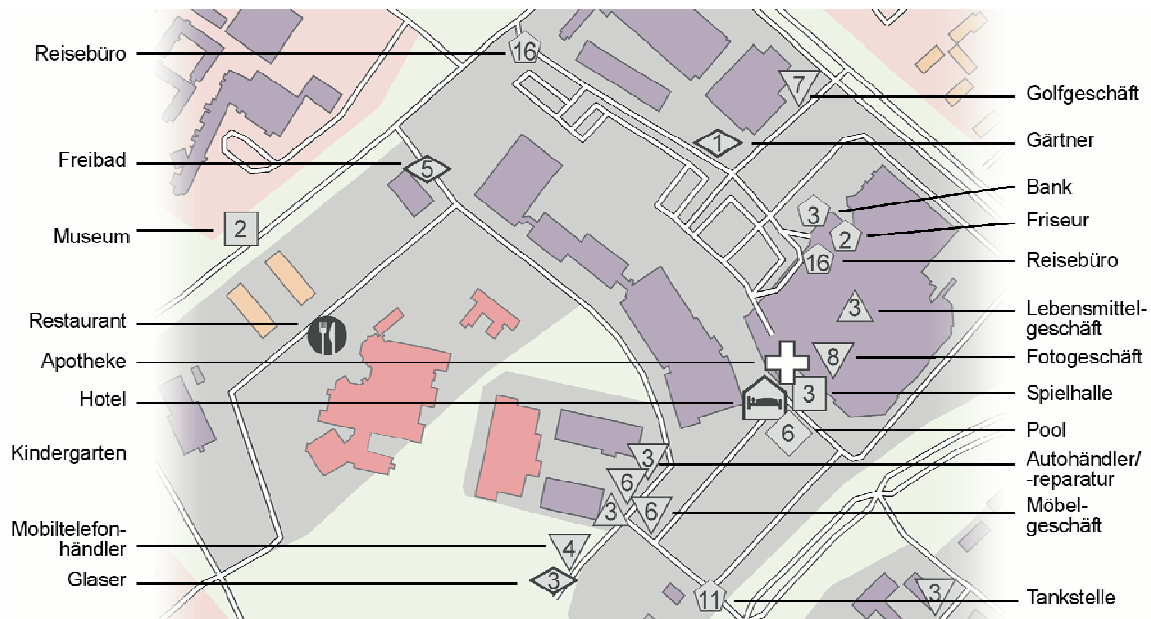
„Points of Interest“

Ebenso wurde ein Shape-File mit so genannten „Points of Interest“⁷ von der Stadt Borlänge bereitgestellt (© 2002 Stadsbyggnadskontoret, Borlänge kommun). Es enthält Punktoobjekte, die Örtlichkeiten wie Restaurants, Buchläden und Tankstellen repräsentieren. Der Datensatz scheint aber kein ausgewogenes Bild der Stadt Borlänge zu repräsentieren. Insgesamt sind beispielsweise Einkaufsgelegenheiten für den täglichen Bedarf deutlich unterrepräsentiert, andere Geschäfte wie Friseure etc. wirken hingegen stark überrepräsentiert. So wurden die in den Gelben Seiten vorhandenen Lebensmittelläden hinzudigitalisiert (Gula Sidorna, 2003). Langfristig sollten aber auch noch andere Einkaufsgelegenheiten wie beispielsweise Bäcker und Fleischer integriert und der Datenbestand sollte insgesamt geprüft und homogenisiert werden.

Anhang C 2: Einbeziehung verschiedener „Points of Interest“ enthält die gesamte Liste der Punkte. Sie wird im Makroteil `clus-analyse.aml` verwendet. Eine genaue Beschreibung des Ablaufs steht in Kapitel 5.4.1.

⁷ Mit „Points of Interest“ werden im Allgemeinen verortete Punkte gemeint, die von besonderer Bedeutung sind. In dieser Arbeit sind es Lokalitäten in der Stadt, an denen die Bewohner bestimmten Aktivitäten nachgehen. Dazu gehören beispielweise Friseure, Banken, Spielhallen oder auch Tankstellen und Campingplätze. Da es für „Point of Interest“ keine Entsprechung im Deutschen gibt, wurde in dieser Arbeit der englische Begriff verwendet.

Abbildung 6 Ausschnitt aus der ‚Points-of-Interest‘-Coverage

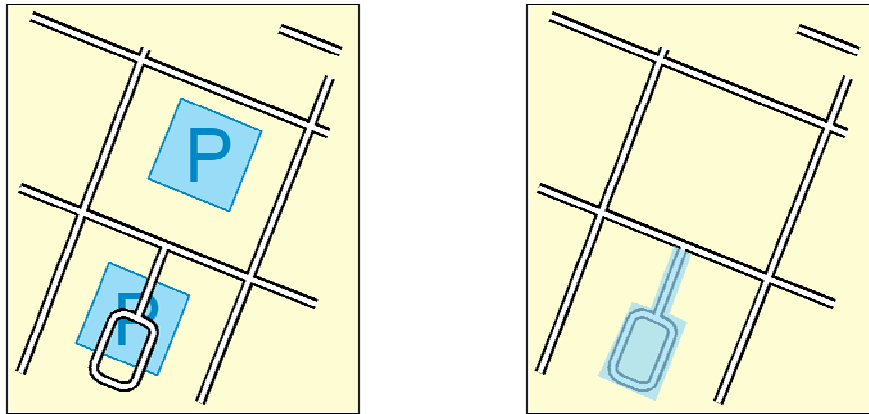


Parkplätze

Stopp-Positionen auf Parkplätzen sollten gesondert behandelt, d.h. alle Stopps auf einem Parkplatz jeweils zu einem Stopp-Cluster zusammengefasst werden. Denn es kann davon ausgegangen werden, dass das Wegziel identisch ist, wenn der Fahrer auf ein und demselben Parkplatz hält. Um dies zu realisieren mussten Parkplätze als Polygonflächen integriert werden. Später konnte dann anhand dieser Polygone analysiert werden, ob Punkte innerhalb oder außerhalb eines Parkplatzes liegen. Die Studienbedingungen hatten zur Folge, dass alle Fahrzeugpositionen auf den Straßen lagen (s.o.). Aus diesem Grund mussten die Parkplätze so digitalisiert werden, dass sie immer alle Straßen einbeziehen, die zum Parkplatz führen bzw. zu ihm gehören. Aus dem gleichen Grund konnten nur Parkplätze aufgenommen werden, denen im digitalen Straßennetz Straßen zugeordnet sind oder die Zugangsstraßen besitzen, weil davon ausgegangen werden kann, dass Autos, die auf diesen Zugangsstraßen ihre Stopp-Position haben, zum Parkplatz gefahren sind. In Spezialfällen können aber auch Parkplätze ohne Zugangsstraßen Parkplatzflächen zugeordnet werden (vgl. Abbildung 7 und Abbildung 8). Das Resultat ist die Coverage „PP2“. Im Code der AML `clus-pp.aml` werden die „Parkplatzpolygone“ einbezogen (vgl. 5.3 -

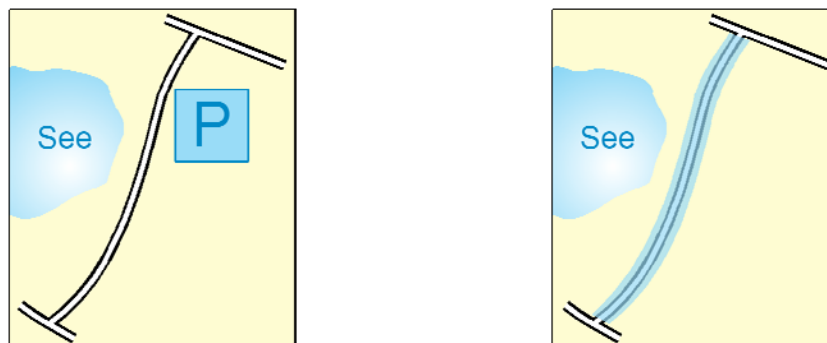
Ablauf der Clusterung und Anhang D).

Abbildung 7 Erläuterung zur Digitalisierung von Parkplätzen mit Zugangsstraßen



Im linken Bild ist die Ausgangssituation zu sehen: der untere Parkplatz besitzt eine Zugangsstraße und eine ihm zugeordnete Straße. In der Abbildung rechts ist die Digitalisierung erfolgt: Zu dem Parkplatz mit zugeordneten Straßen kann ein Polygon zugewiesen werden (hellblau); dem oberen Parkplatz kann keine Parkplatzfläche zugeordnet werden, weil er keine eindeutige Zufahrtsstraße hat.

Abbildung 8 Erläuterung zur Digitalisierung von Parkplätzen ohne Zugangsstraße



In diesem Beispiel hat der Parkplatz zwar keine eigene Zufahrtsstraße. Da es sich aber bei dem See um einen Badesee außerhalb der Stadt handelt, kann angenommen werden, dass Fahrzeuge, die ihre Stopp-Position in dieser Straße haben, auch auf dem Parkplatz parkten oder zumindest ihr Fahrtziel der Badesee war.

Digitalisierte Wohn- und Aktivitätenadressen der Teilnehmer

Die Wohnadressen derjenigen Teilnehmer, die diese bei der an die Verkehrssicherheitsstudie anschließenden Befragung angaben, wurden im IVT digitalisiert. Sie liegen als Shape-File vor und konnten in eine ArcInfo-Coverage umgewandelt und so in den Analyseprozess einbezogen werden. Dazu siehe Kapitel 5.4.1.

Neben den Wohnadressen wurden noch Arbeits- und Freizeitstätten der Teilnehmer digitalisiert, wenn sie sich aus den Angaben rekonstruieren ließen. Da diese Angaben aber bei mehr als 50% der Befragten fehlten oder zu ungenau waren, wurden sie nicht in den weiteren Prozess der Arbeit einbezogen.

4.3.3 Schwedische Verkehrserhebung RES

Die schwedische Verkehrsbefragung RES ist eine nationale Stichtagsbefragung, die umfangreiche Daten über das Mobilitätsverhalten in Schweden liefert. Detailliertere Informationen über diese Erhebung, auch im Vergleich zu anderen nationalen Verkehrsstudien, finden sich bei Sika Institut (2001).

Die Daten aus dieser Befragung standen als Vergleichswerte für die Auswertung der Borlänge-Daten zur Verfügung. Denn bei der RES-Befragung gaben die Teilnehmer natürlich vollständige Informationen zu ihren Wegen an, d.h. auch ihre Wegziele und-zwecke. Und letzteres fehlt bekanntlich in den Borlänge-Daten. So konnten bestimmte ‚Werte‘ aus der RES denen der Autofahrer aus Borlänge gegenübergestellt werden. Diese ‚Werte‘ der Verkehrsbefragung waren die bevorzugten Aktivitäten von bestimmten Bevölkerungsgruppen (Rentner, Vollzeitbeschäftigte und andere) mit bestimmten weiteren soziodemographischen Merkmalen (Geschlecht, Verfügbarkeit von Auto und Führerschein etc.) an bestimmten Wochentagen zu bestimmten Tageszeiten für bestimmte Dauern. Für die Testfahrer aus Borlänge konnten dann ebenso die Informationen zur Bevölkerungsgruppe, zu soziodemographischen Merkmalen und den Tagen, Uhrzeiten und Dauern der von ihnen verfolgten Aktivität abgerufen werden. Aus diesem Vergleich ergaben sich schließlich Vermutungen zu der fehlenden Information über die tatsächlich von den beobachteten Personen ausgeübte Aktivität. Näheres zu diesem Vorgehen wird in Kapitel 5.4.2 beschrieben.

Dieser Vergleichsdatensatz birgt allerdings einige Unsicherheiten in Bezug auf seine Aussagekraft im Rahmen des Borlänge-Projekts. Erstens kann es gut sein, dass sich das Verkehrs-

verhalten der Bewohner von Borlänge von dem durchschnittlichen in Schweden unterscheidet. Es wäre deshalb eigentlich ein Vergleichsdatensatz aus Borlänge wünschenswert. Ein solcher existiert allerdings nicht. Betrachtet man im gesamtschwedischen Datensatz nur Personen aus Borlänge, ist die Stichprobe zu klein (weniger als 100 Personen), um gehaltvolle Aussagen ableiten zu können. Außerdem reicht sie nicht aus, um einen Großteil der im GPS-Testdatensatz vorkommenden Fälle abzudecken.

Außerdem stimmen die Rahmenbedingungen der Borlänge-Studie und der RES-Befragung nicht überein. Die RES-Daten wurden durch eine Stichtagsbefragung gewonnen, die Borlänge-Fahrer nahmen an einer Langzeitstudie teil. Das bedeutet beispielsweise, dass in den RES-Daten die verschiedenen Aktivitätsmuster an Wochentagen und Wochenenden nicht dadurch zustande kommen, dass ein Teilnehmer sein Verhalten geändert hat, sondern dadurch, dass verschiedene Personen an verschiedenen Tagen befragt wurden. Die Tendenzen, die sich aus beiden Studien ableiten lassen, dürften jedoch trotzdem ähnlich sein, sodass der Vergleich Sinn macht.

Fraglich ist vielleicht noch eher, wie gut die aus einer *Verkehrsbefragung* gewonnenen Werte mit denen einer *Verkehrsbeobachtung* übereinstimmen. Murakami und Wagner (1999) stellten in einem Vergleich aus GPS-gestützter Verkehrsbeobachtung und Befragung zu den gleichen Wegen fest, dass die Zeiten und Wege der retrospektiv berichteten Reisen immer länger angegeben werden als sie tatsächlich sind. Eventuell spielen diese Ungenauigkeiten auch keine Rolle, da der Gesamtdatensatz inkl. der Flächennutzungsinformation zu unexakt ist. Gerade bei den kurzen Aktivitäten wie jemanden absetzen oder Einkaufen für den täglichen Bedarf könnte dieser Aspekt aber doch wichtig sein.

Um einen möglichst guten Vergleich zwischen den RES-Daten und den Borlänge-Fahrern zu ermöglichen, wurde aus den RES-Daten eine Modalwerttabelle abgeleitet. Das bedeutet, es wurde die am häufigsten ausgeübte Aktivität für verschiedene in allen möglichen Variationen miteinander kombinierte Parameter ermittelt. Diese Parameter waren

- das Geschlecht,
- der Status (bisher eingeteilt in Rentner, Vollzeitbeschäftigte und andere, eine weitere Aufsplittung könnte ergänzt werden),
- die Verfügbarkeit von Auto und Führerschein,
- die Wochentagskategorie (wochentags, samstags, sonntags),

- die Anfangszeit der Aktivität (eingeteilt in stündliche Startzeiten, d.h. z.B. fallen alle Startzeiten zwischen 7 und 8 Uhr morgens in eine Kategorie) und
- die Dauer der Aktivität (eingeteilt in verschieden lange Dauern: die erste Stunde wird dabei in 15-Minuten-Abschnitte unterteilt, alle Dauern bis zu vier Stunden werden mit einer Genauigkeit von einer halben Stunde gruppiert, alle längeren Aktivitäten dauern werden stündlich zusammengefasst).

Aus der Tabelle kann nun beispielweise abgelesen werden, was eine weibliche Rentnerin, die sowohl Auto als auch Führerschein zur Verfügung hat an Samstagen wahrscheinlich gemacht hat, wenn sie mit ihrer Aktivität in der Stunde von 9 bis 10 Uhr begonnen hat und diese Aktivität zwei bis zweieinhalb Stunden gedauert hat.

Die Einbindung dieser Tabelle in den Analyseprozess wird in Kapitel 5.4.2 ausführlich beschrieben.

5. Automatisierte Auswertung

5.1 Anforderungen und Struktur des automatisierten Auswertungsprozesses

Der Testdatensatz aus der Studie „Rätt Fart“ umfasst Positionsdaten von 400 Autofahrern mit insgesamt mehr als 160 000 Stopp-Positionen. Eine derart große Datenmenge kann nicht mehr manuell ausgewertet werden, da der Arbeitsaufwand zu groß würde. Also muss ein System entwickelt werden, in dem die Analyse so weit wie möglich automatisch abläuft. Da es sich um lokalisierte Daten handelt und in die Auswertung die in Kapitel 4.3 beschriebenen geographischen Informationen einbezogen werden müssen, bietet sich das GIS-Software-Paket ArcInfo mit seiner Makrosprache AML als Entwicklungsumgebung an.

Das zu programmierende Makropaket muss stabil laufen, also sämtliche Sonderfälle in den Datengrundlagen berücksichtigen und mögliche Fehler abfangen. Bei der Analyse sollte es annähernd so gute Ergebnisse liefern wie ein manuelles Verfahren, das von der Erfahrung und dem logischen Kombinationsvermögen der Bearbeiter profitiert. Das heißt, die automatisierte Auswertung muss in mehreren Schritten erfolgen und die logischen Verknüpfungen, die ein Bearbeiter bei anderen Studien für jede Testperson gedanklich erstellt, müssen möglichst verallgemeinert und in den Programmcode integriert werden. Der Analysevorgang kann sicherlich nicht im ersten Versuch in seiner gesamten Komplexität in Programmcode umgesetzt werden. Deshalb sollen die Makros in offenen Modulen ablaufen, die ständige Erweiterungen und somit eine stetige Verbesserung der Analysequalität erlauben.

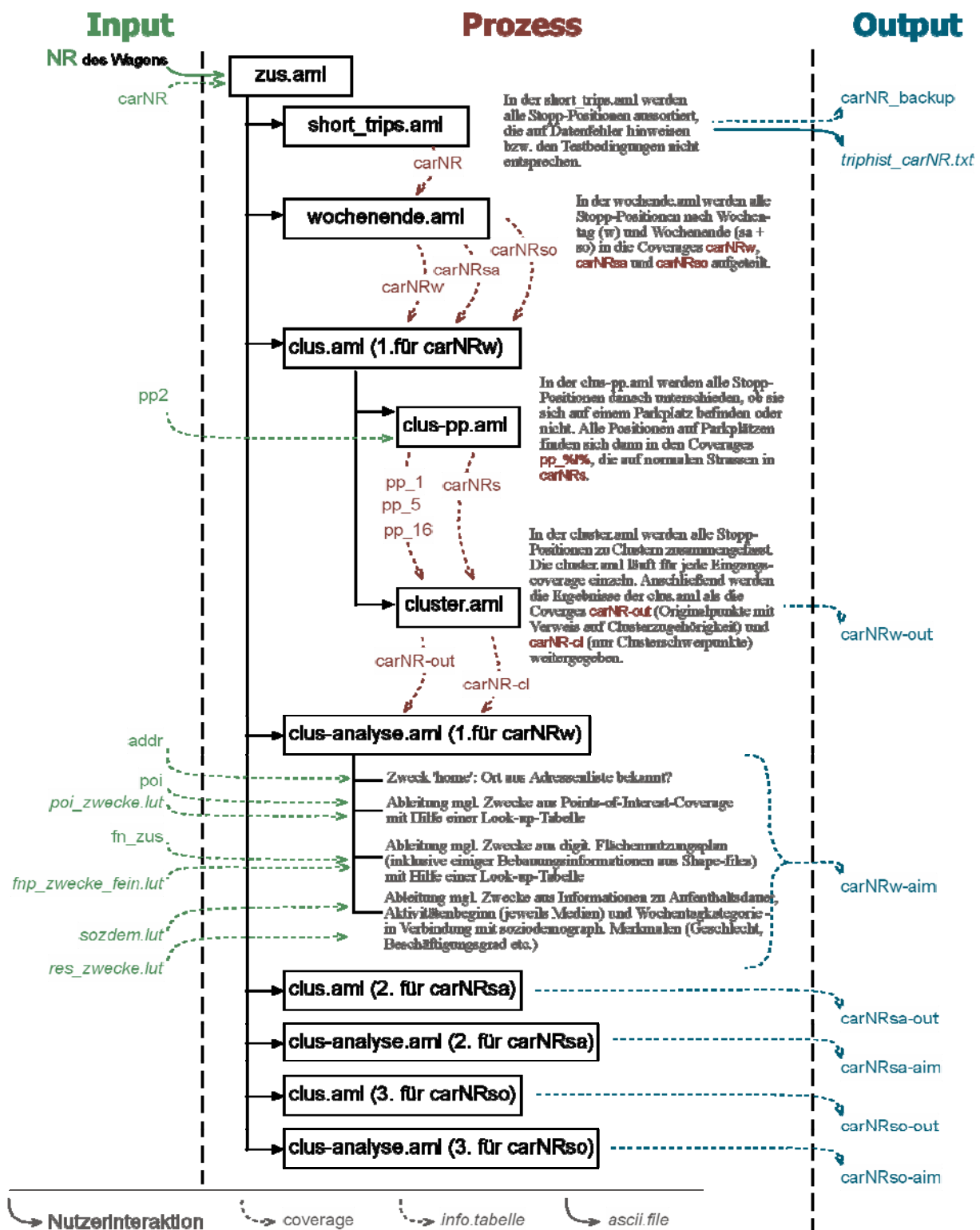
Die grobe Struktur des Makropakets ist in Abbildung 9 zu erkennen: Beim Start des Programms muss eine Testfahrernummer eingegeben werden („NR des Wagens“), woraufhin die entsprechende Coverage (*carNR*⁸) geladen wird. Sie enthält sämtliche Stopp-Positionen des jeweiligen Studienteilnehmers. Aus *zus.aml* werden dann alle Unterprogramme gestartet. Das erste davon ist *short_trips.aml*, hier werden alle fehlerverdächtigen Positionen von der weiteren Berechnung ausgeschlossen. Die bereinigte Coverage (*carNR*) wird an *wochenende.aml* weitergegeben, eine ASCII-Datei (*triphist_carNR.txt*) mit der Anzahl der eli-

⁸ In Klammern stehen die Namen der Coverages bzw. der Dateien.

minierten Stopps sowie ein Backup der Originalcoverage (`carNR_backup`) werden gespeichert. Im Teil `wochenende.aml` werden die Positionen nach ihrer Wochentagskategorie aufgeteilt und im weiteren Verlauf als drei getrennte Coverages behandelt (`carNRw`, `carNRsa` und `carNRso`). `Clus.aml` organisiert daraufhin die Zusammenfassung der einzelnen Stopp-Positionen zu Stopp-Clustern. Dieser Programmteil läuft dreimal hintereinander ab (für jede Wochentagskategorie einmal). Der Clusterprozess in `clus.aml` ist zweigeteilt. Zuerst prüft `clus-pp.aml` bei allen Punkten, ob sie innerhalb der Parkplatzflächen (`pp2`) liegen. Punkte auf Parkplätzen bilden temporäre Coverages (`pp_1`, `pp_2` usw.), Punkte auf dem übrigen Straßennetz eine weitere Coverage (`carNRs`). Im zweiten Teil, `cluster.aml`, findet die eigentliche Clustering statt. Als Ergebnis davon entstehen zwei Coverages (`carNR-out` und `carNR-cl`). In der ersten befinden sich die Originalpositionen mit einem zusätzlichen Item, das die Clusternummer angibt, zu dem der einzelne Punkt gehört. Ein Backup dieser Coverage wird herausgegeben (pro Durchlauf eine, d.h. `carNRw-out`, `carNRsa-out` bzw. `carNRso-out`). Die zweite Coverage, die in `cluster.aml` entsteht, enthält nur noch diejenigen Positionen, die ein Clusterzentrum darstellen. Im eigentlichen Analyseteil `clus-analyse.aml` werden schließlich die wahrscheinlichen Wegzwecke zu jedem Cluster ermittelt. `clus-analyse.aml` läuft ebenfalls für jede Wochentagskategorie einmal. Als erstes wird dabei die Wohnadresse bestimmt (Abgleich mit `addr`). Anschließend werden Wahrscheinlichkeiten aus den ‚Points of Interest‘ (`poi` und `poi.lut`), der Flächennutzung (`fn_zus` und `fnp_zwecke_fein.lut`) und den zeitlichen Informationen (`sozdem.lut` und `res_zwecke.lut`) für den vermutlichen Wegzweck eines Clusters bestimmt. Die so festgelegten wahrscheinlichen Wegzwecke pro Stopp-Cluster werden als Coverage herausgegeben (`carNRw-aim`, `carNRsa-aim` bzw. `carNRso-aim`).

Die Quellcodes der Makros in AML finden sich in Anhang D.

Abbildung 9 Ablauf des Makropakets zur automatisierten Auswertung



5.2 Vorverarbeitung der Daten

Die ersten beiden AMLs des Schemas aus Abbildung 9, `short_trips.aml` und `wochenende.aml`, übernehmen die Vorverarbeitung der Positionsdaten.

In `short_trips.aml` werden sämtliche Stopp-Positionen aussortiert, deren Daten darauf hinweisen, dass sie aus technischen Fehlern bei der GPS-Aufzeichnung oder durch die Aufnahmebedingungen der Geschwindigkeitsstudie resultieren (vgl. Kapitel 4.2). Dazu gehörten Wege

- mit zu kurzer Aktivitätendauer (kürzer als 2 Minuten),
- mit zu kurzer oder unrealistisch langer Reisezeit (weniger als 30 Sekunden oder mehr als 1 Stunde),
- mit zu großen zurückgelegten Entfernungen, die innerhalb der Kleinstadt Borlänge gar nicht zurückgelegt werden können (weiter als 25 km),
- die mit sehr großer Sicherheit zu einem Ziel außerhalb des Beobachtungsgebiets führten, weil sie auf einer der vier großen Ausfallstraßen Borlänges enden (vgl. Kapitel 4.3.1),
- mit einer durchschnittlichen Geschwindigkeit von über 50 km/h, da diese auch nicht innerhalb von Borlänge zurückgelegt worden sein können.

Für den in dieser Arbeit realisierten Analyseschritt (der Zuweisung von wahrscheinlichen Wegzwecken) ist dieses Vorgehen angebracht, um nur Stopp-Positionen Wegzwecke zuzuweisen, bei denen davon ausgegangen werden kann, dass es sich tatsächlich um Wegziele handelt. Ein Problem entsteht demzufolge aber bei der weiteren verkehrswissenschaftlichen Auswertung: da die fehlerhaft scheinenden Wegziele aussortiert werden, entstehen Lücken bei der Analyse von Wegeketten und Tagesabläufen. In einer weiterführenden Arbeit wird bereits geprüft, inwiefern die hier aussortierten Wege nicht gelöscht, sondern korrigiert werden können. Das bedeutet z.B., dass Stopps mit zu kurzen Aktivitätendauern anstatt gelöscht den ihnen nachfolgenden Wegen hinzugefügt werden. Die Fahrzeiten und zurückgelegten Strecken der nachfolgenden Fahrten verlängern sich entsprechend (Axhausen, Schönfelder, Wolf, Oliveira und Samaga, 2003).

Im Makro `wochenende.aml` werden die Stopps nach ihrer Zugehörigkeit zum Wochentag in drei Klassen eingeteilt: `wochentags`, `samstags` oder `sonntags`. Hintergrund dieser Einteilung

ist, dass an Wochentagen tendenziell andere Aktivitäten durchgeführt werden als am Wochenende und an Samstagen wiederum andere als an Sonntagen⁹. Dies ergibt sich aus den Rahmenbedingungen wie Ladenöffnungszeiten, Arbeitszeiten, Möglichkeiten der Freizeitgestaltung etc. Diese Tatsache impliziert, dass an ein und demselben Ort wochentags, samstags und sonntags jeweils verschiedene Tätigkeiten ausgeübt werden können. Deshalb sollten die Stopp-Cluster nach Wochentagskategorien getrennt behandelt und ihnen vor allem unabhängig voneinander mögliche Wegzwecke zugeordnet werden (Schwesig, 1998).

Im Folgenden und auch in den Coveragebezeichnungen steht ‚w‘ für wochentags, ‚sa‘ für samstags und ‚so‘ für sonntags. Sämtliche in den nächsten Abschnitten beschriebene Arbeitsschritte laufen also für die genannten drei Wochentagskategorien getrennt ab.

5.3 Clusterbildung – Stopp-Positionen

Voraussetzungen

Einige der von Verkehrsteilnehmern aufgesuchten Orte werden in mehr oder weniger regelmäßigen Abständen immer wieder aufgesucht (vgl. Abbildung 21). Betrachtet man, wie in dieser Arbeit, nur mit dem Automobil aufgesuchte Orte, die zudem mit GPS-Technik aufgenommen wurden, dann folgt daraus, dass sich identische Fahrtziele nicht zwangsläufig in identischen Stopp-Positionen abbilden müssen. Gründe dafür sind, dass gar nicht jedes Mal der identische Parkplatz genutzt werden konnte, eine Ungenauigkeit innerhalb der GPS-Daten vorhanden ist und außerdem bei diesen Daten nur die aufs Straßennetz gesnapten Positionen bekannt sind. Deshalb sollten Stopp-Positionen zusammengefasst werden, bei denen davon ausgegangen werden kann, dass der Fahrer identische Ziele, jedoch nicht den gleichen Parkplatz aufgesucht hat.

⁹ In einem kleinen fensterlosen Hinterzimmer des Ladens öffnete er eine Pappschachtel und zog zwölf völlig gleich aussehende schwarze Fotoalben daraus hervor. [...] In den letzten zwölf Jahren habe er jeden Morgen um Punkt 7 Uhr an der Ecke Atlantic Avenue und Clinton Street gestanden und jeweils aus genau demselben Blickwinkel ein Farbfoto aufgenommen. [...] Schließlich vermochte ich subtile Unterschiede im Verkehrsfluss zu erkennen, den Rhythmus der einzelnen Tage vorauszuahnen (das Gewühl an Werktagen, die relative Ruhe der Wochenenden, den Kontrast zwischen Samstagen und Sonntagen). [...] (aus: Paul Auster (1990) *Auggie Wren's Christmas Story*, S.2f; Deutsch von Werner Schmitz (1991) *Auggie Wrens Weihnachtsgeschichte*, Rowohlt Verlag GmbH, Reinbek)

Um diese Aufgabe der Clusterung¹⁰ umzusetzen, standen prinzipiell zwei verschiedene Verfahren zur Verfügung. Die Statistiksoftware SAS bietet mit der Funktion *FASTCLUS* eine Möglichkeit, beliebige numerische Daten zu clustern, d.h., es werden mathematisch über die Methode der kleinsten Quadrate Schwerpunkte in Clustern bestimmt (SAS Institute Inc., 1994). In der GIS-Software ArcInfo können derartige Funktionen mithilfe der Makrosprache AML für diesen speziellen Fall entwickelt werden. Letztere Verfahrensweise bietet die beiden Vorteile, dass die Clusterung automatisch erfolgen kann, d.h., der Arbeitsaufwand bei der Anwendung beschränkt sich darauf, das Programm lediglich mit einer Zeile aufzurufen, und, was für diesen Fall aber noch wichtiger ist: der Prozess kann auf die speziellen Bedürfnisse des Anwendungsfalls angepasst werden, z.B. kann mit AML umgesetzt werden, dass Clusterzentren nur auf dem Straßennetz liegen dürfen (vgl. Abbildung 12). Zusätzlich ist es für den weiteren Arbeitsablauf günstig, bereits diesen Schritt in der Sprache zu automatisieren, in der auch die weiteren Teile programmiert werden, um letztlich den gesamten Analyseprozess in einem einheitlichen System automatisiert ablaufen lassen zu können.

Automatisiert man diesen Prozess als AML in ArcInfo, können also beliebige Details oder Zwischenschritte eingefügt werden. In diesem Beispiel betrifft dies vor allem den Umgang mit den Stopp-Positionen auf Parkplätzen. Sie könnten mit der *FASTCLUS*-Clusterung von SAS nur für jeden Parkplatz separat behandelt und bearbeitet werden, nachdem sie vorher durch eine geometrische Analyse im GIS getrennt und dann wieder in SAS eingelesen würden. Da dieser Schritt jedoch lediglich manuell durchgeführt werden könnte, wurde auf die rechentechnisch schnellere Variante von SAS verzichtet und der Vorteil der vollständigen Automatisierung genutzt.

¹⁰ Bei der Clusteranalyse werden einzelne Messwertkombinationen schrittweise so zu Gruppen zusammengefasst, dass sich diese »Cluster« einerseits gut voneinander unterscheiden lassen und andererseits eine im Sinne der jeweiligen Fragestellung befriedigende Reduktion der umfangreichen Ausgangsdaten ergeben. (Brockhaus AG, 2001)

Ablauf der Clusterung

Die Clusterung läuft in dem Makroteil `clus.aml` ab, der wiederum die beiden Unterprogramme `clus-pp.aml` und `cluster.aml` aufruft.

Im ersten Teil, `clus-pp.aml`, wird untersucht, welche Stopp-Positionen eines Autofahrers auf einem der bekannten Parkplätze liegen¹¹. Sie werden temporär in externen Coverages gespeichert, getrennt von den übrigen Positionen geclustert und nach der Clusterung wieder mit den anderen Stopp-Positionen in einer Coverage vereinigt (Abbildung 10).

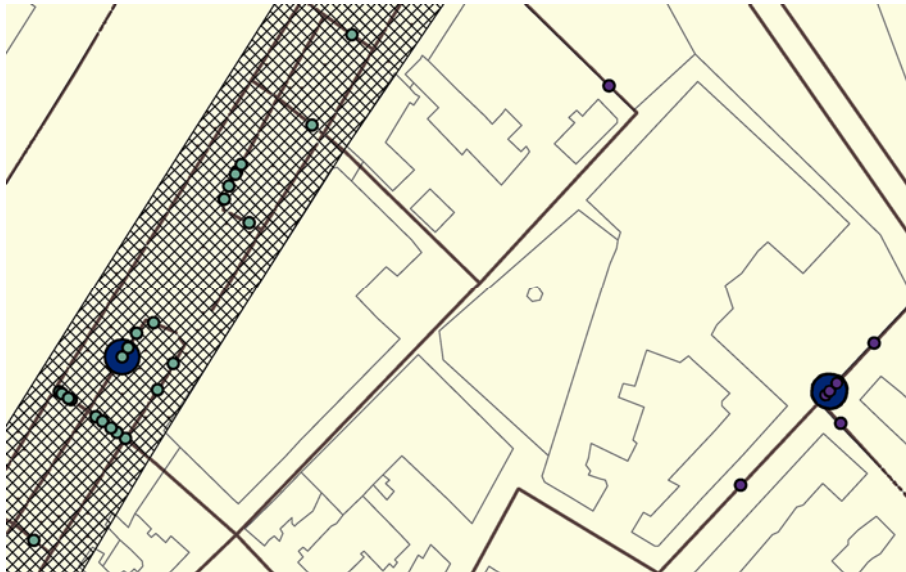
Die eigentliche Clusterung läuft in `cluster.aml` ab. Sie erfolgt in drei wesentlichen Arbeitsschritten. Zuerst werden die Entfernungen von jeder Stopp-Position zu jeder anderen in einem maximalen Abstand von 200 m zu einander berechnet¹². Diejenigen Stopp-Positionen mit den meisten Nachbarn in diesem Radius und der durchschnittlich kürzesten Entfernung zu diesen Nachbarpunkten werden als potentielle Clusterzentren markiert. In einem zweiten Schritt werden Stopps untersucht, die zwischen zwei Clusterzentren liegen. Diese werden dann dem nächstliegenden Zentrum zugeschlagen. Gab es viele solcher Stopp-Positionen, die neu zugeteilt werden mussten, werden in einem dritten Schritt die Lagen der betroffenen Clusterzentren neu berechnet (Abbildung 11).

Somit wird ein bestehender zentralerer Punkt als Schwerpunkt des Clusters bestimmt. Denn bei diesem Verfahren ist wichtig, dass keine geometrischen Mittelwerte berechnet und dadurch neue Punkte als Clusterschwerpunkte erzeugt werden, sondern jeweils bestehende Stopp-Positionen als Clusterzentrum benutzt werden. Es werden also Median-Positionen verwendet, sodass es nicht möglich ist, dass ein Clusterzentrum außerhalb des Straßennetzes liegt (vgl. Abbildung 12).

¹¹ zur Digitalisierung der Parkplätze siehe Kapitel 4.3.2

¹² In einer Kleinstadt wie Borlänge ist es erfahrungsgemäß nicht zu erwarten, dass sehr lange Fußwege in Kauf genommen werden, um einen Aktivitätenort vom Parkplatz des Autos aus zu erreichen. Aus diesem Grund ist die maximale Entfernung von 200 m festgelegt worden, die aber natürlich durch einen anderen Variablenwert verändert werden könnte.

Abbildung 10 Einbeziehung der Parkplätze in die Clusterung



Diese Abbildung zeigt das Ergebnis der Clusterbildung unter Einbeziehung der Parkplatzflächen. Die schraffierte Fläche markiert einen Parkplatz, die hellgrauen Linien stellen Gebäude-mauern dar, die braunen Striche die Straßen. Gleich gefärbte Punkte kennzeichnen die Zugehörigkeit zu einem Cluster. Die großen dunkelblauen Kreise kennzeichnen die Clusterzentren. Es ist zu erkennen, dass Punkte auf einem Parkplatz zu einem Cluster und Stopp-Positionen außerhalb des Parkplatzes zu einem anderen Cluster zählen.

Abbildung 11 Clusterbildungsprozess – Ablauf

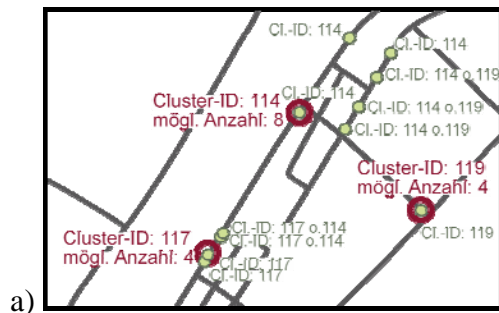
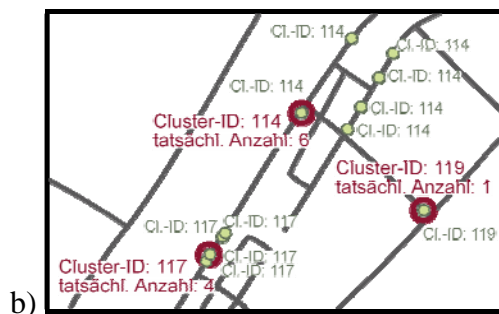
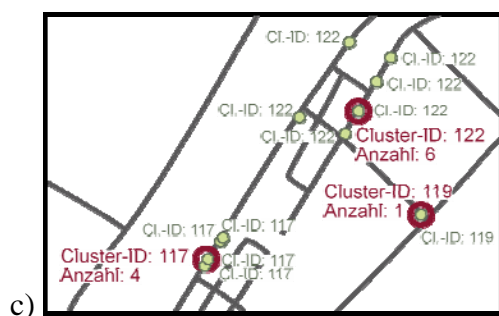


Abbildung a) zeigt das Ergebnis des ersten Arbeitsschrittes zur Bestimmung der Clusterzentren. Die rot umkreisten Stopp-Positionen sind diejenigen mit den meisten Nachbarn (mögliche Anzahl) im Umkreis von 200 m und gleichzeitig den durchschnittlich kürzesten Entfernungen zu diesen. Die Cluster-ID entsteht aus der ArcInfo-internen Nummerierung aller Punkte einer Coverage.

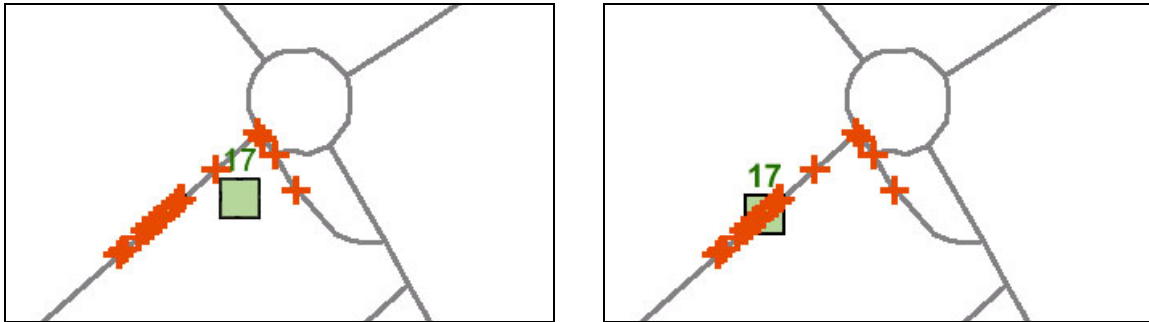


Das Resultat des nächsten Arbeitsschrittes ist in Abbildung b) dargestellt. Die Stopp-Positionen, die zwischen zwei Clusterzentren lagen und zu beiden weniger als 200 m entfernt liegen, wurden jeweils dem ihnen nächsten Clusterzentrum zugeschlagen.



Im dritten Arbeitsschritt, dessen Ergebnis in Abbildung c) zu sehen ist, wurde für diejenigen Cluster ein neues Zentrum ermittelt, bei denen der Unterschied zwischen möglicher Anzahl und tatsächlicher Anzahl zugehöriger Stopp-Positionen größer als 5 % war. Eine Veränderung ergab sich dadurch für das oberste der drei in diesem Ausschnitt zu sehenden Cluster.

Abbildung 12 Clusterungsprozess – Medianpositionen der Zentren



Das linke Bild zeigt die mögliche Position des geometrischen Mittels des Clusterschwerpunktes. An dieser Stelle kann das Fahrzeug allerdings nicht tatsächlich gehalten haben. Die rechte Skizze zeigt das durch den Algorithmus berechnete Clusterzentrum - also die Medianposition, die einer bereits existierenden Stopp-Position entspricht und somit zwangsläufig auf dem Straßennetz liegt.

Der Algorithmus könnte noch verbessert werden, wenn die Stopp-Positionen nicht über ihre Luftliniendistanz zueinander geclustert würden, sondern entlang des Straßennetzes. Das wäre deshalb sinnvoll, weil sich benachbarte Parkplätze nicht durch eine kurze Entfernung zueinander auszeichnen, sondern durch kurze mit dem Fahrzeug zurückzulegende Strecken. Bei der Umsetzung dieser Aufgabe könnten die in ArcInfo implementierten Network-Funktionen genutzt werden. Für die Borlänge-Daten eignet sich das Verfahren allerdings nicht, da, wie bereits oben erwähnt, nicht alle befahrbaren Straßen im digitalen Netzwerk vorhanden sind. Ein weiteres Argument spricht gegen die Clusterung entsprechend der Kantenlängen. Es gibt Fälle, in denen Parkplätze mit kurzer Luftliniendistanz für denselben Wegzweck genutzt werden. Das passiert z.B. wenn ein Gebäude zwischen zwei in etwa parallelen Straßen liegt, von denen jeweils ein Fußweg zu dem Gebäude hinführt (vgl. Abbildung 13).

Abbildung 13 Spezialfälle bei der Clusterung durch zu schlechte Information über das Straßennetz

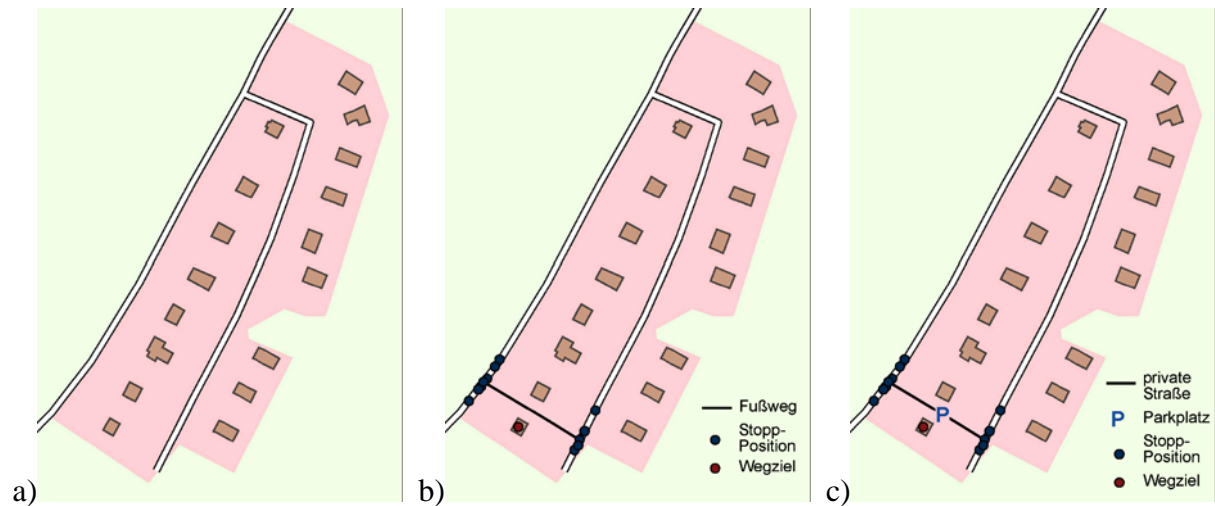


Abbildung a) zeigt die derzeit im digitalen Straßennetz und Flächennutzungsplan erfasste Situation. Die beiden anderen Abbildungen b) und c) zeigen mögliche Szenarien, die sich hinter der momentan dargestellten Situation verbergen könnten.

5.4 Clusteranalyse – wahrscheinliche Wegzwecke

Wie in Kapitel 2.3 bereits erläutert, wurden in bisherigen Verkehrsstudien die von den Verkehrsteilnehmern zurückgelegten Wege und die am Wegziel verfolgten Aktivitäten von den Teilnehmern selbst protokolliert. Auch in anderen GPS-gestützten Verkehrsbeobachtungen wurde die an einem Wegziel durchgeführte Aktivität von den Teilnehmern angegeben. Um Mobilität im Raum zu untersuchen ist diese Information essentiell. In der vorliegenden Studie von Borlänge fehlt sie allerdings. Deshalb müssen diese Wegzwecke, soweit es möglich ist, mithilfe des AML-Makropakets rekonstruiert werden. Die hier ermittelten wahrscheinlichen Aktivitäten sind also als Vermutung zu betrachten, die sich jeweils aus der Kombination der Position des Wegziels, der Ankunftszeit und der Dauer der Aktivität ergeben.

Ein solches Verfahren wurde bisher noch nicht automatisiert angewendet¹³. So lagen bei der Ermittlung sämtlicher numerischer Werte in Look-Up-Tabellen u.ä. keine Referenzwerte vor. Sie wurden speziell für diese Arbeit eingeführt und sind als erste Annäherung zu verstehen. Sie können im Programm jedoch jeder Zeit verändert werden.

Um die Vergleichbarkeit mit anderen Verkehrsstudien wie z.B. *Mobidrive* (PTV AG, 2002) oder *KONTIV* (Infas und DIW, 2001 und 2003) zu gewährleisten, wurde deren Kodierung der Aktivitäten übernommen. Dabei werden die von den Verkehrsteilnehmern Wege nach ihrem Zweck in folgende 10 Kategorien eingeteilt:

- zu Hause
- Jemanden Abholen/ Wegbringen
- Private Erledigung
- Geschäftliche Erledigung
- Ausbildung/ Schule
- Arbeiten
- Einkauf, täglicher Bedarf
- Einkauf, langfristiger Bedarf
- Freizeit
- Sonstiges

Es ist nicht ganz einfach, diese Kategorien auch für die hier angewendete andere Vorgehensweise zu benutzen, da die Einteilung nicht vom Verkehrsteilnehmer selbst vorgenommen wird. Berichtet ein Teilnehmer selbst über seinen Reisezweck, findet er leicht eine passende Kategorie (notfalls ‚Sonstiges‘). Bei der Rekonstruktion dieser Aktivität, bei der nur Informationen über den Ort und die Zeit vorliegen, kommt es aber zu Differenzierungsproblemen, wenn die gleichen Kategorien verwendet werden. So unterscheiden sich z.B. ‚geschäftliche Erledigung‘ und ‚private Erledigung‘ meist weder vom Ort noch von der Zeit signifikant von einander. Trotzdem wurde die Einteilung in diese zehn Aktivitätenkategorien übernommen,

¹³ Wolf (2000) hat in ihrer Dissertation ebenfalls versucht, Wegzwecke aus GPS-Daten nachträglich zu ermitteln. Dabei ist sie aber im Gegensatz zu dieser Arbeit manuell vorgegangen, sodass sich keine Referenzwerte für die automatisierte Anwendung ableiten ließen.

um die Ergebnisse mit bisherigen Studien vergleichen und so auch ihre Plausibilität besser überprüfen zu können. Die Kategorie ‚Sonstiges‘ wurde allerdings in ‚Ohne Ergebnis‘ (‚no result‘) umbenannt. Das soll deutlich machen, dass in diesem Fall kein wahrscheinlicher Wegzweck zugewiesen werden konnte.

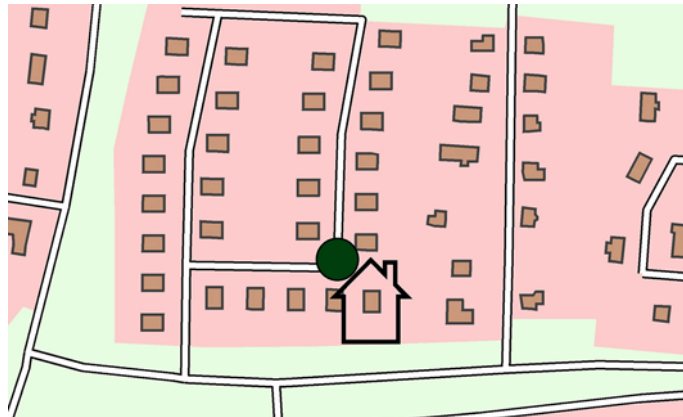
Die gesamte Clusterauswertung und Zuweisung der wahrscheinlichen Wegzwecke läuft im Makro `clus-analyse.aml` ab. Die folgenden Abschnitte beschreiben den Ablauf im Detail.

5.4.1 Einbeziehung der Flächennutzungsinformationen

Für die meisten in dieser Arbeit analysierten Autofahrer waren die Wohnadressen bekannt und durch das IVT digitalisiert worden¹⁴. So konnten in einem ersten Analyseschritt von vornherein die Stopp-Cluster identifiziert werden, an denen der jeweilige Teilnehmer zu Hause war. Sind diese Cluster im Analyseprozess einmal erkannt, kann ihnen kein anderer Aktivitätszweck mehr zugeordnet werden und es kann kein anderes Stopp-Cluster mit der Aktivität ‚zu Hause‘ belegt werden. Um die Wohnadresse festzulegen, wird wieder der 200 m Radius eingesetzt. Allen Clustern, die ihren Schwerpunkt in diesem Radius um das Wohnhaus haben, wird der Wegzweck ‚zu Hause‘ zugeordnet. In der Regel ist dies genau ein Cluster pro Wochentagskategorie und Autofahrer (vgl. Abbildung 14). In wenigen Ausnahmefällen gibt es zwei Cluster, beispielsweise wenn der Fahrer manchmal einen öffentlichen Parkplatz in der Nähe seines Hauses und manchmal einen Stellplatz in der Nebenstraße nutzt. Es ist zwar möglich, dass ein Fahrer in der Nähe seines Hauses auch einer anderen Aktivität nachgegangen ist, es ist jedoch sehr unwahrscheinlich, dass er dafür Strecken unter 200 m mit seinem Auto zurücklegt.

¹⁴ Zur Digitalisierung der Wohnadressen vgl. Kapitel 4.3.2 - Digitalisierte Wohn- und Aktivitätenadressen der Teilnehmer.

Abbildung 14 Wohnadresse eines Autofahrers



Die Hohlfigur in der Form eines Hauses liegt über dem Gebäude, das der Studienteilnehmer als Wohnadresse angegeben hat. Der grüne Punkt markiert einen Clusterschwerpunkt. Ihm wurde aufgrund der Nähe zu diesem Haus die Aktivität „zu Hause“ zugeordnet.

Nach diesem Auswertungsschritt wird der Flächennutzungsplan¹⁵ in die Analyse der Wegzwecke einbezogen. Dabei wird um den Clusterschwerpunkt ein 200 m Puffer gelegt. Die in diesem Puffer enthaltenen oder von ihm angeschnittenen Polygone des Flächennutzungsplans gehen als mögliche Wegziele ein. Um einem Cluster eine Aktivität zuzuordnen, werden Wahrscheinlichkeiten aus einer Look-Up-Tabelle (vgl. Anhang C 1: Einbeziehung des digitalen Flächennutzungsplans) entnommen. In dieser Tabelle sind für alle Flächennutzungsarten Wahrscheinlichkeitspunkte für jede der zehn oben erwähnten Aktivitätenkategorien aufgelistet.

Für die im Puffer vorhandenen Nutzungsarten gehen die Werte für die verschiedenen Aktivitätenkategorien aus der Look-Up-Tabelle dann jeweils *einmal* pro Art ein, d.h., wenn sich beispielsweise zehn Industriebauten im Umkreis des Stopp-Clusters befinden, gehen die Werte aus der Look-Up-Tabelle für Industriebauten nur *einmal* ein. Die Wahrscheinlichkeitspunkte für jeden Wegzweck werden dabei pro Cluster aufsummiert.

¹⁵ Zu den im Flächennutzungsplan enthaltenen Informationen vgl. Kapitel 4.3.2.

Die Flächennutzungspolygone hätten auch entsprechend ihrer Anzahl im Puffer oder entsprechend der von ihnen bedeckten Fläche im Puffer in die Kalkulation eingehen können. Darauf wurde aber verzichtet, da auf diese Weise Sondernutzungen wie Sporthallen und andere öffentliche Gebäude an Bedeutung verloren hätten. Sie bedecken keine großen Flächen und treten im Gegensatz zu anderen Nutzungen, z.B. Einfamilienhäuser, nur als einzelne Gebäude und nicht in Ansammlungen auf. Ihre Bedeutung für die Zuweisung des möglichen Wegzwecks ist aber groß – und verhältnismäßig wichtiger als flächen- oder zahlenmäßig stärker vertretene andere Nutzungsarten (siehe Abbildung 15).

Abbildung 15 Einbeziehung des Flächennutzungsplans mit seinen Sondernutzungen in die Clusteranalyse

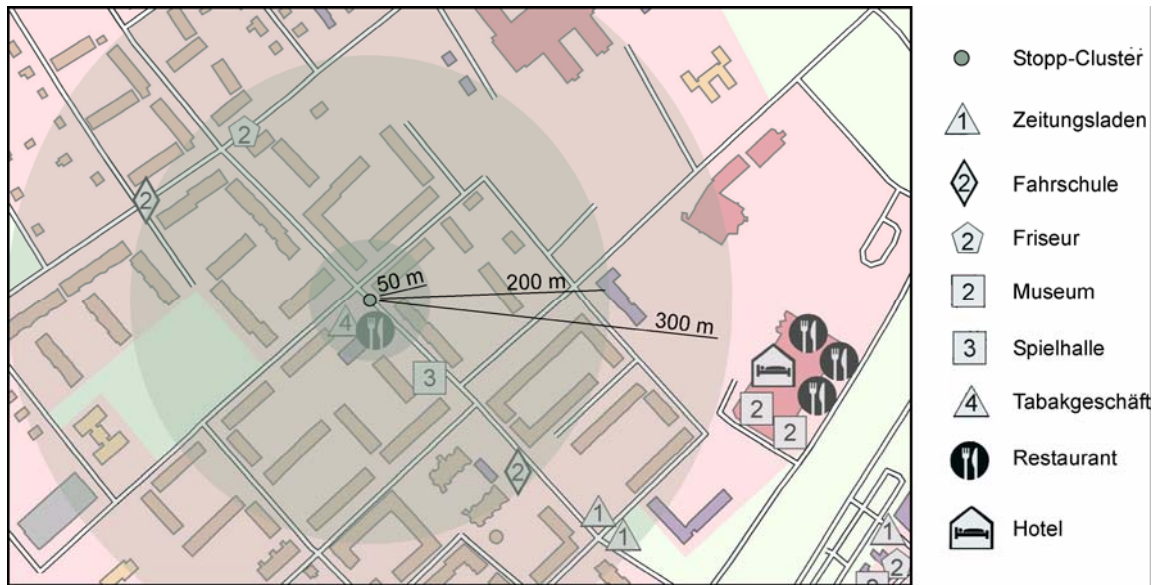


Der dunkelblaue Kreis markiert ein mögliches Stopp-Cluster. Würden die Polygone aus dem Flächennutzungsplan entsprechend ihrer Größe oder ihrer Anzahl im 200m-Puffer einbezogen, würde der Kindergarten als Sondernutzung kaum ins Gewicht fallen.

Auf ähnliche Weise gehen die Points of Interest¹⁶ ein. Auch für sie gibt es eine Look-Up-Tabelle, in der allen Arten von Points of Interest Wahrscheinlichkeitspunkte für den Aktivitätszweck zugeordnet werden (vgl. Anhang C 2: Einbeziehung verschiedener “Points of Interest”). Auch hier werden die Wahrscheinlichkeitspunkte für die verschiedenen Wegzwecke aufaddiert (und zwar zu denen aus der Analyse des Flächennutzungsplans dazu). Der Unterschied zur Flächennutzungsinformation besteht darin, dass hier Punkte in einem 300 m Umkreis gesucht werden und außerdem jeweils ein Faktor für die Entfernung zwischen Stopp-Cluster und Point of Interest eingerechnet wird. Dass die Testperson Orte in einem Umkreis von 50 m von ihrem Parkplatz aufgesucht hat, wurde als sehr wahrscheinlich eingestuft. Also werden die Wahrscheinlichkeitspunkte der Points of Interest in diesem Abstand mit einem Faktor von 1,5 multipliziert. Points of Interest, die weiter als 200 m vom Clusterschwerpunkt entfernt sind, gehen schließlich nur noch mit einem Faktor von 0,4 ein, da es unwahrscheinlich, aber nicht auszuschließen ist, dass ein Autofahrer in einer Kleinstadt wie Borlänge regelmäßig so weit zu seinem Fahrtziel zu Fuß geht.

¹⁶ zu Points of Interest vgl. Kapitel 4.3.2

Abbildung 16 Einbeziehung der ‚Points of Interest‘ in die Clusteranalyse



Die beiden ‚Points of Interest‘ im 50m-Umkreis (in diesem Fall Restaurant und Tabakgeschäft) gehen am stärksten in die Berechnung ein, denn die Wahrscheinlichkeit ist groß, dass der Fahrer einen dieser Orte aufgesucht hat.

Als Ergebnis dieses Analyseteils gibt es pro Clusterzentrum Punktverteilungen für wahrscheinliche Wegzwecke, wobei in der Regel ein Stopp-Cluster als Wohnort erkannt werden konnte. An dieser Stelle des Analysealgorithmus zeichnet sich das zum Wohnort gehörende Stopp-Cluster durch sehr viele Wahrscheinlichkeitspunkte für den Zweck ‚zu Hause‘ aus.

Eine Verbesserung der Analyse aus den Flächennutzungsinformationen ergäbe sich, würde nicht die Luftliniendistanz als Pufferradius und Entfernungsmaß, sondern die Gehentfernung (evtl. auch nicht als Strecke, sondern als Zeit) als Maß genutzt. Die Informationen dafür, d.h. das Fußwegnetz etc., liegen bisher allerdings nicht vor. Bevor die Entfernungsmessung im Algorithmus verbessert wird, müssten allerdings auch die Daten zu den Points of Interest vervollständigt werden (vgl. 4.3.2).

5.4.2 Einbeziehung zeitlicher Informationen

Neben dem wahrscheinlichen Wegzweck, der sich aus der das Clusterzentrum umgebenden Flächennutzung ergab, wird ein weiterer wahrscheinlicher Wegzweck ermittelt, der sich aus den zeitlichen Bedingungen ergibt. Dazu werden die Referenzdaten aus der Verkehrsbefragung RES herangezogen. Sie lagen nach der statistischen Aufbereitung als Look-Up-Tabelle vor¹⁷.

Der Analysealgorithmus verarbeitet diese Information dann folgendermaßen: Für jedes Stopp-Cluster werden die Mediane der Ankunftszeiten und der Aktivitätendauern ermittelt. Dann wird der RES-Look-Up-Tabelle der entsprechende Wert entnommen. Dabei werden neben den zeitlichen Kategorien folgende berücksichtigt:

- die Wochentage (,w', ,sa' oder ,so'),
- der Beschäftigungsstatus des Fahrers (bisher nur Rentner, Vollzeitbeschäftigte und Andere),
- das Geschlecht des Testfahrers und
- die Verfügbarkeit von Führerschein und Auto (bisher nur ,Führerschein *und* Auto verfügbar' implementiert, da nur Haushalte mit einem Autofahrer und einem Auto ausgewählt wurden und so andere Kategorien nicht in Frage kommen).

Der in der RES für diese Parameter am häufigsten genannte Wegzweck wird dann auch als wahrscheinlicher Wegzweck für die automatisierte Auswertung der Borlänge-Daten vermutet. Gibt es keinen passenden Wert in der RES-Look-Up-Tabelle, wird versucht, einen passenden mit um eine Stunde verschobenem Aktivitätenbeginn zu finden.

Die Qualität dieses Analyseabschnittes hängt sehr stark von der Übereinstimmung der Referenzwerte mit der tatsächlichen Studie ab. Vergleichsdaten von Autofahrern aus Borlänge wären von daher äußerst wünschenswert und würden das Ergebnis verbessern.

Eine andere Möglichkeit zur Verfeinerung der Aussage könnte dadurch erzielt werden, dass der zeitliche Vergleich mit der RES nicht über die Mediane des gesamten Stopp-Clusters erfolgt, sondern für jede Stopp-Position einzeln. Aus den sich daraus ergebenden wahrscheinli-

¹⁷ Zu Inhalt und statistischer Aufbereitung der Verkehrsbefragungsdaten vgl. Kapitel 4.3.3.

chen Wegzwecken pro Stopp-Position würde dann ein Modalwert als wahrscheinlichster Wegzweck des Stopp-Clusters bestimmt werden.

5.4.3 Verknüpfung zu wahrscheinlichen Wegzwecken

Als letzter Schritt zur Ermittlung der wahrscheinlichen Wegzwecke werden die in den beiden oben beschriebenen Verfahren (aus der Analyse der Flächennutzungsinformation und Analyse der zeitlichen Komponenten) ermittelten wahrscheinlichen Wegzwecke miteinander verglichen und zu einem Wegzweck verknüpft.

Dabei werden folgende Möglichkeiten in der hier angeführten Reihenfolge berücksichtigt und mit verschiedenen Wahrscheinlichkeitsparametern belegt, um die Aussagekraft dieser Vermutung zu verdeutlichen. Dabei steht eine 1 für die größte Sicherheit bei der Zuweisung und eine 3 für die unsicherste.

- Als erstes wird der Wegzweck ‚zu Hause‘ aus der Adressanalyse vergeben und mit der Sicherheit 1 belegt
- Konnte die Wohnadresse sicher an ein Clusterzentrum vergeben werden, werden alle anderen möglichen Wegzwecke ‚zu Hause‘ (vor allem die aus der zeitlichen Analyse) vom weiteren Verfahren ausgeschlossen.
- Danach werden für die Vollzeitbeschäftigten Arbeitsorte ermittelt und mit dem Wegzweck ‚Arbeiten‘ und der Sicherheit 1 belegt, wenn sie folgende Kriterien erfüllen:
 - das entsprechende Stopp-Cluster wird am zweithäufigsten von allen aufgesucht,
 - es gehört in die Wochentags-Kategorie (,w‘) und
 - aus dem zeitlichen Vergleich mit den RES-Daten ergab sich als zeitlich wahrscheinlicher Wegzweck ‚Arbeiten‘.
- Konnte ein solches Cluster bestimmt werden, wird ebenfalls der Wegzweck ‚Arbeiten‘ für die weitere Zuordnung ausgeschlossen.
- Stimmt der wahrscheinliche Wegzweck aus der Flächennutzungsanalyse und der zeitlichen Analyse überein, wurde eben dieser Zweck übernommen und ebenfalls die Sicherheit 1 vergeben.
- Stimmt die Ergebnisse aus den beiden Analysen nicht überein, wurde die Zuweisung aus der Flächennutzung übernommen, da davon ausgegangen wurde, dass der örtliche Bezug eine größere Wahrscheinlichkeit auf die richtige Zuordnung verspricht. Einzige Ausnahme war, wenn sich aus der zeitlichen Analyse ‚Jemanden Abholen/ Wegbringen‘ ergeben hatte. Dann wurde dieser Wegzweck übernommen, da diese Aktivitätenkategorie aus der Landnutzung kaum herausgelesen werden kann

und für sie hauptsächlich die kurze Aktivitätendauer charakteristisch ist. All diesen Zuordnungen wurde die Wahrscheinlichkeit 2 zugewiesen.

- Für alle restlichen Stopp-Cluster, für die kein Zweck aus der Landnutzung abgeleitet werden konnte, wurde der Zweck aus der zeitlichen Analyse übernommen, aber nur noch mit der Sicherheit 3 belegt.

Langfristig könnte das Verfahren über eine stärkere Betonung der verschiedenen Wahrscheinlichkeitswerte, z.B. durch die Implementierung einer Diskriminanzanalyse, verbessert werden. Bei der Diskriminanzanalyse wird eine zu erklärende, abhängige Variable durch eine Funktion definiert, die eine oder mehrere unabhängige, erklärende Variablen enthält. Dabei würden sich also in diesem Fall die Werte für die wahrscheinlichen Wegzwecke (abhängige Variable) aus einer Funktion der anderen Variablenwerte (Werte aus Flächennutzung und zeitlicher Analyse) ergeben. (Psychologie TU BS, 2003)

6. Ergebnisse

6.1 Visualisierung der Ergebnisse des Makropakets

Am Ende eines Datenverarbeitungsprozesses in einem GIS steht sinnvollerweise eine Visualisierung der Ergebnisse. Doch vor der Visualisierung müssen stets folgende Fragen geklärt werden: ‚Was soll dargestellt werden?‘, ‚Auf welchem Medium soll die Ausgabe erfolgen?‘ und ‚Wie kann dies kartographisch umgesetzt werden?‘¹⁸

Was soll dargestellt werden?

Das eigentliche Thema der resultierenden Karte(n) ist natürlich das Ergebnis der im vorangehenden Kapitel beschriebenen Analyse. Und das AML-Makropaket zur Rekonstruktion wahrscheinlicher Wegzwecke liefert prinzipiell vier Resultate zu jedem Stopp-Cluster. Dies sind

- die Wochentagskategorie,
- die Anzahl der Stopp-Positionen, die in diesem Cluster vereinigt wurden,
- der wahrscheinlichste Wegzweck (eingeteilt in die o.g. Kategorien) und
- die Sicherheit (in drei Stufen), mit der dieser Wegzweck als wahr angenommen werden kann.

Neben der eigentlichen Thematik soll thematischen Karten die inhaltlich entsprechend reduzierte topographische Grundlage zugrunde liegen (Arnberger, 1993). Im Fall der Borlänge-Daten sollten allerdings statt einer eigens für die Visualisierung angefertigten Grundlage lieber die in den Analyseprozess eingeflossenen Daten dargestellt werden. Sie liefern den topographischen Bezug ebenso und ermöglichen zudem, dass bei der Auswertung der Karten der Analyseprozess zumindest im Groben nachvollzogen werden kann. Die darstellbaren eingeflossenen Daten sind der digitalisierte Flächennutzungsplan, die ‚Points of Interest‘ sowie (zumindest für die bisher analysierten Testfahrer) die Wohnadressen. Außerdem soll natürlich das Straßennetz dargestellt werden, auf das sich die gesamte Thematik bezieht.

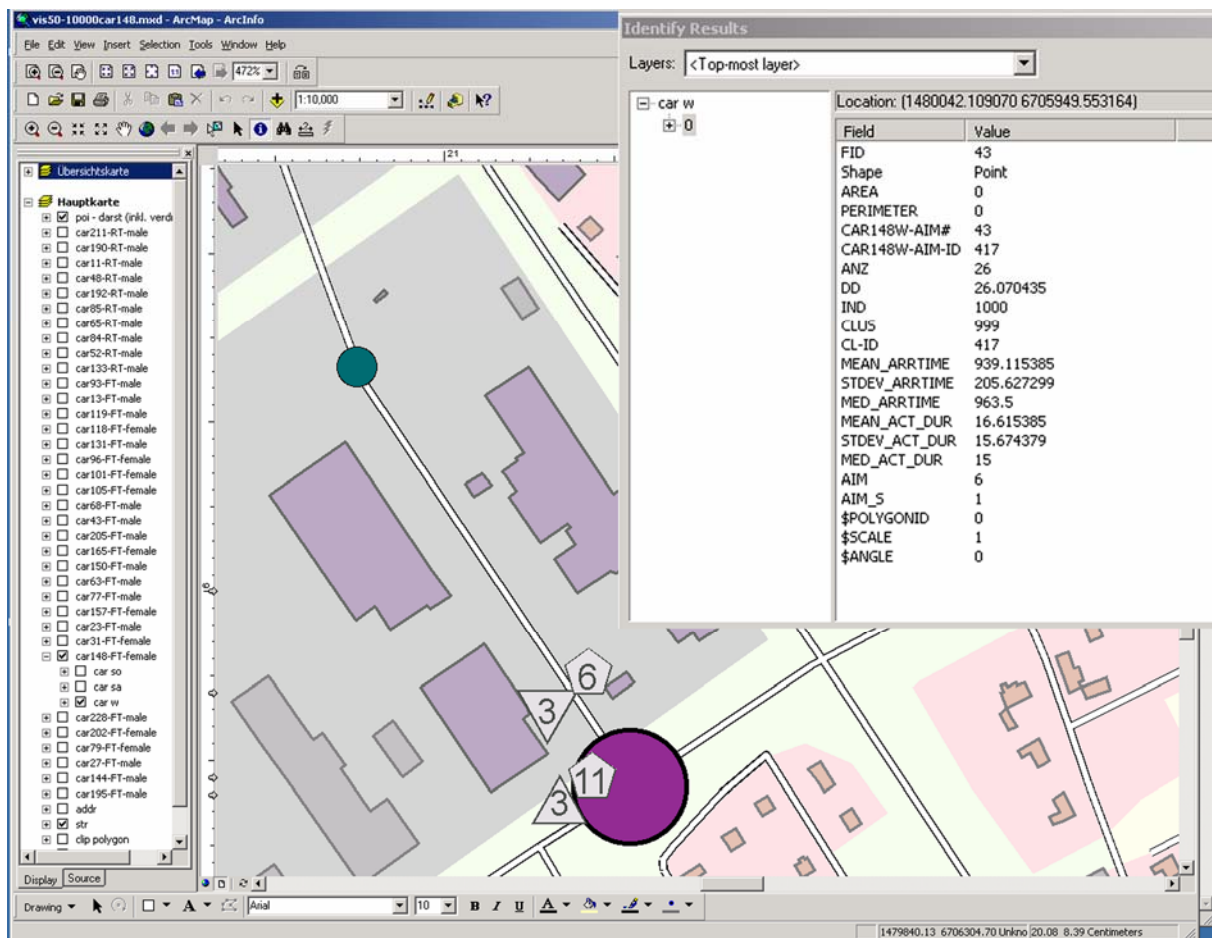
¹⁸ Die Visualisierungsbeispiele befinden sich im Anhang E.

Auf welchem Medium soll die Ausgabe erfolgen ?

Für im GIS erstellte Karten bietet sich prinzipiell sogar die Ausgabe am Bildschirm an. Sie bietet vor allem den Vorteil, dass weitere Informationen zu den dargestellten Objekten per Mausklick abgerufen werden können.

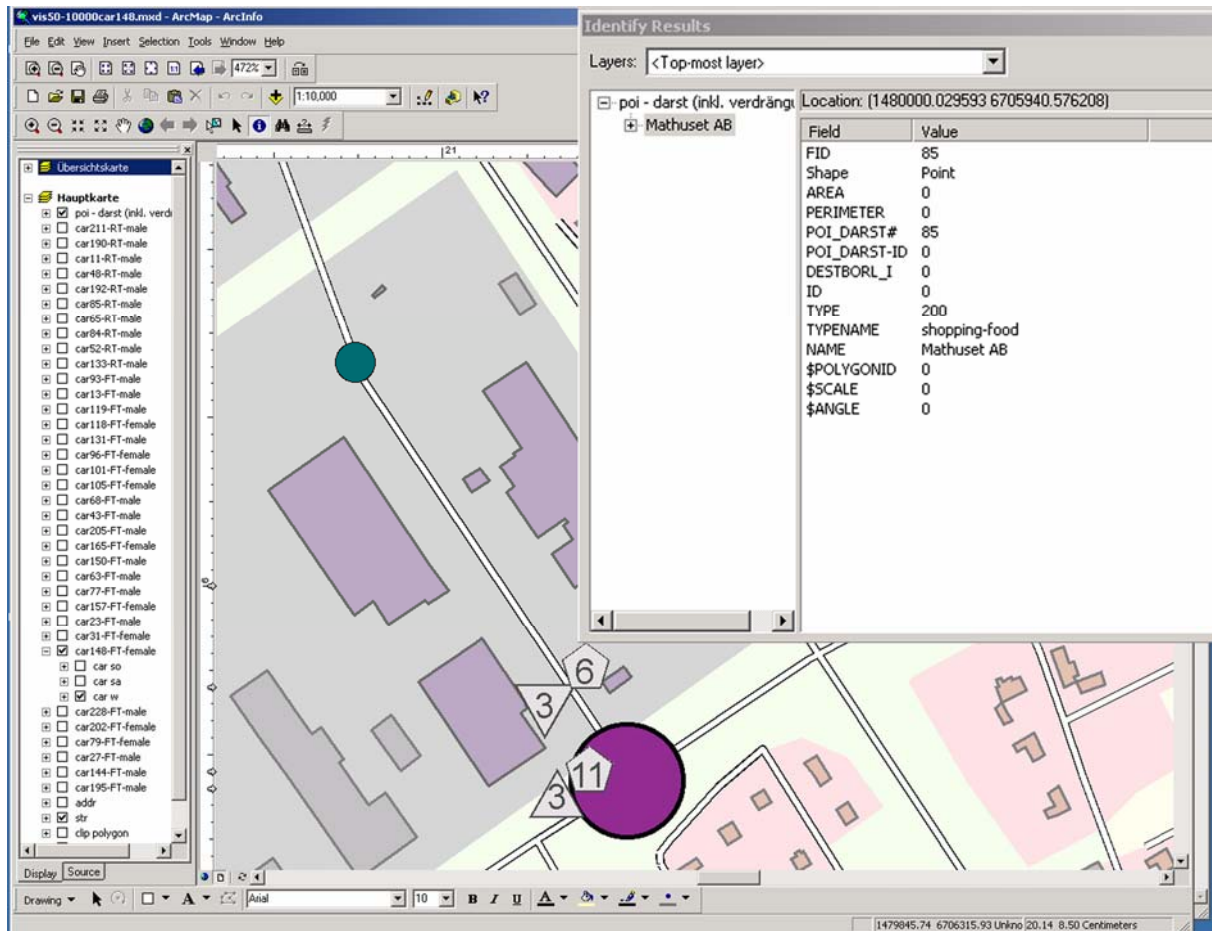
Die Ausgabe auf Papier hat hingegen den Vorteil, hard- und softwareunabhängig verwendbar zu sein. Ergebnisse können so an verschiedenen Orten verschiedenen Personenkreisen zugänglich gemacht werden. Durch die gedruckte Version geht allerdings die Anbindung an weitere Daten verloren und für Aktualisierungen werden jeweils neue Ausgaben benötigt.

So sind die im Anhang E enthaltenen Karten als Druckversionen von Bildschirmkarten zu sehen, bei denen am Bildschirm stets noch weitere Informationen abgerufen werden können (vgl. Abbildung 17 und 18) und die typischen ‚GIS-Karten-Probleme‘ wie sich überlappende Signaturen durch Funktionen wie Zoomen oder Selektieren und Abrufen von Zusatzinformationen gelöst werden können. Das Format von DIN A3 für die Hardcopy-Version wurde vor allem aus technischen Gründen gewählt. Es bietet ausreichend Platz für aussagekräftige Kartenausschnitte und die umfangreiche Legende, was in DIN A4 kaum mehr möglich gewesen wäre, und es ist für die Bindung in einer DIN A4 Arbeit noch praktikabel.

Abbildung 17 Abrufen von Informationen zu Clusterzentren ¹⁹

Diese Abbildung zeigt das Informationsmenü zum unteren der beiden Clusterzentren. ‚ANZ‘ gibt an, wie viele Wegziele in diesem Cluster zusammengefasst wurden. ‚DD‘ steht für die durchschnittliche Entfernung des Clusterzentrums zu den integrierten Wegzielen. ‚MEAN...‘, ‚STDEV...‘ und ‚MED...‘ geben jeweils den Durchschnittswert, die Standardabweichung bzw. den Median der Ankunftszeit (‚...ARRTIME‘) bzw. Aktivitätendauer (‚...ACT_DUR‘) an.

¹⁹ Zur Erläuterung der graphischen Elemente vgl. Karten in Anhang E.

Abbildung 18 Abrufen von Informationen zu den ‚Points of Interest‘²⁰

Diese Abbildung zeigt das Informationsmenü zu einem der ‚Points of Interest‘. So kann die Kategorie („TYPENAME“) und der Name des Geschäfts etc. („NAME“) abgerufen werden.

Wie kann dies kartographisch umgesetzt werden?

Für die kartographische Umsetzung konnten entweder bestehende Visualisierungsmethoden für Mobilitätsdaten oder eine neue Variante angewendet werden. Da es vordergründig um die Darstellung der Ergebnisse der automatisierten Auswertung und nicht der reinen Mobilitäts-

²⁰ Zur Erläuterung der graphischen Elemente vgl. Karten in Anhang E.

daten ging, bot es sich an, eine neue Visualisierungsvariante speziell für diese Thematik zu entwickeln.

Vor der eigentlichen Gestaltung der Karte(n) musste die Frage des Maßstabs geklärt werden. Die Ausgabe sollte in DIN A3 erfolgen (s.o.). So passen im Maßstab 1:50'000 jeweils Übersichtskarten des gesamten Untersuchungsgebiets und in 1:10'000 verschiedene Detailkarten auf je ein Blatt. Die graphischen Vorschriften sollen allerdings einheitlich festgelegt werden, um bei den interaktiven Bildschirmkarten die Möglichkeit des Zooms zur Verfügung zu haben.

Die größte Schwierigkeit stellt die Fülle an Informationen dar. Um das Geflecht überschauen zu können, wurde deshalb in zwei Arbeitsschritten vorgegangen. Als erstes wurde die Grundlagenkarte bearbeitet und so gestaltet, dass sie graphisch in den Hintergrund tritt. In einem zweiten Schritt wurde die Thematik selbst, das Analyseergebnis, graphisch umgesetzt und so dominant gestaltet, dass sie über der Grundlagenkarte zu liegen scheint und so gut wie möglich gelesen werden kann.

Die Grundlagenkarte setzt sich aus den drei Inhalten, dem digitalen Flächennutzungsplan, den ‚Points of Interest‘ und dem Straßennetz zusammen.

- *Flächennutzungsplan*

Der Flächennutzungsplan sollte in dezenten Farben gestaltet werden. Eine farbige Flächenfärbung war deshalb nötig, weil es sich um zu viele Klassen handelte, als dass sie nur durch Grauwerte hätten unterschieden werden können. Auch die Unterscheidung durch verschiedene Muster war ungünstig, weil die gesamte Karte durch die große Informationsdichte sonst insgesamt zu unruhig geworden wäre. Für einige außerhalb der Innenstadt liegende Flächen wurden allerdings zusätzlich zu den Farben Muster eingesetzt, um eine weitere Differenzierung zu ermöglichen ohne zu viele nicht mehr unterscheidbare oder zu kräftige Farben einzusetzen (z.B. für Waldflächen wurde eine Baumsignatur als Flächenmuster verwendet). Des Weiteren wurden einige im Analyseprozess zwar in getrennten Klassen behandelte Flächennutzungsarten bei der Visualisierung zusammengefasst (z.B. Sporthallen zu öffentlichen Gebäuden), da sich bei zu großer Informationsfülle die Klassenunterschiede nur schwer erkennen lassen und die gesamte Karte bei einer anderen Vorgehensweise überladen würde. Deshalb wurden auch spezielle Flächen in den 1:10'000er Karten beschriftet, statt visuell von den anderen Klassen unterschieden zu werden (z.B. die Kläranlage, die visuell den Industrieflächen zugeschlagen wurde). Insgesamt blieben immer noch 15 Nutzungsarten übrig.

- *‚Points of Interest‘*

Das Problem bei der Visualisierung der ‚Points of Interest‘ war ebenso die große Anzahl an Kategorien – ca. 60. Um den Analyseprozess bei der Auswertung noch annähernd nachvollziehen zu können, sollten die Kategorien bei der Betrachtung der Kar-

ten aber klar werden. Deshalb wurden nur wenige Kategorien zusammengefasst. Es blieben 54 Nutzungsarten. Es können aber unmöglich derart viele Signaturen benutzt und dann in der Karte eindeutig wiedererkannt werden – vor allem, wenn ihre Gestaltung so dezent sein soll, dass sie vom eigentlichen Thema noch überlagert werden kann. Also wurden nur für einige wenige charakteristische Nutzungsarten bildhafte Signaturen ausgewählt (z.B. Kirche und Flughafen), deren Positionen auch Orientierungswert haben. Alle anderen Kategorien konnten in Gruppen gemäß ihrer Bedeutung für mögliche Aktivitäten an diesen Orten zusammengefasst werden. Und für jede dieser Gruppen wurde dann eine einfache geometrische Figur als Basis gewählt und mit einer Zahl für die jeweilige Kategorie versehen. Es wurde bewusst darauf verzichtet, Farben für die Differenzierung zu verwenden, da die graphischen Variablen ‚Farbe‘ und ‚Helligkeit‘ noch für das eigentliche Thema der Stopp-Cluster verwendet werden sollte. Ebenso wurde darauf verzichtet, Kombinationen aus verschiedenen geometrischen Figuren zu verwenden, da dadurch schnell Quantitäten oder Qualitäten suggeriert werden, welche aus den Daten gar nicht hervorgehen. Um das eigentliche Thema auch noch angemessen in den Vordergrund treten lassen zu können, wurden die ‚Points of Interest‘ nur in Grautönen dargestellt. Außerdem war die Datenbasis geometrisch nicht unverändert darzustellen, da teilweise bis zu sechs ‚Points of Interest‘ an ein und demselben Punkt lagen. So wurde eine extra ‚Point-of-Interest-Coverage‘ für die Darstellung erstellt, in der die Punkte gegeneinander verdrängt wurden.

- *Straßennetz*

In ArcMap ist die Funktion integriert, bestimmte Themen je nach Maßstab an- und auszuschalten und so das Erscheinungsbild zu verändern. So konnten die Straßen im Maßstab 1:10'000 als ästhetisch etwas ansprechendere Doppellinien und in 1:50'000 als einfache dunkelgraue Linien dargestellt werden, die bei dem begrenzten Platz angemessener erschienen, um das Kartenbild nicht zu überlasten.

In dem Abschnitt ‚Was soll dargestellt werden?‘ wurden bereits die vier Aussagen über die Stopp-Cluster aufgezählt, die als die eigentliche thematische Ebene visualisiert werden sollten. Im Folgenden soll auf diese vier Parameter pro Stopp-Cluster eingegangen und die jeweilige Umsetzung in eine graphische Variable vorgestellt werden.

- *Anzahl der im Cluster vereinten Stopp-Positionen*

Die Eigenschaft, deren graphische Umsetzung am einfachsten und eindeutigsten ist, dürfte die Anzahl der im Cluster vertretenen Stopp-Positionen sein. Quantitäten können immer am günstigsten durch die Variable ‚Größe‘ dargestellt werden. Am günstigsten ist dabei ein kontinuierlicher flächenproportionaler Signaturenmaßstab, wenn die Werte (und die Software) dies zulassen. Probleme entstehen, falls es zu große und/oder zu kleine Werte für eine einheitliche Skala gibt. Die hier vorliegenden Werte lassen einen flächenproportionalen Wertemaßstab bedingt zu. D.h. für sehr mobile Autofahrer werden die Signaturen sehr groß und überdecken sehr viel. Da die Software ArcMap kontinuierliche Skalen auch nur ermöglicht, wenn eine einfache geometrische Signatur verwendet wird, die nicht noch durch weitere Parameter und somit weitere graphische Elemente beeinflusst werden soll, konnte eine streng propor-

tional kontinuierliche Skala sowieso nicht umgesetzt werden. Also wurde eine *streng proportional gestufte* Skala angewendet. (vgl. Arnberger, 1993, S.83ff)

- *Wochentagskategorien*

Die drei Wochentagskategorien sollten graphisch das gleiche Gewicht haben, weil sie lediglich qualitative Unterschiede repräsentieren und es deshalb auf keinen Fall zu einer Gewichtung kommen darf. Möglich zur Unterscheidung von Qualitäten sind die drei graphischen Variablen Form, Richtung und Farbe (vgl. Bollmann und Koch, 2001, S.350f). Die Variable Farbe sollte noch offen bleiben für die Wegzwecke (s.u.). Die Variation der Richtung wäre eine Möglichkeit gewesen. Es treten dabei aber zwei Probleme auf: wählt man eine symmetrische Ausgangsfigur, lassen sich die resultierenden Figuren im Kartenbild nur schlecht unterscheiden, sie heben sich schlecht von einander ab. Benutzt man unsymmetrische Figuren, wirken die Richtungen schnell wertend (aufsteigend, absteigend etc.). So erwies sich die Variation der Variable ‚Form‘ als am günstigsten.

- *wahrscheinlichste Wegzwecke und die Sicherheit ihrer Zuordnung*

Wie oben bereits angedeutet, wurden für die zehn Kategorien von Wegzwecken Farben verwendet. Sie repräsentieren ebenfalls qualitative Unterschiede, sodass sich die Variable Farbe anbot. Der Vorteil dieser Variablen besteht darin, dass sie gegenüber dem übrigen graphischen Gefüge durch eine kräftige Gestaltung dominant hervortreten kann. Das ist wichtig, weil die Information über die angenommenen Wegzwecke das eigentliche Analyseergebnis enthalten. Außerdem konnten die einzelnen Farbtöne in ihrer Helligkeit variiert werden und somit die Sicherheiten der Zuordnung in direkter logischer und graphischer Verbindung mit den Wegzwecken selbst dargestellt werden. Die Anzahl der hierfür insgesamt verwendeten Farben ist zwar recht groß (27 Farbtöne und 1 Grauwert für nicht zugeordnete Cluster), das System funktioniert aber insofern, als die zehn Wegzweck-Kategorien gut von einander unterschieden und stets eindeutig zugeordnet werden können – diese Aussage ist die wichtigste. Die Zuordnung der Sicherheitskategorien ist hingegen optisch komplizierter, aber inhaltlich auch weniger wichtig.

In weiterführenden Arbeiten wäre es interessant, weitere Auswertungen der Ergebnisse darzustellen wie die Regelmäßigkeit von Aktivitäten oder abgeleitete Aktivitätenräume (vgl. Schönfelder und Samaga, 2003). Als Weiterführung wären auch perspektivische Darstellungen möglich, die Mobilität als dritte Dimension umsetzen. Die Funktionen, die der ArcView-3D-Analyst Nachfolger ArcGIS-ArcScene bietet, sind derzeit noch so beschränkt, dass nur sehr grobe Eindrücke vermittelt werden können. Zum Beispiel werden bei dem Versuch, oben beschriebene kartographische zweidimensionale Umsetzung perspektivisch darzustellen, sämtliche Farbeinstellungen durch die Schattenberechnung so stark verändert, dass sie nicht mehr voneinander zu unterscheiden sind.

Die Sprache für die Schrift in den Karten sollte Englisch sein. Dadurch kann davon ausgegangen werden, dass die Karten sowohl in Schweden als auch im deutschsprachigen Raum ver-

standen werden können. So können die Karten auch den Datenlieferanten als Ergebnis vorgelegt werden.

6.2 Qualität der Ergebnisse aus verkehrswissenschaftlicher Sicht

Für die 39 ausgewählten Testfahrer²¹ wurden bereits erste Plausibilitätsprüfungen aus verkehrswissenschaftlicher Sicht am IVT durchgeführt und in Schönfelder und Samaga (2003) veröffentlicht. Dabei wurden die Ergebnisse des Algorithmus, d.h. die den Wegen zugeordneten Wegzwecke, in ihren Wechselwirkungen mit Aktivitätendauern, als Wegketten, Tagesabläufe etc., mit Angaben aus anderen Verkehrserhebungen verglichen.

Doch bevor die Qualität des Algorithmus beurteilt werden kann, sollen noch die Rahmenbedingungen dieses Datensatzes bzw. seine Unterschiede zu anderen Verkehrsstudien betont werden, um seine Repräsentativität besser abschätzen zu können:

- Es wurde bisher nur der relativ kleine Testdatensatz von 39 Autofahrern ausgewertet. Werden mehr Teilnehmer in diese nachfolgende Auswertung einbezogen, haben ‚Ausreißer‘ (d.h. Autofahrer mit ungewöhnlichem Mobilitätsverhalten) nicht mehr so großen Einfluss auf das Ergebnis.
- Der andere Unterschied zu bisherigen Studien ist die einmalig lange Beobachtungsdauer von bis zu einem halben Jahr pro Autofahrer. Dadurch, dass es noch keine vergleichbaren Langzeitbeobachtungen einzelner Verkehrsteilnehmer gibt, kann nicht immer geklärt werden, was Ungereimtheiten in der Datenbasis, was Fehler bei der automatisierten Auswertung und was einfach Besonderheiten einer solchen Langzeitstudie sind. Beispielsweise sind die Testfahrer aus Borlänge im Mittel häufiger pro Tag unterwegs und fahren im Mittel kürzere Strecken und kürzere Zeiten am Tag mit dem Auto als die Fahrer aus dem gesamtschwedischen Vergleichsdatsatz RES (vgl. Tabelle 2). Ob diese Fahrweise nun für die Autofahrer in Borlänge tatsächlich charakteristisch ist oder ob eventuell zu viele Stopps aus den ursprünglichen GPS-Daten extrahiert wurden, die gar keine Stopps, sondern nur Datenungenauigkeiten oder beispielsweise Stopps an roten Ampeln sind, bleibt in weiteren Untersuchungen zu klären.

²¹ 28 Vollzeitbeschäftigte und 11 Rentner, vgl. Kapitel 4.2

Tabelle 2 Bearbeitete Borlänge GPS-Daten im Vergleich mit den 2000/2001 RES-Daten*: Ausgewählte Mobilitätscharakteristika

Variable	Rentner		Vollzeitbeschäftigte	
	RES	Borlänge GPS	RES	Borlänge GPS
Median der durchschnittlichen Anzahl täglicher Fahrten (Std.)	3.8 (3.8)	4.3 (2.8)	4.8 (5.7)	3.9 (2.4)
Median der durchschnittlichen täglichen Wegdistanzen [km] (Std.)	24.6 (33.7)	16.3 (12.5)	32.0 (45.3)	13.2 (11.3)
Median der durchschnittlichen täglichen Fahrzeit [min] (Std.)	44.1 (55.2)	31.8 (22.4)	58.7 (82.1)	24.4 (18.2)

*Vergleich entsprechender Fahrten entsprechender Bevölkerungsgruppen;
Std. = Standardabweichung;
Datensätze: RES: N(Vollzeitbeschäftigte) = 1516, N(Rentner) = 440;
Borlänge: N(Vollzeitbeschäftigte) = 28, N(Rentner) = 11
Quelle: Schönfelder und Samaga (2003), S.18

Für die eigentliche Qualitätsanalyse des Algorithmus wurden bisher vier Konsistenzprüfungen vorgenommen. Beim ersten Verfahren wurden die Anteile der in der automatisierten Auswertung ermittelten Wegzwecke (aller Testfahrer zusammen) mit denen der nationalen Verkehrsbefragung verglichen. In einer zweiten Begutachtung wurden tägliche Aktivitätsmuster einzelner Testfahrer untersucht. Weitere Kriterien waren das Verhältnis der Anzahl zurückgelegter Wege zur Anzahl der dabei aufgesuchten Orte und die Häufigkeit, mit der Orte erstmalig aufgesucht wurden. Diese beiden Kriterien wurden bereits für die *Mobidrive*-Untersuchung²² gemessen und konnten für einen Vergleich mit den Borlänge-Daten herangezogen werden.

²² Die *Mobidrive*-Studie wurde im Herbst 1999 in den Städten Karlsruhe und Halle/Saale durchgeführt. Sie untersucht das Mobilitätsverhalten von 317 Personen aus 139 Haushalten. Die Teilnehmer berichteten jeweils sechs Wochen lang die von ihnen zurückgelegten Wege und ihre Aktivitäten. (vgl. Zimmermann, A., K.W. Axhausen, K.J. Beckmann, M. Düsterwald, E. Fraschini, T. Haupt, A. König, A. Kübel, G. Rindsfüser, R. Schlich, S. Schönfelder, A. Simma und T. Wehmeier (2001); PTV AG (2002))

Die prozentualen Verteilungen der verschiedenen Aktivitätenarten in der RES und die der ausgewerteten Borlänge-Testfahrer entsprechen sich weitgehend (vgl. Tabelle 3). Auffällig sind große Differenzen bei den Wegzwecken ‚private Erledigung‘, ‚geschäftliche Erledigung‘ und ‚Einkaufen für den täglichen Bedarf‘. Eine Ursache dafür kann sein, dass an verschiedenen Lokalisationen mehrere Aktivitäten ausgeführt werden können. Wie bereits in 5.4 beschrieben, sind die Wegzwecke im Nachhinein nicht immer klar zu definieren und den vorgegebenen Kategorien zuzuordnen. So ist es beispielsweise nicht zu unterscheiden, ob eine Person zu privaten oder beruflichen Zwecken bei einer Bank war. Die Unsicherheit wird dann noch gesteigert, wenn sich im gleichen oder in benachbarten Gebäuden andere ‚Points of Interest‘, z.B. ein Schnellrestaurant, befinden, die wieder auf einen anderen Wegzweck, z.B. ‚Freizeit‘, hinweisen. Außerdem existieren noch große Lücken bei den Flächennutzungsdaten im Bereich des täglichen Bedarfs wie Bäcker, Fleischer etc. (s.o.). Dass diesem Wegzweck (‚Einkaufen für den täglichen Bedarf‘) zu wenig Stopp-Cluster zugeordnet wurden, verwundert von daher wenig. Es würde sich deshalb lohnen, diese Ungenauigkeit zu beheben. Denn es ist davon auszugehen, dass die Ergebnisse zum ‚Einkaufen für den täglichen Bedarf‘ mit einer konsistenteren Points-of-Interest-Information deutlich verbessert würden. Und eine Verbesserung für den einen Wegzweck bedeutet automatisch eine Veränderung bei der prozentualen Aufteilung der anderen Wegzwecke.

Tabelle 3 Wegzwecke nach automatisierter Auswertung des Borlänge-Datensatzes im Vergleich mit Wegzwecken der 2000/2001 RES-Daten: Wegzweckanteile in [%]

Wegzweck	Rentner		Vollzeitbeschäftigte		
	RES	Borlänge GPS	RES	Borlänge	GPS
Jemanden Abholen/ Wegbringen	6.8	7.0	8.9		8.6
Private Erledigung	4.6	10.0	3.7		8.2
Geschäftliche Erledigung	0.1	9.8	8.3		5.7
Zur Ausbildung/ Schule	0.0	0.1	0.0		0.1
Zum Arbeitsplatz	0.3	0.0	16.0		10.2
Einkauf, täglicher Bedarf	12.4	4.4	6.1		1.5
Einkauf, langfristiger Bedarf	8.6	7.3	5.7		7.9
Freizeit	20.7	23.6	10.6		20.4
Sonstiges/ Nicht zugeordnet	5.2	0.0	4.1		0.0
Zu Hause	41.3	37.8	36.6		37.4

*Vergleich entsprechender Fahrten entsprechender Bevölkerungsgruppen; alle Wochentage; Datensätze: RES: N(Vollzeitbeschäftigte) = 1516, N(Rentner) = 440;

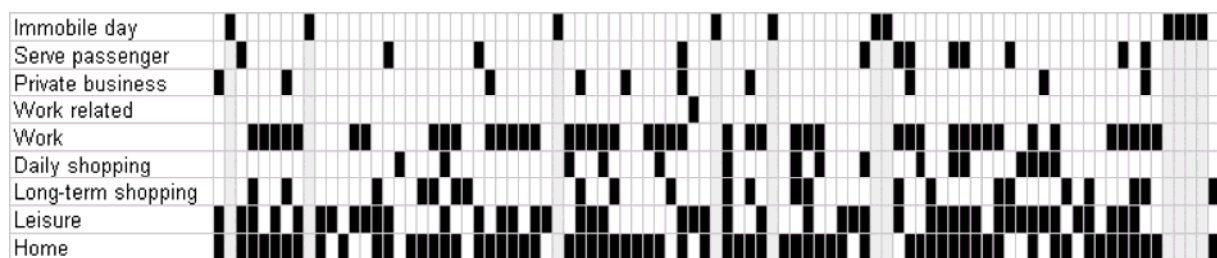
Borlänge: N(Vollzeitbeschäftigte) = 28, N(Rentner) = 11

Quelle: Schönfelder und Samaga (2003), S.19

Um die Konsistenz der Resultate weiter zu überprüfen, wurden außerdem Aktivitätenmuster einzelner Testfahrer untersucht. Diese wurden graphisch ausgewertet (vgl. Abbildung 19). Dabei repräsentiert eine Kästchenspalte immer einen beobachteten Tag. Die schwarze Füllung bedeutet, dass an diesem Tag die jeweilige Aktivität durchgeführt wurde, wobei es keine Rolle spielt, wie lange und wie oft dieser Beschäftigung nachgegangen wurde. Die grau gefärbten Kästchen kennzeichnen die immobilen Tage. Daraus lassen sich nun verschiedene Ergebnisse ableiten. Erstens kann man Blöcke von fünf Tagen Arbeit und dann zwei Tagen Wochenende erkennen. An Tagen ohne Mobilität konnte natürlich auch kein Weg zu der Aktivität ‚zu Hause‘ führen. Allerdings sollte an allen anderen Tagen mindestens ein Weg am Abend nach Hause geführt haben. Es ist zwar möglich, dass eine Person auch einmal an einem anderen Ort übernachtet bzw. erst nach Mitternacht heimkehrt, dieser Fall kommt in der Borlänge-

Auswertung jedoch zu häufig vor (bis zu 25% für einen Testfahrer). Diejenigen weißen Felder beim Zweck ‚zu Hause‘ lassen also Fehler im Algorithmus bzw. bei der Datenbereinigung vermuten. Es wurden wahrscheinlich zu viele Wege vor der Analyse aussortiert. Die andere mögliche Fehlerquelle, dass im Analysealgorithmus nicht alle Fahrten nach Hause also solche erkannt wurden, kann so gut wie ausgeschlossen werden, da für die bisher analysierten Testfahrer jeweils die Wohnadresse bekannt war.

Abbildung 19 Aktivitätsmuster eines vollzeitbeschäftigten Testfahrers über den gesamten Beobachtungszeitraum



Quelle: nach Schönfelder und Samaga (2003), S.20

Die oben bereits angeführten Vergleiche mit den *Mobidrive*-Auswertungen zeigen Folgendes: Aus den *Mobidrive*-Daten wurde bereits eine Korrelation zwischen der Anzahl zurückgelegter Wege und der dabei aufgesuchten Orte abgeleitet. Es konnte dort festgestellt werden, dass sehr mobile Personen in der Regel nicht nur viele Wege (zu wenigen Orten) zurücklegen, sondern auch eine Vielzahl von Aktivitätenorten nutzen. In der *Mobidrive*-Studie wurde dabei ein Verhältnis von ca. 1:5 ermittelt. Das bedeutet, dass die von einer Person genutzten Aktivitätenorte in einer Stadt im Schnitt einem Fünftel der von ihr insgesamt gefahrenen Wege entspricht. Ein ähnlicher Wert kann für die Borlänge-Daten abgeleitet werden. (vgl. Abbildung 20; Schönfelder und Axhausen, 2003)

Abbildung 20 Anzahl zurückgelegter Wege im Verhältnis der Anzahl aufgesuchter Orte

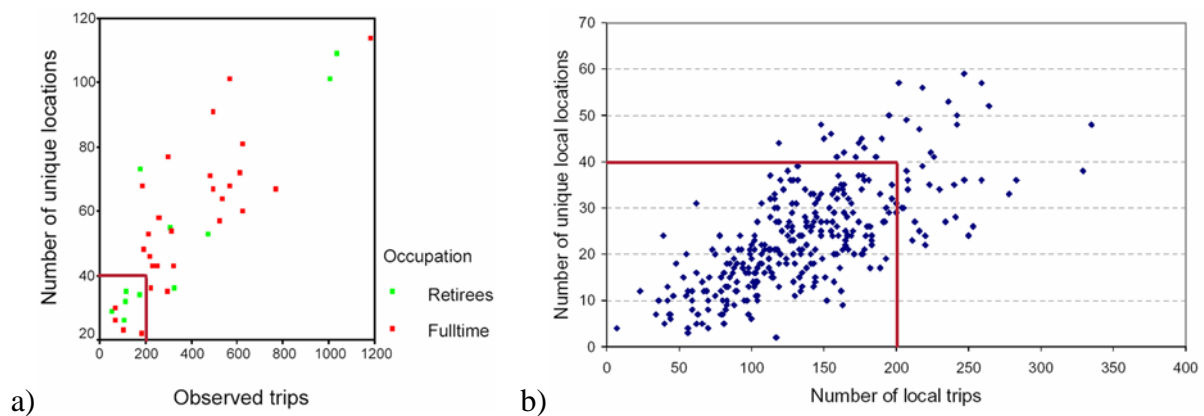


Abbildung a) zeigt das Verhältnis für die bereits ausgewerteten Borlänge-Fahrer, b) das Verhältnis aus den Mobidrive-Daten. Die dunkelroten Striche sind als Skalenvergleich eingefügt worden.

Quelle: a) nach Schönfelder und Samaga (2003), S.21
b) nach Schönfelder und Axhausen (2003), S.18

Außerdem wurde sowohl in der Mobidrive- als auch in der Borlänge-Untersuchung analysiert, in welchen zeitlichen Abständen immer wieder neue Orte aufgesucht werden (vgl. Abbildung 21). Die Mobidrive-Daten zeigten bereits, dass zumindest in einem Zeitraum von sechs Wochen noch kontinuierlich, wenn auch in leicht abnehmender Anzahl, neue Orte hinzukommen. Diese Aussage trifft auch für die Borlänge-Daten zu. Wie sich diese Zahlen nach sechs Wochen weiterentwickeln, war bisher unbekannt. Wie bereits oben erwähnt, ist bei diesen neuen Daten ohne Vergleichswerte aus anderen Studien stets zu hinterfragen, ob neue Erkenntnisse gewonnen werden können oder ob Fehler in den Daten oder bei der Interpretation vorhanden sind. Eine mögliche Fehlerquelle in diesem Fall könnten nämlich die falsch interpretierten GPS-Daten sein, bei denen Stopps an Ampeln, in Staus etc. eventuell als Wegziele behandelt wurden, die keine Wegziele sind. Auf diese Weise würde natürlich aus jedem fälschlicherweise als Wegziel identifizierten Stopp ein neu aufgesuchter Ort. Und dass diese Orte über die Stadt verteilt an immer wieder neuen Stellen auftreten, verwundert nicht. Allerdings wurden bisher scheinbar tendenziell eher zu viele Wege aussortiert (s.o.), sodass man zumindest für die hier betrachteten Testfahrer davon ausgehen kann, dass die Fehlinterpretation von Stopppositionen keine Rolle spielt. Ein Fehler in der Datengrundlage kann in diesem Fall ebenfalls ausgeschlossen werden, denn im Gegensatz zu Verkehrsbefragungen, bei denen ein Teilnehmer mehr oder weniger über seine Mobilität berichten kann, wird bei der GPS-Beobachtung

stets die tatsächliche Mobilität erfasst. Aus diesen Gründen kann davon ausgegangen werden, dass tatsächlich auch nach sechs Wochen immer wieder neue Orte aufgesucht werden. Eine mögliche Erklärung für diesen speziellen Fall könnte aber auch sein - da die meisten bisher ausgewerteten Teilnehmer an der Studie von Oktober/November 2000 bis ca. Mai 2001 teilgenommen haben (vgl. Anhang B) - , dass zu verschiedenen Jahreszeiten verschiedene Orte aufgesucht werden, weil dann z.B. andere Freizeitmöglichkeiten zur Verfügung stehen.

Abbildung 21 Zeitliche Verteilung neu aufgesuchter Orte

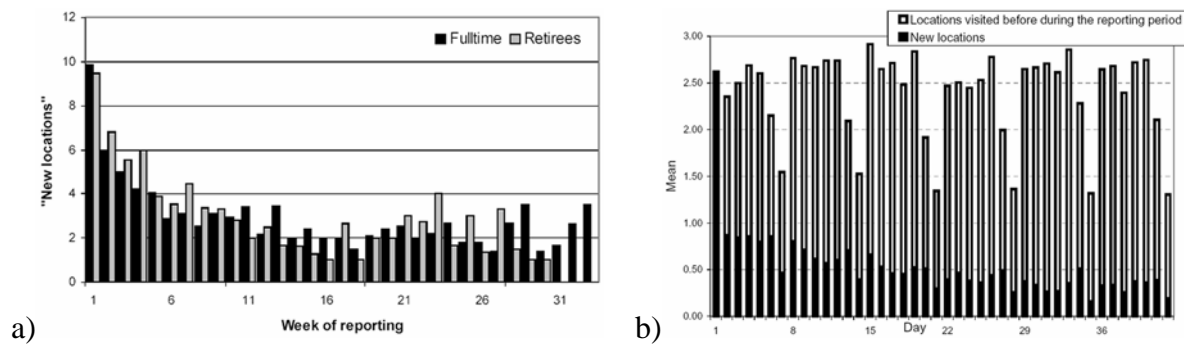


Abbildung a) zeigt den Median der erstmalig aufgesuchten Orte während der Rätt fart-Studie pro Woche, Abbildung b) zeigt den Median der neu aufgesuchten Orte pro Tag der Mobidrive-Untersuchung.

Quelle: a) nach Schönfelder und Samaga (2003), S.21
b) nach Schönfelder und Axhausen (2003), S.19

7. Ausblick

Die obige kurze Auswertung aus Sicht der Verkehrswissenschaft hat gezeigt, dass sowohl die durch die Verkehrssicherheitsstudie ‚Rätt Fart‘ gewonnenen Daten als auch der hier angewendete Analysealgorithmus interessante und verwendbare Ergebnisse liefern. Es sollte deshalb noch einmal die Einzigartigkeit dieses Datensatzes betont werden, der erstmals Mobilitätsverhalten einzelner Verkehrsteilnehmer über einen so langen Zeitraum untersucht und beweist, dass Mobilitätsstudien auch in Zukunft auf diese Art und Weise durchgeführt werden können. Allerdings sollten dafür einige Punkte zu den Studienbedingungen, zur Datenaufbereitung vor der Auswertung und zum Analysealgorithmus selbst berücksichtigt werden, die in dieser Arbeit noch nicht zufriedenstellend waren.

7.1 Verbesserungen bei den Studienbedingungen

Das Charakteristische dieser Studie ist, dass sie ausschließlich Autofahrer in ihrem Mobilitätsverhalten untersucht und von diesen nur den Teil der täglichen Bewegung, der mit dem Auto zurückgelegt wird. Aussagekräftige Studien über das Verkehrsverhalten müssten ebenso andere Verkehrsmittel einbeziehen. In Zukunft sollten die technischen Probleme wie die Stromversorgung keine so große Rolle mehr spielen, da entsprechende portable Geräte produziert werden (GeoStats, 2003). Der zweite Hinderungsgrund, derartige Langzeitstudien für alle Verkehrsmittel durchzuführen, bleibt allerdings, wie bei den herkömmlichen Befragungen, der Aufwand für die Teilnehmer. Werden sie gebeten, sich immer mit den portablen Geräten zu bewegen, entstehen wieder große zusätzliche Belastungen.

Unabhängig vom Verkehrsmittel könnte die Genauigkeit einer nächsten Studie dadurch verbessert werden, dass der GPS-Empfänger nicht nur angeschaltet ist, während der Motor läuft, sondern permanent Signale annimmt. Dadurch würde die Unsicherheit wegfallen, wie lange das Gerät jeweils gebraucht hat, um die erste Position zu ermitteln, und die Schwellwerte für minimale Reise- und Aktivitätendauern könnten überdacht und neu definiert werden. Die Präzision der Daten wird sich zudem durch differentielles GPS (DGPS) verfeinern. Bei den Borlänge-Daten kam es aber noch nicht zum Einsatz.

Eine weitere Verbesserung für die Aussagekraft einer solchen Studie wäre eine interaktive Eingabemöglichkeit für die Identifizierung des tatsächlichen Fahrers, eventuell ergänzt durch

die Abfrage, wie viele Personen außer ihm im Auto sitzen. Es müssten sicherlich nicht, wie in bisherigen anderen Studien, vollständige Beschreibungen der zurückgelegten Wege erbracht werden. Dadurch sind die Studienteilnehmer stets belastet und es könnte wieder Motivationsprobleme für Langzeitstudien geben. Durch einmalige kurze Identifizierung des Fahrers entstünde nicht viel Aufwand für die Teilnehmer und bei der Auswertung könnten die Wege effektiv der richtigen Testperson zugeordnet werden. Eventuell könnte diese Lücke auch, wie Wolf (2003) vorschlägt, durch GIS-basierte Interfaces im Internet geschlossen werden, bei denen Studienteilnehmer die von ihnen zurückgelegten Wege kontrollieren und mit Zusatzinformationen versehen können. Das erfordert allerdings wieder eine zeitliche Zusatzbelastung für die Teilnehmer.

In Bezug auf die konkrete Situation in Borlänge wäre eine weitere Verbesserung in der Auswertung beim Arbeitsschritt der Clusterung möglich wenn die originalen GPS-Positionen bekannt wären.

7.2 Verbesserungen bei der Aufbereitung der Rohdaten

Die Mängel in den Rohdaten wie die zu kurzen Stopps, Wegziele mit unwahrscheinlich kurzer Aktivitätendauer etc. sollten, wie bereits oben erwähnt, nicht einfach eliminiert, sondern korrigiert werden, was derzeit auch untersucht wird (Axhausen *et al.*, 2003). Anschließend könnte noch kontrolliert werden, ob andere Schwellwerte für die zu kurzen Reisezeiten und Aktivitätendauern etc. verwendet werden müssten.

Eine etwas akkuratere Aussage über die wahrscheinlichen Wegzwecke innerhalb der Wochentagskategorien könnte dadurch gewonnen werden, dass allgemeine Feiertage ermittelt und noch mit zur Kategorie ‚Sonntag‘ gezählt werden würden.

7.3 Verbesserungen im Analyseprozess

Eine erste wesentliche Verbesserung im Analyseprozess würde sich ergeben, wenn die Informationen über die Flächennutzung und die ‚Points of Interest‘ konsistenter wären (vgl. Kapitel 5.4.1). Gleiches gilt für die Straßen. Die Ergänzung mit privaten Wegen und eventuell auch Fußwegen lohnt sich allerdings nur, wenn Originalpositionen der Wagen zur Verfügung stehen.

Stünden diese Weginformationen, vor allem die Fußwege, zur Verfügung, könnte der Prozess dahingehend verbessert werden, dass für die Cluster- und alle Pufferprozesse nicht mehr die Luftliniendistanz, sondern abgeschätzte Gehzeiten als Entfernungskriterium genutzt würden.

Eventuell würde das Ergebnis noch leicht verbessert, wenn der Abgleich zwischen den Borlänge-Daten und den zeitlichen Informationen aus der RES genauer würde. Dies könnte erreicht werden, indem statt der Mediane der Ankunftszeiten und Aktivitätendauern eines Clusterzentrums der Abgleich für jede in diesem Clusterzentrum integrierte Stopp-Position durchgeführt würde und erst aus diesem Ergebnis der Modalwert als wahrscheinlicher Wegzweck ermittelt würde. Eine größere Genauig- oder Verlässlichkeit der Aussagen würde sich allerdings durch passendere Vergleichsdaten ergeben, d.h. beispielsweise eine passend große Stichprobe der RES-Daten aus Borlänge selber.

Eventuell könnten gerade diese letzten Schritte des automatisierten Prozesses dahingehend gänzlich umgestaltet werden, dass komplexere statistische Modelle umgesetzt würden wie z.B. eine Diskriminanzanalyse (dazu vgl. Kapitel 5.4.3).

Außerdem sollten noch weitere zeitliche Regelmäßigkeiten als der Vergleich mit den RES-Daten einbezogen werden. Dies kann aber erst mit einem komplexeren Analyse-Modell geschehen. Im bisherigen Modell konnte Stopp-Clustern, die eine große zeitliche Regelmäßigkeit aufzeigen (z.B. wenn 9 von 10 Aktivitäten eines Clusterzentrums in der Nähe der Sporthalle immer am Dienstagabend für zweieinhalb Stunden stattfanden), keine entsprechende Bewertung im Prozess eingeräumt werden. Denn an dem Abgleich mit der einzigen zeitlichen Referenzinformation, die zur Verfügung steht – den RES-Daten –, ändert sich dadurch nichts. Das Ergebnis bliebe gleich, da sowieso der Median aller im Cluster enthaltenen Werte als Ausgangswert zählt. Von daher ist das Verfahren langfristig vielleicht doch dahingehend zu verändern, nicht ausschließlich auf die zusammengefassten Cluster, sondern auch auf die einzelnen Stopp-Positionen zuzugreifen. Dies erfordert aber größere Speicherkapazitäten. Und dabei dürfen nicht das Ziel, Mobilitätsmuster zu erkennen (d.h. Trends und Verteilungen) und die Ausgangssituation und -genauigkeit, aus der die Daten stammen (nämlich die vollständig computergestützte Verkehrsbeobachtung, bei der eh niemals ein einzelner Wegzweck gänzlich zweifelsfrei rekonstruiert werden kann), aus den Augen verloren werden. Behält man diese Ausgangssituation der Daten und die Zielstellung, möglichst plausible Annahmen zu treffen, Muster zu erkennen und den Prozess für möglichst viele Fälle sinnvoll zu automatisieren, stets in Erinnerung, ist es wieder legitim, die Vermutung anzustellen, dass ein gleiches Wegziel auch auf einen gleichen Wegzweck hindeutet. Stimmt diese Vermutung in einigen Fällen

nicht, kann der tatsächliche Wegzweck sehr wahrscheinlich sowieso nicht rekonstruiert werden.

Der in diesem Diplom erarbeitete Ansatz für die automatisierte Auswertung von GPS-basierten Mobilitätsbeobachtungen ohne Teilnehmerinteraktion zeigt also ein erstes grobes Konzept dafür. Er liefert das Gerüst, in das oben beschriebene und andere das Verfahren verfeinernde Arbeitsschritte integriert werden können.

8. Literatur

- Adler, T., L. Rimmer, G. Bandy und D. Schellinger (2000) Use of respondent-interactive Geocoding in Baltimore, Maryland, mode choice survey, *Transportation Research Record*, **1719**, 154-158.
- Arnberger, E. (1993) Thematische Kartographie, Westermann Schulbuchverlag GmbH, Braunschweig.
- Axhausen, K.W. (2003) <http://gallium.ethz.ch/vrp/lesung/vpl3.pdf>, April 2003.
- Axhausen, K.W. und M. Bierlaire (2001) Structure and use of human activity spaces, Projektantrag, IVT, ETH, Zürich und ROSO, EPF, Lausanne.
- Axhausen, K.W., S. Schönfelder, J. Wolf, M. Oliveira und U. Samaga (2003) 80 weeks of GPS-traces: Approaches to enriching the trip information, Arbeitsberichte Verkehrs- und Raumplanung, **178**, IVT, ETH Zürich, Zürich.
- Bahrenberg, G., E. Giese und J. Nipper (1992) Statistische Methoden in der Geographie: ‚Multivariate Statistik‘, B. G. Teubner, Stuttgart.
- Bill, R. und D. Fritsch (1991) Grundlagen der Geo-Informationssysteme: ‚Hardware, Software und Daten‘, Wichmann, Karlsruhe.
- Bollmann, J. und W.G. Koch (Hrsg.) (2001) Lexikon der Kartographie und Geomatik in zwei Bänden, Spektrum Akademischer Verlag GmbH, Heidelberg, Berlin.
- Brockhaus AG (2001) Der Brockhaus multimedial 2002, Bibliographisches Institut & F.A. Brockhaus AG, Mannheim.
- Buhmann E., R. Bachhuber und J. Schaller (1996) ArcView: GIS-Arbeitsbuch, Wichmann Verlag, Heidelberg.
- Columna (2003) <http://www.columna.se/default.asp>, Januar 2003.
- Canzler, W. (1999) Der anhaltende Erfolg des Automobils. Zu den Modernisierungsleistungen eines außergewöhnlichen technischen Artefakts, in *Technik und Gesellschaft – Jahrbuch 10: Automobil und Automobilismus*, 19-40, Campus Verlag, Frankfurt/Main.
- Doherty S.T., N. Noël, M. Lee-Gosselin, C. Sirois und F. Theberge (2001) Moving beyond observed outcomes: integrating global positioning systems and interactive computer-based travel behaviour surveys, http://www.wlu.ca/~sdoherty/downloads/pap_gis.pdf, *Transportation Research Circular "Personal Travel: The Long and Short of It"*, Transportation Research Board, Washington.
- Downs, R. und D. Stea (1977) Maps in mind – reflections on cognitive mapping, Harper & Row, New York.
- Draijer, G., N. Kalfs und J. Perdok (2000) Global Positioning System as data collection method for travel research, *Transportation Research Record*, **1719**, 147-153.

- Drosdowski G., W. Müller, W. Scholze-Stubenrecht und M. Wermke (Hrsg.) (1990) Fremdwörterbuch, 5. Aufl., Dudenverlag, Mannheim, Wien und Zürich.
- FHWA (1994) <http://ntl.bts.gov/DOCS/surveys.html>, Januar 2003.
- FHWA (2003a) <http://www.fhwa.dot.gov/ohim/lextrav.pdf>, Januar 2003.
- FHWA (2003b) <http://www.fhwa.dot.gov/ohim/womens/post4.pdf>, Januar 2003.
- French, S.P., W. Bachman und L.D. Frank (2001) Regional land use database: Descriptive analysis, Atlanta.
- GeoStats (2003) <http://www.geostats.com/products.php>, Januar 2003.
- Golledge, R.G. and R.J. Stimson (1997) *Spatial Behavior*, The Guilford Press, New York.
- GSM Association (2003) <http://www.gsmworld.com/technology/faq.shtml>, Januar 2003.
- Gula Sidorna (2003) <http://www.gulasidorna.se/?lang=en>, Februar 2003.
- PTV AG (2002) <http://www.ptv.de/mobidrive/MobiDriv/welcome.html>, April 2003.
- Infas und DIW (Hrsg.) (2001) KONTIV 2001, Kontinuierliche Erhebung zum Verkehrsverhalten, Methodenstudie, Endbericht, Infas (Institut für angewandte Sozialwissenschaft) und DIW (Deutsches Institut für Wirtschaftsforschung), Bonn und Berlin.
- Infas und DIW (2003) <http://www.kontiv2002.de/index.htm>, April 2003.
- Jan, O., A.J. Horowitz und Z.-R. Peng (2000) Using global positioning system data to understand variations in path choice, *Transportation Research Record*, **1725**, 37-44.
- Jennrich, R.I. und F.B. Turner (1969) Measurement of non-circular home range, *Journal of Theoretical Biology*, **22**, 227-237.
- Knoflacher, H. (1999) Wo uns das Auto erwisch hat, in *Technik und Gesellschaft – Jahrbuch 10: Automobil und Automobilismus*, 121-128, Campus Verlag, Frankfurt/Main.
- Kreitz, M. (2002) CHASE-GIS: Erhebung raumbezogener Daten mit einem computer- und GIS-gestützten Aktivitätentagebuch in einer Haushaltserhebung zum Verkehrsverhalten, *Dissertation*, Rheinisch-Westfälische Technische Hochschule Aachen.
- Landmäteriet (2002) <http://butiken.lmv.lm.se/butiken/fastigh.htm>, September 2002.
- Lee-Gosselin, M. (2000) Update on the GPS travel survey prototyping projects at Université Laval, Quebec City, Canada, www.grimes.ulaval.ca/grimes_v3/documents/gps_sst.pdf (Stand 18.11.2002).
- Lee-Gosselin, M. (2002) Some reflections on GPS-supported travel survey methods in an increasingly ICT-rich environment, Vortrag beim Workshop on ICT, Innovation and the Transport System, Arlington, Januar 2002.
- Lynch, D. (1960) *The Image of the City*, MIT-Press, Cambridge.

- Murakami, E. und D.P. Wagner (1999) Can using global positioning system (GPS) improve trip reporting?, *Transportation Research Part 7C*, 149-165.
- Murakami, E., D.P. Wagner und D.M. Neumeister (1997) Using global positioning system and personal digital assistants for personal travel surveys in the United States, http://www.nas.edu/trb/publications/ec008/session_b.pdf, *International Conference on Transport Survey Quality and Innovation*, Grainau, Mai 1997.
- Psychologie TU BS (2003) http://www.psychologie.tu-bs.de/studium/manuale/spss/25_Diskriminanzanalyse.pdf, TU Braunschweig, Februar 2003.
- Rätt fart i Borlänge (2003a) <http://www.rattfart.com>, Januar 2003.
- Rätt fart i Borlänge (2003b) <http://www.rattfart.com/pdf/isateknik.pdf>, Januar 2003.
- Rätt fart i Borlänge (2003c) <http://www.rattfart.com/pdf/isa8sids.pdf>, Januar 2003.
- SAS Institute Inc. (1994) SAS/STAT®, User's Guide, Volume 1, Version 6, 4th Edition.
- Schönfelder, S. und K.W. Axhausen (2002) Measuring the size and structure of human activity spaces – The longitudinal perspective, *Arbeitsbericht Verkehrs- und Raumplanung*, **135**, IVT, ETH Zürich.
- Schönfelder, S. und K.W. Axhausen (2003) Activity spaces: Measures of social exclusion?, *Arbeitsbericht Verkehrs- und Raumplanung*, **140**, IVT, ETH Zürich.
- Schönfelder, S., K.W. Axhausen, N. Antille und M. Bierlaire (2002) Exploring the potentials of automatically collected GPS data for travel behaviour analysis – A Swedish data source, <http://www.ivt.baum.ethz.ch/vrp/ab124.pdf>, *Arbeitsbericht Verkehrs- und Raumplanung*, **124**, IVT, ETH Zürich.
- Schönfelder, S. und U. Samaga (2003) Where do you want to go today? – More observations on daily mobility, *3rd Swiss Transport Research Conference*, Monte Verità / Ascona, März 2003.
- Schmidt, G. (1999) „Die Zeit fährt Auto“, in *Technik und Gesellschaft – Jahrbuch 10: Automobil und Automobilismus*, 7-14, Campus Verlag, Frankfurt/Main.
- Schwarze, B. und S. Schönfelder (2001) ArcView-Extension VISAR – Visualisierung von Aktionsräumen, Version 1.6, *Arbeitsbericht Verkehrs- und Raumplanung*, **95**, IVT, ETH Zürich.
- Schwesig, R. (1988) Räumliche Strukturen von Außerhausaktivitäten – Ein Konzept zur Analyse räumlichen Verhaltens und empirische Überprüfung am Beispiel der Aktionsräume von Bewohnern der Stadtregion Hamburg, *Dissertation*, Universität Hamburg.
- Sika Institut (Hrsg.) (2001) RES 2000, Den nationella resundersökningen, Statistiska Centralbyrån, Stockholm.
- Waßmuth, V. (2001) Modellierung der Wirkungen verkehrsreduzierender Siedlungskonzepte, *Dissertation*, Universität Karlsruhe.

- Wolf, J. (2000) Using GPS data loggers to replace travel diaries in the collection of travel data, *Dissertation*, Georgia Institute of Technology, Atlanta.
- Wolf, J. (2003) Tracing people and cars with GPS and diaries: Current experience and tools, *Vortrag ETH Zürich*, Zürich, Februar 2003.
- Wolf, J., R. Guensler, L. Frank und J. Ogle (2000) The use of electronic travel diaries and vehicle instrumentation packages in the year 2000 Atlanta regional household travel survey: test results, package configurations, and deployment plans, *9th International Association of Travel Behaviour Research Conference*, Gold Coast, Queensland, Australia.
- Wolf, J., R. Guensler und W. Bachman (2001a) Elimination of the travel diary: An experiment to derive trip purpose from GPS travel data, *Transportation Research Board*, 80th Annual Meeting, Washington, D.C.
- Wolf, J., M. Loechl, J. Myers und C. Arce (2001b) Trip rate analysis in GPS-enhanced personal travel surveys, *International Conference on Transport Survey Quality and Innovation*, Kruger Park, South Africa, August 2001.
- Zimmermann, A., K.W. Axhausen, K.J. Beckmann, M. Düsterwald, E. Frascini, T. Haupt, A. König, A. Kübel, G. Rindsfuser, R. Schlich, S. Schönfelder, A. Simma und T. Wehmeier (2001) *Mobidrive: Dynamik und Routinen im Verkehrsverhalten: Pilotstudie Rhythmik*, Forschungsbericht an das Bundesministerium für Forschung und Technologie, PTV AG, Institut für Verkehrsplanung, Transporttechnik, Strassen- und Eisenbahnbau (ETH Zürich) und Institut für Stadtbauwesen (RWTH Aachen), Karlsruhe, Zürich und Aachen.

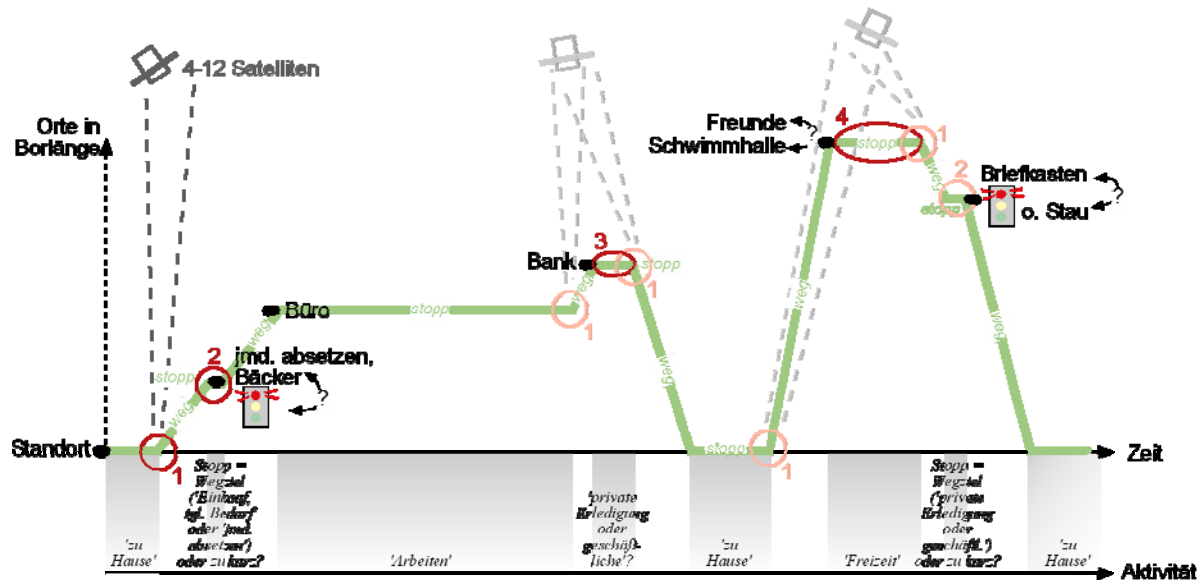
Anhang A: Glossar

A 1: Begriffe der Mobilitätsforschung

Aktivität	Eine Aktivität ist verkehrswissenschaftlich eine bedeutungsvolle Interaktion mit Menschen oder Umwelt bei gleichbleibendem räumlichen Umfeld und sozialer Konstellation. Sie wird durch ihre Art (was tut jemand?), durch den Zweck und Sinn (wozu? und warum?), durch eine bestimmte Zeit (wann?), ihre Dauer (wie lange?) und den Ort (wo?), an dem sie stattfinden, charakterisiert.
Aktivitäten-kategorie	Verschiedene Aktivitäten werden zusammengefasst, um Studienergebnisse verallgemeinern zu können. Kategorien sind dabei z.B. ‚Arbeiten‘, ‚Ausbildung‘ und ‚Freizeit‘.
Aktivitäten-raum	Der Aktivitätenraum einer Person wird durch die Orte aufgespannt, an denen sie ihren Aktivitäten nachgeht.
Bewegung	Durch Bewegung wird stets der Ort der nächsten Aktivität erreicht.
Etappe	Eine Etappe ist die Bewegung mit einem Verkehrsmittel, respektive Fahrzeug, inklusive der reinen Wartezeiten unmittelbar vor und während der Bewegung. In der Borlänge-Untersuchung sind allerdings nur die Etappen mit dem eigenen PKW aufgezeichnet, sodass sie hier den Wegen entsprechen und nicht extra behandelt werden.
Reise	Eine Reise besteht aus einer Folge von Wegen vom Standort aus und wieder zurück.
Standort	Der Standort wird als jeweiliger räumlicher Bezugspunkt definiert, von dem Reisen ausgehen und wieder enden. Das kann ein Hotel o.ä. sein, in das immer der letzte Weg eines Tages führt. In der Untersuchung in Borlänge sind es jedoch die Wohnungen der Studienteilnehmer.
Stopp	Stops sind in GPS-Daten durch gleichbleibende Ortsangaben zu erkennen, d.h. das beobachtete Fahrzeug hat sich für eine bestimmte Dauer nicht bewegt. Die Schwierigkeit besteht darin, zu erkennen, wann dieser Stopp auch ein Wegziel darstellt und wann er aus den verkehrstechnischen Bedingungen entsteht (vgl. Abbildung 22).
Stopp-Cluster	Zu einem Cluster werden diejenigen Stopp-Positionen zusammengefasst, deren Wege vermutlich das gleiche Wegziel hatten.
Stopp-Position	Eine Stopp-Position ist die Ortsangabe, bei der in den GPS-Daten festgestellt wird, dass keine Bewegung mehr stattfindet.
Tour	Eine Tour geht von einem beliebigen Ort aus und wieder an ihn zurück.
Weg/ Fahrt	Ein Weg ist die Folge von Etappen zwischen zwei Aktivitäten. D.h. in den Borlänge-Daten entsprechen die ermittelten Etappen den Wegen, da keine Verkehrsmittel außer eigenen PKW erfasst wurden.
Wegziel/ Wegzweck	Ein Wegziel ist dann erreicht, wenn das Fahrzeug stoppt, weil der Fahrer einer Aktivität nachgehen will. Somit erhält der zurückgelegte Weg einen „Wegzweck“. Dieser entspricht der Aktivität, der der Fahrer nachgeht.

Quelle: allgemeine Definitionen aus Axhausen (2003), ergänzt durch spezielle für die Situation in Borlänge

Abbildung 22 Zusammenhang Wege-Stopps-Aktivitäten anhand einer möglichen Konstellation für den Tagesverlauf eines Testfahrers



- Q₁ Es ist nie bekannt, wie lange es dauert, die erste Position nach dem Starten des Fahrzeugs durch das GPS zu bestimmen. So entstehen Ungenauigkeiten über die Zeitdauern der vorhergehenden Aktivität und des nächsten Wegs.
- Q₂ Aus den GPS-Daten lässt sich lediglich erkennen, dass sich das Fahrzeug für eine kurze Zeit nicht bewegt hat. Es ist dabei unklar, ob es gehalten hat, weil ein Wegziel/ Stopp erreicht war oder ob es sich lediglich um einen verkehrstechnischen Stopp (durch rote Ampeln, Stau etc.) handelt.
- Q₃ An dieser Stelle dürfte aufgrund der Stopp-Dauer zu vermuten sein, dass der Fahrer anhält, um einer Aktivität nachzugehen. Da sich in unmittelbarer Nähe des Parkplatzes eine Bank befindet, ist es sehr wahrscheinlich, dass sie das Wegziel war. Bei der Zuordnung der Aktivität kann man aber kaum entscheiden, ob der Studienteilnehmer dort private oder geschäftliche Erledigungen getätigt hat.
- Q₄ Am Abend hält der Testfahrer in einem Wohngebiet in der Nähe einer Schwimmhalle. Es kann deshalb möglich sein, dass er Freunde besucht oder schwimmen geht. Die zugehörige Aktivitätenkategorie ist in jedem Fall ‚Freizeit‘.

A 2: Abkürzungsverzeichnis

AML	ArcMacroLanguage
DGPS	Differential GPS
GIS	GeoInformationSystem
GPS	Global Positioning System
GSM	Global System for Mobile Communications
IKA	Institut für Kartographie, ETH Zürich
IfK	Institut für Kartographie, TU Dresden
ISA	Intelligent Speed Adaption
IVT	Institut für Verkehrsplanung und Transportsysteme, ETH Zürich
KONTIV	Kontinuierliche Erhebung zum Verkehrsverhalten
PDA	Personal Digital Assistant
POI	Point of Interest
RES	Schwedische Verkehrserhebung
SAS	Statistical Analysis System (Statistiksoftware)
SMARTAQ	Strategies for Metropolitan Atlanta's Regional Transportation and Air Quality

Anhang B: „Datengeschichte“

Tabelle 4 Datengeschichte nach Testfahrern

Testfahrer (NR)	Vollzeit- besch. (FT)/ Rentner (RT)	Erfasste Trips	Eliminierte Trips bei Datenim- port in GIS	Eliminierte Trips wegen zu gering- er Reisezeit, Ak- tivitätendauer etc.	Anzahl Tage Studienteil- nahme (von - bis Datum)		Anzahl Tage	Durch- schnitt- liche Anzahl Wege pro Tag
Car13	FT	332	0	-103	10.11.2000	28.05.2001	200	1.15
Car23	FT	799	0	-177	07.11.2000	28.05.2001	203	3.06
Car25	FT	943	0	-465	09.10.2000	27.05.2001	231	2.07
Car27	FT	882	0	-32	16.10.2000	27.05.2001	224	3.79
Car31	FT	678	0	-108	15.12.2000	28.05.2001	165	3.45
Car43	FT	306	0	-11	01.01.2001	28.05.2001	148	1.99
Car56	FT	251	0	-64	04.04.2001	09.06.2001	67	2.79
Car63	FT	319	-1	-97	05.04.2001	26.06.2001	83	2.66
Car68	FT	280	0	-45	08.12.2000	23.04.2001	137	1.72
Car77	FT	681	0	-190	18.10.2000	17.04.2001	182	2.70
Car79	FT	110	0	-6	08.01.2001	04.02.2001	28	3.71
Car93	FT	114	0	-49	04.04.2001	13.06.2001	71	0.92
Car96	FT	266	-1	-41	08.11.2000	17.04.2001	161	1.39
Car101	FT	395	-1	-84	10.11.2000	14.05.2001	186	1.67
Car105	FT	545	0	-185	26.10.2000	28.05.2001	215	1.67
Car118	FT	87	-6	-13	13.04.2001	03.06.2001	52	1.31
Car119	FT	242	0	-66	11.10.2000	22.02.2001	135	1.30
Car131	FT	641	-1	-335	12.10.2000	28.05.2001	229	1.33
Car140	FT	305	0	-51	26.03.2001	28.05.2001	64	3.97
Car144	FT	351	0	-39	27.03.2001	28.05.2001	63	4.95
Car148	FT	702	-1	-133	20.10.2000	26.03.2001	158	3.59
Car150	FT	282	0	-31	14.02.2001	28.05.2001	104	2.41
Car157	FT	737	0	-102	16.10.2000	28.05.2001	225	2.82
Car165	FT	572	-1	-27	12.10.2000	28.05.2001	229	2.38
Car195	FT	1532	0	-354	16.10.2000	21.05.2001	218	5.40
Car202	FT	187	0	-4	12.03.2001	30.04.2001	50	3.66
Car205	FT	673	-1	-155	13.10.2001	21.05.2001	221	2.34
Car228	FT	710	0	-88	10.12.2000	28.05.2001	170	3.66
Car11	RT	479	0	-170	03.04.2001	14.08.2001	148	2.09

Car48	RT	292	0	-118	03.04.2001	12.06.2001	71	2.45
Car52	RT	1176	0	-171	11.10.2000	27.05.2001	229	4.39
Car65	RT	131	0	-14	19.03.2001	23.04.2001	36	3.25
Car84	RT	130	0	-17	20.03.2001	19.04.2001	31	3.65
Car85	RT	552	-1	-77	20.10.2001	17.04.2001	180	2.63
Car133	RT	1138	0	-103	01.11.2000	28.05.2001	209	4.95
Car190	RT	347	0	-18	10.11.2000	21.05.2001	193	1.70
Car192	RT	252	0	-72	05.04.2001	12.06.2001	69	2.61
Car209	RT	87	0	-34	29.03.2001	21.05.2001	54	0.98
Car211	RT	140	0	-32	07.04.2001	12.06.2001	67	1.61

Anhang C: Codierungen zwischen Flächennutzungen und Reisezwecken

C 1: Einbeziehung des digitalen Flächennutzungsplans

Tabelle 5 Typen der Flächennutzung und ihre Einbeziehung in die Auswertung

Flächennutzung	Mögliche Zwecke									
	Pick up/drop off	Private Business	Work related Business	School	Work	Shopping, Daily	Shopping, Long Term	Leisure	Other	Home
Geschlossene Bebauung	1	1	1			1				
Offene Bebauung	1	1	1			1	1	1		
Wohnbebauung	1							2		2
Industriegelände		1	2		1		1			
Arbeitsplätze					2					
Wald und anderes offenes Land								2		
Hauptsächlich Wald								2		
Umweltschutzgebiet								2		
Hauptsächlich Wald im Umweltschutzgebiet								2		
Acker								1		
Wasserflächen								3		
Wasserflächen im Industriegebiet					1					
Friedhof								3		
Trabrennbahn								3		
Flughafen	3	1	1							
Steinbruch					2					
Baumschule					1		1			
Altersheim	1							2		
Campingplatz	1							2		
Krankenhaus	1							2		

Kläranlage					1					
Kraftwerk					1					
Kompostieranlage					1					

zusätzliche Flächennutzung aus Einzelgebäuden der Shape-Files:

After-school Centre	2			1				1		
Block of Flats								1		1
Culture Building								2		
Detached House								1		1
Hotel								2		
Industrial Building		1	1		1		1			
Infant School	2									
Office- and Commerce Building		1	1			2	2			
Public Use Building		1	1							
School				3						
Sport Hall								3		

C 2: Einbeziehung verschiedener “Points of Interest”

Tabelle 6 Typen der Points of Interest und ihre Einbeziehung in die Auswertung

Point of Interest-Typ	Mögliche Zwecke									
	Pick up/drop off	Private Business	Work related Business	School	Work	Shopping, Daily	Shopping, Long Term	Leisure	Other	Home
Hotel	1		2					3		
Youth Hostel	1		2					3		
Camping	1									
Gas station		2	2							
Car retail/ repair		2	2				2			
Car rental		2	2							
Car care		2	2							
Car towing		2	2							
Traffic school				5						
Carpenter		2	3				1			
Glazier		2	3				1			
Boat retailer		3	2				1			
Estate agent		4	1							
Bank		3	3							
Paint shop		3	3				2			
Locksmith		3	3							
Furniture store							5			
Electronics store							5			
Solarium								5		
Gardener		2	2				5			
Auctioneer		2					1	1		
Badminton								5		
Pool								5		
Cinema								5		
Bowling								5		
Shoppingcenter						5	5	1		
Golf store							5			
Golf court								5		
Spelombud								5		

Theater								5		
Catering			2					4		
Travel agency		5	2							
Restaurant			2					5		
Pizza Restaurant			1					5		
Pharmacy							5			
Church								5		
Police Station		3	3							
Educational association				5						
Photo retailer							5			
Newspaper store						5				
Tobacco store						5				
Delivery		2	2							
Bus company		2	2							
Airport	1	2	2							
Shipping company		2	2							
Taxi		2	2							
Transport agency		2	2							
Road carrier		2	2							
Bookstores							5			
Museums				3				5		
Barbershop		5								
Outdoor bath								5		
Indoor bath								5		
Cellphone retailer							5			
Hospital								5		
Shopping - Food						5	2			

Anhang D: Quellcodes der Makros

Zus.aml

```
/* Zentral-AML, die alle Prozesse aufruft
&arg nr cov
&ty zus.aml laeuft - HAUPTPROGRAMM
&ty andere routinen werden von hier aus aufgerufen.....

&r short_trips %nr% %cov%
&r wochenende %cov%

&r clus %nr% %cov%w
&r clus-analyse %nr% %cov%w 1 /* 1 = w
&r clus %nr% %cov%sa
&r clus-analyse %nr% %cov%sa 2 /* 2 = sa
&r clus %nr% %cov%so
&r clus-analyse %nr% %cov%so 3 /* 3 = so

&return
```

Short_trips.aml

```
/* aml zur Ermittlung der Stopps mit zu kurzen Weg/ Export in File: triphist_carNR.txt
&arg nr cov
&if [exists %cov%_backup -cover] &then
    &do
        &ty backup existiert bereits
        kill %cov% all
        copy %cov%_backup %cov%
    &end
    &else copy %cov% %cov%_backup
&if [exists triphist_car%nr%.txt -file] &then
    &do
        &if [delete triphist_car%nr%.txt -file] = 0 &then &ty triphist_car%nr%.txt deleted!!
        &else
            &do
                &ty problem!!
                &pause
            &end
        &end
    &else &ty triphist_car%nr%.txt existiert nicht.
&ty -----
&ty ... very short trips eliminieren ...
&ty -----
&sv shorttrip = 1.30 /* [response 'Minuten Aktivitätendauer, die von Berechn. ausgeschlossen werden sollen']
&sv shorttrip2 = 30 /* [response 'Sek. Reisezeit, die von Berechn. ausgeschlossen werden sollen (max. 60)']
&sv fileunit = [OPEN triphist_car%nr%.txt openstatus -APPEND]
&if %openstatus% ne 0 &then &ty ERROR beim OEFFNEN!!!!!! fehlernr = %openstatus%
ae
    edit %cov% point
    select ACT DUR INT <= %shorttrip% * 100 OR ACT DUR INT >= 240000
    aselect TotalkortID <= %shorttrip2% OR TotalkortID >= 3600
    aselect Koerningstracka >= 25
    aselect ( Koerningstracka / ( totalkortid + 1 ) ) >= ( 1 / 72 )
    aselect laenk_ID2 in {2132, 2128, 3600, 2551, 2550, 3604, 3535, 3602, 3608, 3534, 2121, 2122}

    &sv number_sel = [show number select]
    &sv klappts = [WRITE %fileunit% %number_sel%]
    &if %klappts% = 0 &then &ty erfolgreich eingetragen
    &else &ty problem!!!
    &if %number_sel% > 0 &then
        &do
            delete
            q
            y
            y
        &end
    &else
        q
    &sv klappts = [CLOSE -ALL]
    &if %klappts% ne 0 &then &ty files NOT closed succesfully
    &else &ty files closed succesfully
&return
```

Wochenende.aml

```
&arg cov
&ty wochenende.aml laeuft
&ty Stopp-Positionen werden nach ihrem Wochentag sortiert,
&ty d.h. mo-fr = Woche (w), Samstag (sa) und Sonntag (so)
&ty -----
&if [exists %cov%w -cover] &then kill %cov%w all
&if [exists %cov%s -cover] &then kill %cov%s all

ap
    reselect %cov% points d_o_w < 6
    writeselect %cov% week.sel
    nselect %cov% points
    reselect %cov% points d_o_w = 6
    writeselect %cov% sa.sel
```

```

nselect %cov% points
reselect %cov% points d_o_w = 7
writeselect %cov%_so.sel
q

reselect %cov% %cov%w point %cov%_week.sel
reselect %cov% %cov%sa point %cov%_sa.sel
reselect %cov% %cov%so point %cov%_so.sel

&if [delete %cov%_week.sel -File] = 0 &then &ty %cov%_week.sel deleted!!
&if [delete %cov%_week.sex -File] = 0 &then &ty %cov%_week.sex deleted!!
&if [delete %cov%_sa.sel -File] = 0 &then &ty %cov%_sa.sel deleted!!
&if [delete %cov%_sa.sex -File] = 0 &then &ty %cov%_sa.sex deleted!!
&if [delete %cov%_so.sel -File] = 0 &then &ty %cov%_so.sel deleted!!
&if [delete %cov%_so.sex -File] = 0 &then &ty %cov%_so.sex deleted!!

&return

```

Clus.aml

```

/* diese AML vereint die wg. Parkplatz/nichtParkplatz getrennten Coverages wieder zu einer
&args nr cov
&ty clus aml laeuft
&ty organisiert die Clusterung.....
&ty -----

&r clus-pp %cov%

&if [exists %cov%-cl -cover] &then kill %cov%-cl all
&if [exists %cov%-out -cover] &then kill %cov%-out all

&sv i = 2
&do &until %i% = 21
    &if [exists pp_%i% -cover] &then &r cluster %nr% pp_%i%
    &sv i = %i% + 1
&end

&r cluster %nr% %cov%

Append %cov%-cl point
    %cov%-cl
    &if [exists pp_2-cl -cover] &then pp_2-cl
    &if [exists pp_3-cl -cover] &then pp_3-cl
    &if [exists pp_4-cl -cover] &then pp_4-cl
    &if [exists pp_5-cl -cover] &then pp_5-cl
    &if [exists pp_6-cl -cover] &then pp_6-cl
    &if [exists pp_7-cl -cover] &then pp_7-cl
    &if [exists pp_8-cl -cover] &then pp_8-cl
    &if [exists pp_9-cl -cover] &then pp_9-cl
    &if [exists pp_10-cl -cover] &then pp_10-cl
    &if [exists pp_11-cl -cover] &then pp_11-cl
    &if [exists pp_12-cl -cover] &then pp_12-cl
    &if [exists pp_13-cl -cover] &then pp_13-cl
    &if [exists pp_14-cl -cover] &then pp_14-cl
    &if [exists pp_15-cl -cover] &then pp_15-cl
    &if [exists pp_16-cl -cover] &then pp_16-cl
    &if [exists pp_17-cl -cover] &then pp_17-cl
    &if [exists pp_18-cl -cover] &then pp_18-cl
    &if [exists pp_19-cl -cover] &then pp_19-cl
    &if [exists pp_20-cl -cover] &then pp_20-cl
end

Append %cov%-out point
    %cov%-out
    &if [exists pp_2-out -cover] &then pp_2-out
    &if [exists pp_3-out -cover] &then pp_3-out
    &if [exists pp_4-out -cover] &then pp_4-out
    &if [exists pp_5-out -cover] &then pp_5-out
    &if [exists pp_6-out -cover] &then pp_6-out
    &if [exists pp_7-out -cover] &then pp_7-out
    &if [exists pp_8-out -cover] &then pp_8-out
    &if [exists pp_9-out -cover] &then pp_9-out
    &if [exists pp_10-out -cover] &then pp_10-out
    &if [exists pp_11-out -cover] &then pp_11-out
    &if [exists pp_12-out -cover] &then pp_12-out
    &if [exists pp_13-out -cover] &then pp_13-out
    &if [exists pp_14-out -cover] &then pp_14-out
    &if [exists pp_15-out -cover] &then pp_15-out
    &if [exists pp_16-out -cover] &then pp_16-out
    &if [exists pp_17-out -cover] &then pp_17-out
    &if [exists pp_18-out -cover] &then pp_18-out
    &if [exists pp_19-out -cover] &then pp_19-out
    &if [exists pp_20-out -cover] &then pp_20-out
end

&sv i = 2
&do &until %i% = 21
    &if [exists pp_%i% -cover] &then kill pp_%i% all
    &if [exists pp_%i%-cl -cover] &then kill pp_%i%-cl all
    &if [exists pp_%i%-out -cover] &then kill pp_%i%-out all
    &sv i = %i% + 1
&end

&if [exists %cov%-cover] &then kill %cov% all
&if [exists %cov%-cl -cover] &then kill %cov%-cl all
&if [exists %cov%-out -cover] &then kill %cov%-out all

&return

```

Clus-pp.aml

```

/* diese AML sortiert vor der Clusterung nach Punkten auf Parkplätzen

```

```

/* und anderen, um die "Parkplatzpunkte" gesondert zu behandeln
&args cov
&ty clus-pp.aml laeuft
&ty Aufenthaltsorte werden nach ihrer Lage auf oder neben Parkplaetzen sortiert.
&ty -----
&if [exists %cov%pp -cover] &then kill %cov%pp all
&if [exists %cov%s -cover] &then kill %cov%s all

&sv i = 2
ap
&do &until %i% = 21
&if [exists pp2_%i% -cover] &then arc kill pp2_%i% all
reselect pp2 polys pp2# = %i%
writeselect pp2_%i%.sel
arc reselect pp2 pp2_%i% poly pp2_%i%.sel
nselect pp2 polys
&if [delete pp2_%i%.sel -FILE] = 0 &then
&ty pp2_%i%.sel deleted!!
&if [delete pp2_%i%.sex -FILE] = 0 &then
&ty pp2_%i%.sex deleted!!
&sv i = %i% + 1
&end
q
ae
ec %cov% point
&sv i = 2
&do &until %i% = 21
arcplot
clearselect
&ty i = %i%
&if [exists pp_%i% -cover] &then arc kill pp_%i% all
RESELECT %cov% points OVERLAP pp2_%i% polys
q
selectget
&sv nr = [SHOW NUMBER SELECT]

&if %nr% > 0 &then &do
arcplot
ClearSELECT
RESELECT %cov% points OVERLAP pp2_%i% polys
writeselect pp_%i%.sel
arc reselect %cov% pp_%i% point pp_%i%.sel point
&if [delete pp_%i%.sel -FILE] = 0 &then
&ty pp_%i%.sel deleted!!
&if [delete pp_%i%.sex -FILE] = 0 &then
&ty pp_%i%.sex deleted!!
ClearSELECT
q
&end
&sv i = %i% + 1
&end

arcplot
clearselect
RESELECT %cov% points OVERLAP pp2 polys
NSELECT %cov% points
writeselect %cov%.sel
arc reselect %cov% %cov%.point %cov%.sel point
&if [delete %cov%.sel -FILE] = 0 &then
&ty %cov%.sel deleted!!
&if [delete %cov%.sex -FILE] = 0 &then
&ty %cov%.sex deleted!!
q
&sv i = 2
&do &until %i% = 21
&if [exists pp2_%i% -cover] &then kill pp2_%i% all
&sv i = %i% + 1
&end
&return

```

Cluster.aml

```

/* diese AML clustert die diversen Aufenthaltspunkte jeweils eines Autofahrers
&args nr inputcover
&ty cluster.aml laeuft
&ty
&ty Aufenthaltsorte werden geclustert.....
&ty Output sind carNR-out und carNR-clus mit Verbindungs-
&ty item "cl-id"

&call vorbereitung
&call analyse
&call outputcover
&call IST_zentrum
&call fehlerkorrr
&call IST_zentrum
&call neue_schwerp
&call Output_allepunkte
&call aufräumen
&return

/* ROUTINE VARIABLENVEREINBAREN & AUFRÄUMEN
/* -----
&routine vorbereitung
&ty cluster.aml:
&ty jetzt vorbereitung...
&SV outputcover = %inputcover%-cl /* [response 'Outputcover-Name fuer Cluster-Zentren (bitte mit [m]-Angabe)']
&SV outputpoints = %inputcover%-out /* [response 'Outputcover-Name fuer Orig-Punkte']

&if [exists ptall -cover] &then kill ptall all
copy %inputcover% ptall
&if [exists %outputcover% -cover] &then kill %outputcover% all
&if [exists %outputpoints% -cover] &then kill %outputpoints% all
&if [exists %outputcover%.anz -info] &then killinfo %outputcover%.anz
&if [exists %outputcover%_write.anz -info] &then killinfo %outputcover%_write.anz

```

```

&ty      ... Distanztabelle berechnen ...
&ty      ... #a/#b in bezug zu my_id stellen!!! ...
&sv dist = ptall.dist      /* info-file mit berechneten Dist. Zw. allen Punkten der inputcoverage in bestimmten radius %d%
&if [exists %dist% -info] &then killinfo %dist%
&sv d = 200
pointdistance ptall ptall %dist% %d%

relate add my_A ptall.pat INFO ptall#a ptall# linear rw
relate add my_B ptall.pat INFO ptall#b ptall# linear rw

tables
  additem %dist% a_my_id 4 10 i
  additem %dist% b_my_id 4 10 i
  select %dist%
  calculate a_my_id = my_a//my_id
  calculate b_my_id = my_b//my_id
  calculate ptall#a = a_my_id
  calculate ptall#b = b_my_id
  dropitem %dist% a_my_id b_my_id
q

&ty      ... Tabelle fuer Statistik anlegen ...
&ty d.h. welcher Punkt hat wieviele Nachbarn in welcher
&ty durchschnittlichen Entfernung?
  tables
    define %outputcover%.anz
      cluster-id 6 6 i      /* punkt-id des clusterpkts
      anz 6 6 i            /* anzahl punkte in Distanzbereich %d%
      dd 4 10 f 5          /* Distanzenberechnung
      ind 4 10 i           /* Index für bereits stellv. Pkt oder nicht
      clus 4 10 i          /* Index für "Clusterzentrum ja o. nein"
      cl-ID 4 10 i         /* ID des Clusterzentrums, zu dem der Punkt gehört
    ~
  q

&return

/* ROUTINE ANALYSE
/* -----
/* in dieser Routine werden die Distanzen aus dem oben berechneten File ausgewertet.
/* Dabei werden die Anzahl der Nachbarn im oben bestimmten Radius und die durchschnittliche
/* Entfernung zu ihnen ermittelt.
&routine analyse
&ty      cluster.aml:
&ty      jetzt ANALYSE ...
/* von welchem Punkt aus welche durchschnittliche Entfernung zu wievielen Nachbarpunkten in %d%m Radius??

&sv curnr = - 999
&if [exists %dist%.stat -info] &then killinfo %dist%.stat
statistics %dist% %dist%.stat ptall#A
  sum distance
end

tables
  additem %dist%.stat dd 8 18 F 6
  additem %dist%.stat freq2 4 5 B
  select %dist%.stat
  reselect FREQUENCY > 1
  calculate freq2 = FREQUENCY - 1
  calculate dd = SUM-DISTANCE / FREQ2
  nselect
  calculate dd = SUM-DISTANCE
q

cursor STAT declare %dist%.stat info ro      /* cursor auf Statistiktabelle mit Distanzwerten und Häufigkeiten
cursor STAT open
cursor cur3 declare %outputcover%.anz info rw      /* Cursor auf Statistiktabelle der Outputcoverage,
/* die mit Werten aus Distanzberechnung gefüllt wird

cursor cur3 open
&do &while %:STAT.aml$next%
  cursor cur3 insert
  &sv :cur3.cluster-id = %:STAT.ptall#A%
  &sv :cur3.anz = %:STAT.FREQUENCY%
  &sv :cur3.dd = %:STAT.dd%
  cursor STAT next
&end

cursor cur3 close
cursor cur3 remove
cursor STAT close
cursor STAT remove

&ty      %outputcover%.anz angelegt!!!
&ty      (ptall#a = cluster-id, daraus spaeter cl-id)
&ty      (frequency = anz)
&ty      (dd = dd)

&return

/* ROUTINE CLUSTER-COVERAGE ERZEUGEN
/* -----
/* in dieser Routine werden überall dort Punkte in einer neuen Cluster-Coverage erzeugt, an denen
/* die Clusterzentren viele Nachbarn haben und die in noch keinem anderen Cluster vertreten sind
&routine outputcover
&ty      jetzt Cluster-Punkte in coverage erzeugen
/* sortieren nach
/* a) Anzahl der Nachbarpunkte und
/* b) durchschnittlicher Entfernung zu diesen
/* so kann %...%.anz als look-up für mgl. Clusterzentren dienen
/* und wird dann der Reihe nach "abgearbeitet"/"überprüft"

tables
  select %outputcover%.anz
  sort anz (d) dd
q

cursor cur3 declare %outputcover%.anz info rw
cursor cur3 open
&sv i = 1
&do &while ( %:cur3.aml$next% )
  &sv cursor_nr = %:cur3.$recno%
  &sv abbr = %:cur3.ind%
  &ty abbr = %abbr%

```

```

if %abbr% = 0 &then
  &do
    /* Zentrum bearbeiten, d.h. Punkt in outputcover erzeugen
    &sv akt_id = %:cur3.cluster-id%
    &ty akt_id = %akt_id%
    cursor cur declare ptall points ro my_id = %akt_id%
    cursor cur open
      &sv x = %:cur.x-coord%
      &sv y = %:cur.y-coord%
      precision double
      generate %outputcover%
      points
      %akt_id% %x% %y%
    end
    quit
    &sv :cur3.ind = 999
    &sv :cur3.clus = 999
    &sv :cur3.cl-ID = %akt_id%
    &ty akt_id = %akt_id% und x = %x% und y = %y%
  cursor cur close
  cursor cur remove

  /* "betroffene Punkte" bearbeiten & items in %outputcover%.anz eintragen (cur3)
  /* d.h. ind + 1 für gehört schon zu cl-zentrum
  cursor cur2 declare %dist% info ro ptall#a = %akt_id%
  cursor cur2 open
    &do &while %:cur2.aml$next%
      cursor cur3 close
      cursor cur3 remove
      cursor cur3 declare %outputcover%.anz info rw cluster-id = %:cur2.ptall#b%
      cursor cur3 open
        &sv :cur3.ind = %:cur3.ind% + 1
        &sv :cur3.cl-ID = %akt_id%
      cursor cur3 close
      cursor cur3 remove
      cursor cur3 declare %outputcover%.anz info rw
      cursor cur3 open
      cursor cur2 next
    &end
  cursor cur2 close
  cursor cur2 remove
&end
&else
  &ty leider kein moegliches clusterzentrum!!!

  &sv i = %i% + 1
  &ty cursor_nr = %cursor_nr%
  &ty i = %i%
  &ty
  cursor cur3 %cursor_nr%
  cursor cur3 next
&end
cursor cur3 close
cursor cur3 remove

build %outputcover% point
if [exists %outputcover% write.anz -info] &then killinfo %outputcover% write.anz
copyinfo %outputcover%.anz %outputcover%_write.anz /* Korrektur wird in %...% write.anz geschrieben, um Arbeitsschritte
/* nachvollziehen zu können (Vergleich %...%.anz und %...%_write.anz)
/* und weil es eh' ein "look-up-files" gebraucht hätte

&return

/* ROUTINE IST-Zentrum oder NICHT ?!
/* an die Distanz-Tabelle wird Item angehängt, das angibt, ob ein Punkt Clusterzentrum ist oder nicht.
&routine IST_zentrum
&ty jetzt IST-ZENTRUM ODER NICHT?
&ty items aus ANALYSE an %dist% anhaengen...

tables
dropitem %dist% anz dd ind clus cl-ID
select %outputcover% write.anz
alter cluster-id ptall#a
~
~
~
q
joinitem %dist% %outputcover%_write.anz %dist% ptall#a
tables
select %outputcover% write.anz
alter ptall#a cluster-id
~
~
~
q
&return

/* ROUTINE Fehlerkorrektur
/* -----
/* Fehler, die durch die mehrfache Zuordnung eines Stopp-Punktes zu mehreren Clustern
/* gezählt wurden, werden eliminiert
&routine fehlerkorr
&ty jetzt Fehlerkorr

cursor DOPPELT declare %outputcover%.anz info ro ind > 1 and ind < 1000
cursor DOPPELT open /* cursor DOPPELT = auf allen Punkten, die in mehreren Clustern

&do &while %:DOPPELT.aml$next%
  &ty START 1. SCHLEIFE

  &sv wieviele = %:DOPPELT.ind% /* wieviele = Anzahl, in wievielen Clustern Punkt vertreten ist
  &ty wieviele = %wieviele%

  &sv akt = %:DOPPELT.$recno%
  &sv mindist = %d% + 1 /* Ausgangswert für nächste Nachbarn finden, dabei Start bei d+1, damit für alle
  /* möglichen Distanzen "d" genügend groß
  /* wichtig: bei jedem schleifendurchgang neu setzten!

  if [exists %dist% copy -info] &then killinfo %dist%_copy
  copyinfo %dist% %dist%_copy /* Kopie der Distanzen-Tabelle, damit Löschen für jeden Schleifendurchgang möglich

  cursor FINDMIN declare %dist%_copy info rw ptall#B = %:DOPPELT.cluster-ID% and ptall#a NE ~
  %:DOPPELT.cluster-ID% and clus = 999
  cursor FINDMIN open /* cursor FINDMIN auf alle Punktpaare in dist-File, die betreffenden Punkt als ...#b
  /* haben, damit dann klar werden kann, ob die ...#a cl-zentren oder nicht und zu

```

```

/* Schleife, in der korrr. Cl-Zentrum ermittelt u. anschließend gelöscht wird
/* HIER MUSS AUCH KORR-Cl-Zentrum-Nummer ermittelt & in %...%_write.anz bei cl-ID eingetragen werden
&do &while %:FINDMIN.aml$next%
    &ty START 2. SCHLEIFE
    &ty
    &ty dist = %:FINDMIN.distance%
    &ty if %:FINDMIN.distance% < %mindist% &then
        &do
            &ty IF-BEDINGUNG/2.SCHLEIFE
            &ty
            &sv mindist = %:FINDMIN.distance%          /* min. Distanz, die für Cl-Zentrum spricht
            &sv central = %:FINDMIN.ptall#a%            /* zugehöriges Cluster-Zentrum
            &sv recno = %:FINDMIN.$recno%                /* zugehörige rec-no
        &end
    &ty cursor FINDMIN next
&end
cursor FINDMIN %recno%
cursor FINDMIN delete
cursor FINDMIN first

cursor WRITE declare %outputcover%_write.anz info rw $recno = %akt%
cursor WRITE open
&sv :WRITE.cl-ID = %central%
cursor WRITE close
cursor WRITE remove

/* Schleife, in der "falsche" Cl-Zentren ermittelt u. anschließend gelöscht werden
&sv i = 2
&do &while %i% <= %wieviele%
    &ty START 3. SCHLEIFE
    &ty
    &sv mindist = %d% + 1
    &ty i = %i% und wieviele = %wieviele%
    &ty cursor FINDMIN first
    &do &while %:FINDMIN.aml$next%          /* Schleife, in der jeweils nächstes "falsches" Zentrum ermittelt wird
        &ty START 3. SCHLEIFE/INNEN
        &ty
        &ty if %:FINDMIN.distance% < %mindist% &then
            &do
                &ty IF-BEDINGUNG/3.SCHLEIFE/INNEN
                &ty
                &sv mindist = %:FINDMIN.distance%          /* min. Distanz, die für nächstes Zentrum spricht
                &sv central = %:FINDMIN.ptall#a%            /* zugehöriges Cluster-Zentrum
                &sv recno = %:FINDMIN.$recno%                /* zugehörige rec-no
            &end
        &ty cursor FINDMIN next
    &end

    cursor FINDMIN %recno%

    cursor WRITE declare %outputcover%_write.anz info rw cluster-id = %central%
    cursor WRITE open
    &sv :WRITE.anz = %:WRITE.anz% - 1
    &ty :WRITE.anz = %:WRITE.anz% (- 1)
    cursor WRITE close
    cursor WRITE remove

    cursor WRITE declare %outputcover%_write.anz info rw $recno = %akt%
    cursor WRITE open

    &sv :WRITE.ind = %:WRITE.ind% - 1
    cursor WRITE close
    cursor WRITE remove

    cursor FINDMIN delete          /* entfernen des entspr. records für nächsten durchlauf 3.schleife
    /* "-1" in %outputcover%.anz - INFO-File

    &sv i = %i% + 1
&end

cursor FINDMIN close
cursor FINDMIN remove
cursor DOPPELT %akt%
cursor DOPPELT next
&end

cursor DOPPELT close
cursor DOPPELT remove

&if [exists %dist%_copy -info] &then killinfo %dist%_copy

&return

/* ROUTINE neue Schwerpunkte
/* -----
/* nach Klärung der Zugehörigkeiten (welche Pt zu welchen Clustern) müssen bei großen
/* Veränderungen neue Clusterzentren in einem Clusterhaufen bestimmt werden
&routine neue.schwerp
&ty      Jetzt neue Schwerpunkte...
&if [exists ids.tp -info] &then killinfo ids.tp
tables
define ids.tp
ids
2
4
b
anz_alt
2
4
b
~
q

&ty -----TEIL CL-ZENTREN ERMITTELN-----
&ty -----
cursor ANZ declare %outputcover%.anz info ro clus = 999          /* cursor auf ursprüngl. Anz-File
cursor ANZ open
cursor ANZWRITE declare %outputcover%_write.anz info ro clus = 999 /* cursor auf bearb. Anz-File
cursor ANZWRITE open
/* vergleich der Anzahl pro ursprünglichem clusterzentrum mit der nach Bereinigung wenn, dann eintrag in
/* ids.tp, um diese betroffenen dann später abzuarbeiten

```

```

&do &while %:ANZ.aml$next%
  &if %:ANZ.anz% <> %:ANZWRITE.anz% &then
    &do
      &type sie sind nicht gleich %:ANZ.anz% <> %:ANZWRITE.anz%
      &if [calc 0.95 * %:ANZ.anz%] >= %:ANZWRITE.anz% &then
        &do
          &type !!10!!10!!10!!10!!10!!10!!10!!10!!10!!
          &type betroffenes zentrum [calc 0.95 * %:ANZ.anz%] >= %:ANZWRITE.anz%
          &type !!10!!10!!10!!10!!10!!10!!10!!10!!10!!
          &sv cl-ID-10 = %:ANZWRITE.cl-ID%
          &sv anz_alt = %:ANZWRITE.anz%
          tables
            select ids.tp
            add
              %cl-ID-10%
              %anz_alt%
            ~
          q
        &end
      &end
    cursor ANZ next
    cursor ANZWRITE next
  &end

cursor ANZ close
cursor ANZ remove
cursor ANZWRITE close
cursor ANZWRITE remove

&ty -----TEIL betroffene Punkte ERMITTELN-----
&ty -----
tables
  dropitem %dist% count dd2 dd2-anz
  additem %dist% count 2 3 i /* für - a&b betroffen??
  additem %dist% dd2 4 10 f 5 /* neuer Durchschnitt
  additem %dist% dd2-anz 3 3 i /* für neu max-Nachbarn
  ~
q

cursor IDS declare ids.tp info ro
cursor IDS open
&do &while %:IDS.aml$next%
  &if %:IDS.anz_alt% > 1 &then
    &do
      cursor ANZWRITE declare %outputcover%_write.anz info ro cl-ID = %:IDS.ids%
      cursor ANZWRITE open /* alle fraglichen Clusterzentren abfragen
      &sv name = [value :IDS.ids]
      &if [exists tp$name%.tp -info] &then killinfo tp$name%.tp
      tables /* Tabelle für betroffene ID-s anlegen, wobei betroffen = in Cluster enthalten, das Änderung > 10 % hatte!!
        define tp$name%.tp
          id
          2
          4
          b
          ~
        dropitem %dist% count
        additem %dist% count 2 3 i /* für - a&b betroffen??
      q
      &sv i = 0
      &do &while %:ANZWRITE.aml$next%
        &sv wert = [value :ANZWRITE.cluster-id]
        tables /* Tabelle mit betroffenen ID-s füllen
          select tp$name%.tp
          add
            %wert%
          ~
        q
        &sv i = %i% + 1 /* Index für Anzahl der betroffenen Punkte
        cursor ANZWRITE next
      &end

      &ty i = %i%

      &sv j = 0
      &do &until %j% = %i%
        &sv j = %j% + 1
        &ty j = %j%
        cursor TP declare tp$name%.tp info ro $recno = %j%
        cursor TP open
        &sv id%j% = [value :TP.id]
        &ty id%j% = [value id%j%]
        tables
          select %dist%
          reselect ptall#a = [value id%j%] XOR ptall#b = [value id%j%]
          calculate count = count + 1
        q
        cursor TP close
        cursor TP remove
      &end

      &ty kleinste DURCHSCHNITTliche Entfernung BERECHNEN!!!
      &ty -----
      &sv j = 0
      &do &until %j% = %i%
        &sv j = %j% + 1
        &ty j = %j%
        &ty id%j% = [value id%j%]

        cursor dd2 declare %dist% info rw count > 1 and ptall#a = [value id%j%]
        cursor dd2 open
        &sv dd2 = 0
        &sv dd2-anz = 0
        &do &while %:dd2.aml$next%
          &sv dd2 = %dd2% + %:dd2.distance%
          &sv dd2-anz = %dd2-anz% + 1
          &sv :dd2.dd2 = 444
          cursor dd2 next
        &end
        &ty dd2-anz = %dd2-anz%
        &ty dd2 = %dd2%

```



```

        &if %dd2-anz% = 0 &then &sv dd2-anz = 1
        &sv dd2 = %dd2% / %dd2-anz%
        &ty dd2 = %dd2%
    cursor dd2 close
    cursor dd2 remove
    &ty DURCHSCHNITTliche Entfernungen in %dist% eintragen %j%
    tables
        select %dist%
        reselect ptall#a = [value id%j%] and ptall#b = [value id%j%]
        calculate dd2 = %dd2%
        calculate dd2-anz = %dd2-anz%
    q
&end

&if [exists tmp.stat -info] &then killinfo tmp.stat
&if [exists tmp2.stat -info] &then killinfo tmp2.stat

tables
copy %dist% tmp.stat
select tmp.stat
reselect cl-ID = %name% and ptall#a = ptall#b
sort dd2-anz (d) dd2
nselect
purge
y
q

cursor MIN2 declare tmp.stat info ro
cursor MIN2 open
&sv min_id = %MIN2.ptall#a%
&ty min_id = %min_id%
cursor MIN2 close
cursor MIN2 remove

&if %min_id% = %name% &then
    &do
        &ty CLUSTERTENTRUM BLEIBT!!!!mglkt.2!!!!
    &end
&else
    &do
        &sv anz_alt = %IDS.anz_alt% /* Anzahl der beteiligten Punkte des neuen Cl-Zentrums muss = ermittelter Anzahl
                                   /* der altens entsprechen, da dort bereits doppelte Zuordnungen eliminiert und nur
                                   /* noch mit dieser Auswahl weiter

        &ty anz_alt = %anz_alt%
        tables
            select %outputcover% write.anz
            reselect cl-ID = %name%
            calculate cl-ID = %min_id%
            select %outputcover% write.anz
            reselect cluster-ID = %name%
            calculate ind = 1
            calculate clus = 0
            select %outputcover% write.anz
            reselect cluster-ID = %min_id%
            calculate ind = 1000
            calculate clus = 999
            calculate anz = %anz_alt%
        q
        cursor ptorig declare ptall points ro my_id = %min_id%
        cursor ptorig open
        &sv x = %ptorig.x-coord%
        &sv y = %ptorig.y-coord%
        precision double
        generate %outputcover%
            points
            %min_id% %x% %y%
        end
        quit
        cursor ptorig close
        cursor ptorig remove

        build %outputcover% point

        ae
            edit %outputcover% point
            select %outputcover%-ID = %name%
            delete
        q
        y
        y
    &end
    cursor ANZWRITE close
    cursor ANZWRITE remove
&end
&else
    &do
        &ty CLUSTERTENTRUM BLEIBT!!!!mglkt1!!!!
        &ty weil anz_alt <= 1
    &end
    cursor IDS next
&end
cursor IDS close
cursor IDS remove

cursor IDS declare ids.tp info ro /* tp-TEMP-FILES wieder Löschen!!!!
cursor IDS open
&do &while %IDS.aml$next%
    &sv name = [value :IDS.ids]
    &if [exists tp$name%.tp -info] &then killinfo tp$name%.tp
    cursor IDS next
&end
cursor IDS close
cursor IDS remove

&if [exists tmp.stat -info] &then killinfo tmp.stat
&if [exists tmp2.stat -info] &then killinfo tmp2.stat
&if [exists ids.tp -info] &then killinfo ids.tp
&return

/* ROUTINE output-allePunkte

```

```

/* -----
/* Zugehörigkeit aller Punkten zu den einzelnen Zentren in *.pat eintragen (einmal in Cluster.pat und in OutputPoints.pat)
&routine Output_allepunkte
&ty      jetzt OUTput-allePunkte...

/* Umbenennen des Items "cluster-id", damit joinitem mit outputcover mit flexiblen namen möglich wird

tables
  dropitem %outputcover%.pat anz dd ind clus cl-ID
  select %outputcover%.write.anz
  alter cluster-id %outputcover%-id
  ~
  ~
q

joinitem %outputcover%.pat %outputcover%.write.anz %outputcover%.pat %outputcover%-id

tables
  select %outputcover%.write.anz
  alter %outputcover%-id my_id
  ~
  ~
q

&if [exists %outputpoints%.cover] &then
  dropitem %outputpoints%.pat %outputpoints%.pat anz dd ind clus cl-ID
  &else copy ptall %outputpoints%
joinitem %outputpoints%.pat %outputcover%.write.anz %outputpoints%.pat my_id

tables
  select %outputcover%.write.anz
  alter my_id cluster-id
  ~
  ~
q

&return

/* ROUTINE AufräumenII
/* -----
/* Überflüssige temporäre Coverages/Infofiles löschen

&routine aufräumen
&ty      jetzt Aufräumen... (hinterher)
&ty

&if [exists %outputcover%.write.anz -info] &then killinfo %outputcover%.write.anz
&if [exists ptall -cover] &then kill ptall all
&if [exists %dist%.stat -info] &then killinfo %dist%.stat
&return

```

Clus-analyse.aml

```

/* diese AML wertet 1. numerische Informationen zu den Clustern aus und ermittelt 2. wahrscheinliche Wegzwecke
&args nr cov dow
&ty      clus-analyse.aml laeuft
&ty      Aufenthaltscluster werden analysiert und ihr wahrschein-
&ty      licher Aufenthaltszweck ermittelt.....

&sv outputcover = %cov%-cl
&sv outputpoints = %cov%-out
&sv clzweck = %cov%-tmp          /* TMP-AIM-Coverage für Zuordnung der Zwecke
&sv clzweck2 = %cov%-aim        /* OUTPUT-Coverage mit zugeordneten Zwecken

  &call ITEMS_ANLEGEN1
  &call STATISTISCHE_AUSWERTUNG
  &call ITEMS_ANLEGEN2
  &call HOME                      /* dieser Teil benötigt die Coverage: addr
  &call POI                      /* dieser Teil benötigt die Coverage: POI und das File: poi_zwecke.lut
  &call FNP                      /* dieser Teil benötigt die Coverage: FN_zus und das File: fnp_zwecke_fein.lut
  &call FN_ZWECK                 /* dieser Teil benötigt die beiden Files: sozdem.lut und res_zwecke.lut
  &call TIMES
  &call Zweck_Ermitteln
  &call aufräumen
&return

/* -----
/*      ROUTINEN-TEIL
/* -----

/* Routine Items-anlegen: hier werden die Relate Items in der "out-coverage" für die Verwaltung der Uhrzeiten in
/* Stunden/Minuten/Sekunden eingeführt
/* -----
&routine ITEMS_ANLEGEN1
&ty ITEMS_ANLEGEN1 laeuft
/* d.h. Items in %outputpoints%.pat anlegen und ausfüllen für Analyse der Zeiten und in %outputcover%.pat Items anlegen, in
/* die Mittelwerte aus %outputpoints%.pat-Zeiten eingetragen werden.

tables
/* Zeiten in Minuten umrechnen
dropitem %outputpoints%.pat a_end_hh a_end_mm actdur_hh actdur_mm deptime_hh deptime_mm arptime_hh arptime_mm
select %outputpoints%.pat
  redefine
  35
  a_end_hh
  2
  3
  1
  37
  a_end_mm
  2
  3
  1

```

```

63
actdur_hh
2
3
1
1
65
actdur_mm
2
3
1
1
181
deptime_hh
2
3
1
1
183
deptime_mm
2
3
1
1
197
arrtime_hh
2
3
1
1
199
arrtime_mm
2
3
1
1
~

/* Items für Analyse/Zweckbestimmung in die covXXY-out.pat einführen
dropitem %outputpoints%.pat a_end2 actdur2 deptime2 arrtime2
additem %outputpoints%.pat a_end2 6 7 i
additem %outputpoints%.pat actdur2 6 7 i
additem %outputpoints%.pat deptime2 6 7 i
additem %outputpoints%.pat arrtime2 6 7 i

select %outputpoints%.pat
calculate a_end2 = a_end hh * 60 + a_end mm
calculate actdur2 = actdur hh * 60 + actdur mm
calculate deptime2 = deptime hh * 60 + deptime mm
calculate arrtime2 = arrtime hh * 60 + arrtime mm

/* Items für Analyse/Zweckbestimmung in die covXXY-cl.pat einführen
dropitem %outputcover%.pat mean_ARRTIME stdev_ARRTIME mean_ACT_DUR stdev_ACT_DUR
dropitem %outputcover%.pat med_ARRTIME med_ACT_DUR

additem %outputcover%.pat mean_ARRTIME 8 18 F 6
additem %outputcover%.pat stdev_ARRTIME 8 18 F 6
additem %outputcover%.pat med_ARRTIME 8 18 F 6

additem %outputcover%.pat mean_ACT_DUR 8 18 F 6
additem %outputcover%.pat stdev_ACT_DUR 8 18 F 6
additem %outputcover%.pat med_ACT_DUR 8 18 F 6
q
&return

/* Routine statistische Auswertung: hier werden für jedes Clusterzentrum durchschnittliche Ankunfts/Abfahrts/Wegezeiten und
/* Aktivitätendauern ermittelt und an die "Cl-PAT" angehängt
/* -----
&routine STATISTISCHE AUSWERTUNG
&ty STATISTISCHE_AUSWERTUNG laeuft

/* teil für mittelwert (mean) & standartabweichung
&if [exists %outputpoints%.stat -info] &then killinfo %outputpoints%.stat

statistics %outputpoints%.pat %outputpoints%.stat cl-id
mean ARRTIME2
STANDARDDEVIATION ARRTIME2
mean ACTDUR2
STANDARDDEVIATION ACTDUR2
END

relate add cl-stat %outputpoints%.stat INFO cl-id cl-id linear rw

tables
select %outputcover%.pat
calculate mean arrtime = cl-stat//MEAN-ARRTIME2
calculate stdev arrtime = cl-stat//STD-ARRTIME2
calculate mean ACT_DUR = cl-stat//MEAN-ACTDUR2
calculate stdev ACT_DUR = cl-stat//STD-ACTDUR2
q

&if [exists %outputpoints%.stat -info] &then killinfo %outputpoints%.stat

/* teil für median
&if [exists median.stat -info] &then killinfo median.stat

/* teil A) für arr-time
copyinfo %outputpoints%.pat median.stat
tables
select median.stat
sort cl-id arrtime2
q

&sv clu-nr = 0
ap
cursor arr declare median.stat info
cursor arr open
&do &while %:arr.aml$next%
&sv clu-nr = %:arr.cl-id%
&sv rec-nr = %:arr.$recno%
cursor arr close
&ty clu-nr = %clu-nr% rec-nr = %rec-nr%
reselect median.stat info cl-id = %clu-nr% /* alle auswählen, die in dieses Clusterzentrum gehören
cursor arr open
&sv cl-anz = %:arr.AML$NSEL% /* wie viele in diesem Zentrum??
&ty anzahl = %cl-anz%
&ty gerade oder ungerade ??? [mod %cl-anz% 2] /* Bestimmung ob gerade oder ungerade, d.h. wahrer oder Pseudomedian

```

```

cursor arr close

&sv rec-med = 0
&sv med-arr = 0
&sv trunc = %cl-anz% / 2
&sv rec-med = %rec-nr% + [TRUNC %trunc%]
/* für ungerade
&if [mod %cl-anz% 2] = 1 &then &do
    &ty ungerade !!!!!
    reselect median.stat info $recno = %rec-med% /* cursor auf Median Position
    cursor arr open
    &sv med-arr = %:arr.arrrtime2%
    &end
/* für gerade -> Pseudomedian berechnen
    &else &do
        &ty gerade !!!!!
        reselect median.stat info $recno = %rec-med% - 1 or $recno = %rec-med%
        cursor arr open
        &sv med-arr1 = %:arr.arrrtime2%
        cursor arr next
        &sv med-arr2 = %:arr.arrrtime2%
        &sv med-arr = ( %med-arr1% + %med-arr2% ) / 2
    &end

arc cursor med declare %outputcover%.pat info rw cl-id = %clu-nr%
arc cursor med open
&sv :med.med ARRTIME = %med-arr%
arc cursor med close
arc cursor med remove

cursor arr close
clearselect
reselect median.stat info cl-id > %clu-nr%
cursor arr open

&end
cursor arr remove

q

/* teil B) für act-dur
tables
select median.stat
sort cl-id actdur2
q

&sv clu-nr = 0
ap
cursor dur declare median.stat info
cursor dur open

&do &while %:dur.aml$next%

    &sv clu-nr = %:dur.cl-id%
    &sv rec-nr = %:dur.$recno%
    cursor dur close
    &ty clu-nr = %clu-nr% rec-nr = %rec-nr%
    reselect median.stat info cl-id = %clu-nr% /* alle auswählen, die in dieses Clusterzentrum gehören
    cursor dur open
    &sv cl-anz = %:dur.AML$NSEL% /* wie viele in diesem Zentrum??
    &ty anzahl = %cl-anz%
    &ty gerade oder ungerade ??? [mod %cl-anz% 2] /* Bestimmung ob gerade oder ungerade, d.h. wahrer oder Pseudomedian
    cursor dur close

    &sv rec-med = 0
    &sv med-dur = 0
    &sv trunc = %cl-anz% / 2
    &sv rec-med = %rec-nr% + [TRUNC %trunc%]
    /* für ungerade
    &if [mod %cl-anz% 2] = 1 &then &do
        &ty ungerade !!!!!
        reselect median.stat info $recno = %rec-med% /* cursor auf Median Position
        cursor dur open
        &sv med-dur = %:dur.actdur2%
        &end
    /* für gerade -> Pseudomedian berechnen
        &else &do
            &ty gerade !!!!!
            reselect median.stat info $recno = %rec-med% - 1 or $recno = %rec-med%
            cursor dur open
            &sv med-dur1 = %:dur.actdur2%
            cursor dur next
            &sv med-dur2 = %:dur.actdur2%
            &sv med-dur = ( %med-dur1% + %med-dur2% ) / 2
        &end

    arc cursor med declare %outputcover%.pat info rw cl-id = %clu-nr%
    arc cursor med open
    &sv :med.med ACT_DUR = %med-dur%
    arc cursor med close
    arc cursor med remove

    cursor dur close
    clearselect
    reselect median.stat info cl-id > %clu-nr%
    cursor dur open

&end
cursor dur remove

q
&if [exists median.stat -info] &then killinfo median.stat
&return

&routine ITEMS ANLEGEN2
&ty ITEMS ANLEGEN2 laeuft
/* nach S&atistischer Auswertung der Zeiten pro cluster in "statistische Auswertung"
/* Items für Analyse/Zweckbestimmung in die covXXY-zweck.pat einführen
&if [exists %clzweck% -cover] &then kill %clzweck% all
copy %outputcover% %clzweck%

dropitem %clzweck%.pat %clzweck%.pat home pick_drop priv_bus wrel_bus school work shop_daily-
shop_lterm leisure other

```

```

dropitem %clzweck%.pat %clzweck%.pat fn_aim fn_kat
dropitem %clzweck%.pat %clzweck%.pat time_aim time_kat aim aim_s anz_index

additem %clzweck%.pat %clzweck%.pat home 2 4 b
additem %clzweck%.pat %clzweck%.pat pick_drop 2 4 b
additem %clzweck%.pat %clzweck%.pat priv_bus 2 4 b
additem %clzweck%.pat %clzweck%.pat wrel_bus 2 4 b
additem %clzweck%.pat %clzweck%.pat school 2 4 b
additem %clzweck%.pat %clzweck%.pat work 2 4 b
additem %clzweck%.pat %clzweck%.pat shop_daily 2 4 b
additem %clzweck%.pat %clzweck%.pat shop_lterm 2 4 b
additem %clzweck%.pat %clzweck%.pat leisure 2 4 b
additem %clzweck%.pat %clzweck%.pat other 2 4 b

additem %clzweck%.pat %clzweck%.pat fn_aim 2 4 b
additem %clzweck%.pat %clzweck%.pat fn_kat 2 4 b

additem %clzweck%.pat %clzweck%.pat time_aim 2 4 b
additem %clzweck%.pat %clzweck%.pat time_kat 2 4 b
additem %clzweck%.pat %clzweck%.pat aim 2 4 b
additem %clzweck%.pat %clzweck%.pat aim_s 2 4 b
additem %clzweck%.pat %clzweck%.pat anz_index 2 4 b

if [exists anz.tmp -info] &then killinfo anz.tmp
copyinfo %clzweck%.pat anz.tmp
tables
select anz.tmp
sort anz (D)
dropitem %clzweck%.pat anz_index
additem %clzweck%.pat anz_index 2 4 b
calculate anz_index = $rcno
relate add ANZ anz.tmp info cl-id cl-id linear ro
select %clzweck%.pat
calculate anz_index = ANZ//anz_index
q
if [exists anz.tmp -info] &then killinfo anz.tmp
&return

/* Routine Home: hier wird für diejenigen Testfahrer, für die die Wohnadresse
/* als Punkt in "Addr-Coverage" bekannt ist, die Aktivität "home" festgelegt
&routine HOME
&ty HOME laeuft

if [exists cltmp -cover] &then kill cltmp all
copy %clzweck% cltmp
if [exists addr%nr% -cover] &then kill addr%nr% all

ae
ec addr point
arcplot
reselect addr points hhnr = %nr%
q
selectget
&sv anz = [SHOW NUMBER SELECT]
&ty anz = %anz%
if %anz% > 0 &then &do
arcplot
reselect addr points hhnr = %nr%
writeselect hhnr.sel
clearselect
q
arc reselect addr addr%nr% point hhnr.sel
if [delete hhnr.sel -file] = 0 &then &ty hhnr.sel geloescht
if [delete hhnr.sex -file] = 0 &then &ty hhnr.sex geloescht
&end
q

if [exists addr%nr% -cover] &then
&do
if [exists home.dist -info] &then killinfo home.dist
pointdistance addr%nr% cltmp home.dist 200
if [exists home.dist -info] &then
&do
list home.dist
relate add home home.dist INFO cltmp# cltmp# linear rw
tables
select cltmp.pat
calculate home = home//distance
reselect home > 0
calculate home = 100
q
&end
&end
if [exists %clzweck% -cover] &then kill %clzweck% all
copy cltmp %clzweck%
if [exists home.dist -info] &then killinfo home.dist
if [exists addr%nr% -cover] &then kill addr%nr% all
if [exists cltmp -cover] &then kill cltmp all
&return

/* Routine POI: hier werden Aktivitäten aus der "Point-of-Interest-Coverage"
/* abgeleitet
&routine POI
&ty POI laeuft

if [exists poi.dist -info] &then killinfo poi.dist
if [exists cltmp -cover] &then kill cltmp all
copy %clzweck% cltmp

pointdistance cltmp poi poi.dist 300

if [exists poi.dist -info] &then &do
relate add POIPAT poi.pat info poi# poi# linear ro
/* relate von distanz-tabelle auf poi.pat, um type herauszubekommen
relate add POIZWECK poi_zwecke.lut info type typenr linear ro
/* relate von poi.pat auf poi_zwecke.lut, um punkte für zwecke herauszubekommen

&sv faktor = 1
cursor CL-ZWECK declare cltmp.pat info rw
cursor CL-ZWECK open
&do &while %:CL-ZWECK.aml$next%

```

```

&ty cl-zweck# = %:CL-ZWECK.cltmp##
cursor CLNR declare poi.dist info ro cltmp# = %:CL-ZWECK.cltmp#
cursor CLNR open
&do &while %:CLNR.aml$next%

    &if %:CLNR.distance% <= 50 &then &sv faktor = 1.5
    &if %:CLNR.distance% > 50 AND %:CLNR.distance% <= 100 &then &sv faktor = 1.0
    &if %:CLNR.distance% > 100 AND %:CLNR.distance% <= 200 &then &sv faktor = 0.7
    &if %:CLNR.distance% > 200 AND %:CLNR.distance% <= 300 &then &sv faktor = 0.4

    &sv %:CL-ZWECK.pick_drop = %:CL-ZWECK.pick_drop% + %faktor% * %:CLNR.poipat//poizweck//pick_drop%
    &sv %:CL-ZWECK.priv_bus = %:CL-ZWECK.priv_bus% + %faktor% * %:CLNR.poipat//poizweck//priv_bus%
    &sv %:CL-ZWECK.wrel_bus = %:CL-ZWECK.wrel_bus% + %faktor% * %:CLNR.poipat//poizweck//wrel_bus%
    &sv %:CL-ZWECK.school = %:CL-ZWECK.school% + %faktor% * %:CLNR.poipat//poizweck//school%
    &sv %:CL-ZWECK.work = %:CL-ZWECK.work% + %faktor% * %:CLNR.poipat//poizweck//work%
    &sv %:CL-ZWECK.shop_daily = %:CL-ZWECK.shop_daily% + %faktor% * %:CLNR.poipat//poizweck//shop_daily%
    &sv %:CL-ZWECK.shop_lterm = %:CL-ZWECK.shop_lterm% + %faktor% * %:CLNR.poipat//poizweck//shop_lterm%
    &sv %:CL-ZWECK.leisure = %:CL-ZWECK.leisure% + %faktor% * %:CLNR.poipat//poizweck//leisure%
    &sv %:CL-ZWECK.other = %:CL-ZWECK.other% + %faktor% * %:CLNR.poipat//poizweck//other%

    cursor CLNR next
&end
cursor CLNR close
cursor CLNR remove
cursor CL-ZWECK next
&end
cursor CL-ZWECK close
cursor CL-ZWECK remove
&if [exists %clzweck% -cover] &then kill %clzweck% all
copy cltmp %clzweck%
&if [exists cltmp -cover] &then kill cltmp all
&if [exists poi.dist -info] &then killinfo poi.dist
relate drop POIPAT POIZWECK
&end

&return

/* Routine FNP: hier werden Aktivitäten aus dem digitalisierten
/* Flächennutzungsplan(FNP) abgeleitet
&routine FNP
&ty FNP laeuft

&if [exists cltmp -cover] &then kill cltmp all
copy %clzweck% cltmp

&sv i = 0
cursor CL declare cltmp.pat info rw
cursor CL open
&do &while %:CL.aml$next%
    &sv i = %i% + 1
    &ty i = %i%
    &sv akt_id = %:CL.CL-ID% /* gerade bearbeitetes Clusterzentrum
    &ty akt_id = %akt_id%

    &ty Teil, in dem aktuelles Clusterzentrum gebuffert wird
    &ty -----
    ap
        &if [exists clbuf%i% -cover] &then arc kill clbuf%i% all
        &if [exists clbuf%i%b -cover] &then arc kill clbuf%i%b all
        clearselect
        reselect cltmp points CL-ID = %akt_id% /* Clusterzentrum wird aus clustercoverage ausgewählt
        writeselect cl_buf.sel
        arc reselect cltmp clbuf%i% point cl_buf.sel /* temp-coverage für dieses Zentrum
        &ty +++ Punkt ausgelesen !!!! +++

        &if [delete cl_buf.sel -File] = 0 &then &ty cl_buf.sel deleted!!
        &if [delete cl_buf.sex -File] = 0 &then &ty cl_buf.sex deleted!!

        arc buffer clbuf%i% clbuf%i%b # # 200 # point round /* Zentrum gebuffert
        &ty +++ Punkt gebuffert !!!! +++

    &ty Teil, in dem Flächeneininformation aus fn_zus verarbeitet wird
    &ty -----
    reselect fn_zus poly OVERLAP clbuf%i%b poly
    infofile fn_zus poly fn_zus_%i%.selected code init
q
tables
select fn_zus_%i%.selected
reselect Code = 0
purge
y
select fn_zus_%i%.selected
sort code
q

relate add FNPZWECKE fnp_zwecke_fein.lut info code typenr linear ro

cursor FNPRV declare fn_zus_%i%.selected info ro
cursor FNPRV open
&sv FN = 0 /* kontrollvariable, damit jeder benachbarte flächentyp nur je einmal in die berechnung eingeht
&do &while %:FNPRV.aml$next%
    &if %fn% NE %:FNPRV.FNPZWECKE//typenr% &then
        &do
            &sv %:CL.home = %:CL.home% + %:FNPRV.FNPZWECKE//home%
            &sv %:CL.pick_drop = %:CL.pick_drop% + %:FNPRV.FNPZWECKE//pick_drop%
            &sv %:CL.priv_bus = %:CL.priv_bus% + %:FNPRV.FNPZWECKE//priv_bus%
            &sv %:CL.wrel_bus = %:CL.wrel_bus% + %:FNPRV.FNPZWECKE//wrel_bus%
            &sv %:CL.school = %:CL.school% + %:FNPRV.FNPZWECKE//school%
            &sv %:CL.work = %:CL.work% + %:FNPRV.FNPZWECKE//work%
            &sv %:CL.shop_daily = %:CL.shop_daily% + %:FNPRV.FNPZWECKE//shop_daily%
            &sv %:CL.shop_lterm = %:CL.shop_lterm% + %:FNPRV.FNPZWECKE//shop_lterm%
            &sv %:CL.leisure = %:CL.leisure% + %:FNPRV.FNPZWECKE//leisure%
            &sv %:CL.other = %:CL.other% + %:FNPRV.FNPZWECKE//other%
        &end
        &sv FN = %:FNPRV.FNPZWECKE//typenr%
        cursor FNPRV next
    &end
cursor FNPRV close

```

```

cursor FNPRV remove

&ty Teil, in dem aufgeraeumt wird
relate drop $all

&if [exists clbuf%i% -cover] &then arc kill clbuf%i% all
&if [exists clbuf%i%b -cover] &then arc kill clbuf%i%b all
&if [exists fn_zus_%i%.selected -info] &then killinfo fn_zus_%i%.selected
cursor CL next
&end
cursor CL close
cursor CL remove

&if [exists %clzweck% -cover] &then kill %clzweck% all
copy cltmp %clzweck%
&if [exists cltmp -cover] &then kill cltmp all

&return

/* Routine FN_ZWECK hier werden wahrscheinliche Zwecke aus vorhergehender Analyse ermittelt
/* -----
&routine FN_ZWECK
&ty FN_ZWECK laeuft

&if [exists cltmp -cover] &then kill cltmp all
copy %clzweck% cltmp

cursor FN_ZWECK declare cltmp.pat info rw
cursor FN_ZWECK open
&do &while %:FN_ZWECK.aml$next%
&sv ho = %:FN_ZWECK.home%
&sv pd = %:FN_ZWECK.pick_drop%
&sv pb = %:FN_ZWECK.priv_bus%
&sv wb = %:FN_ZWECK.wrel_bus%
&sv sh = %:FN_ZWECK.schoOl%
&sv wo = %:FN_ZWECK.work%
&sv sd = %:FN_ZWECK.shop_daily%
&sv sl = %:FN_ZWECK.shop_lterm%
&sv ls = %:FN_ZWECK.leisure%

&sv max = [max %ho% [max %pd% [max %pb% [max %wb% [max %sh% [max %wo% [max %sd% [max %sl% %ls%]]]]]]]]
&ty max = %max%

&sv i = 0
&if %ho% = %max% &then
&do
&sv maxitem = 10
&sv i = %i% + 1
&end
&if %pd% = %max% &then
&do
&sv maxitem = 1
&sv i = %i% + 1
&end
&if %pb% = %max% &then
&do
&sv maxitem = 2
&sv i = %i% + 1
&end
&if %wb% = %max% &then
&do
&sv maxitem = 3
&sv i = %i% + 1
&end
&if %sh% = %max% &then
&do
&sv maxitem = 4
&sv i = %i% + 1
&end
&if %wo% = %max% &then
&do
&sv maxitem = 5
&sv i = %i% + 1
&end
&if %sd% = %max% &then
&do
&sv maxitem = 6
&sv i = %i% + 1
&end
&if %sl% = %max% &then
&do
&sv maxitem = 7
&sv i = %i% + 1
&end
&if %ls% = %max% &then
&do
&sv maxitem = 8
&sv i = %i% + 1
&end

&if %i% = 1 &then
&do
&sv :FN_ZWECK.fn_aim = %maxitem%
&sv :FN_ZWECK.fn_kat = 1
&end
&if %i% = 2 &then
&do
&sv :FN_ZWECK.fn_aim = %maxitem%
&sv :FN_ZWECK.fn_kat = 2
&end
cursor FN_ZWECK next
&end
cursor FN_ZWECK close
cursor FN_ZWECK remove

&if [exists %clzweck% -cover] &then kill %clzweck% all
copy cltmp %clzweck%
&if [exists cltmp -cover] &then kill cltmp all
&return

```

```

/* Routine TIMES: hier werden Aktivitäten aus den Aktivitätendauern und -startzeiten
/* in Zusammenhang mit den Soziodemographischen Merkmalen rekonstruiert
/* -----
&routine TIMES
&ty TIMES laeuft

&if [exists cltmp -cover] &then kill cltmp all
copy %clzweck% cltmp
&ty Bestimmung der festen Parameter für Zwecksuche in res_zwecke.lut

/* pv_avail muss hier erstmal = 3 sein, weil erstmal nur testfahrer selber, die auto haben
/* pv_avail = 2 dann später für "mit Führerschein aber kein Auto"
&sv pv_avail = 3

/* Bestimmung Wochentag -> durch wochenende.aml bereits festgelegt, deshalb hier als Var übergeben
/* wenn %dow% = 1 übergeben, dann wochentags
/* wenn %dow% = 2 übergeben, dann samstags
/* wenn %dow% = 3 übergeben, dann sonntags
&ty
&ty woche/wochenende?
&if %dow% = 1 &then &ty Aktivitaeten fanden wochentags statt
&if %dow% = 2 &then &ty Aktivitaeten fanden samstags statt
&if %dow% = 3 &then &ty Aktivitaeten fanden sonntags statt
&ty

/* Bestimmung soziodemograph. MM aus sozdem.lut, d.h. aus Angaben von Testpersonen
cursor SOZDEM declare sozdem.lut info ro fordonID = %nr%
cursor SOZDEM open
&if %:SOZDEM.status% = 5 &then
&do
&sv .status = 2
&ty der testfahrer ist ein RENTNER :- )
&if %:SOZDEM.sex% = 0 &then
&do
&sv sex = 2
&ty die RENTNERIN IST WEIBLICH
&end
&else
&do
&sv sex = 1
&ty der RENTNER IST MAENNLICH
&end
&end
&if %:SOZDEM.status% = 8 &then
&do
&sv .status = 3
&ty der testfahrer arbeitet VOLLZEIT :- )
&if %:SOZDEM.sex% = 0 &then
&do
&sv sex = 2
&ty die ARBEITNEHMERIN IST WEIBLICH
&end
&else
&do
&sv sex = 1
&ty der ARBEITNEHMER IST MAENNLICH
&end
&end
&if %:SOZDEM.status% NE 5 and %:SOZDEM.status% NE 8 &then
&do
&sv .status = 1
&ty der testfahrer ist weder RENTNER noch VOLLZEITbeschaeftigt :- )
&if %:SOZDEM.sex% = 0 &then
&do
&sv sex = 2
&ty die TESTPERSON IST WEIBLICH
&end
&else
&do
&sv sex = 1
&ty der TESTPERSON IST MAENNLICH
&end
&end

cursor SOZDEM close
cursor SOZDEM remove

&sv i = 0
&sv j = 0
&sv k = 0
&sv l = 0
cursor CL declare cltmp.pat info rw
cursor CL open
&do &while %:CL.aml$next%
&ty Bestimmung der variablen Parameter für Zwecksuche, d.h. Zeitkategorien
&ty fuer jedes Cluster getrennt
&sv match = 0
&sv actdur_cat = 0
&sv i = %i% + 1
&ty i = %i%
&sv akt_id = %:CL.CL-ID% /* gerade bearbeitetes Clusterzentrum
&ty akt_id = %akt id%
&ty anz = %:CL.ANZ%

&sv arrtime_cat = 1 + [truncate [calc %:CL.med_ARRTIME% / 60 ] ]
&ty arrtime_cat = %arrtime_cat%
&if %:CL.med_ACT_DUR% <= 60 &then
&do
&sv actdur_cat = 1 + [truncate [calc %:CL.med_ACT_DUR% / 15 ] ]
&ty actdur unter 1 Stunde
&end
&else
&do
&if %:CL.med_ACT_DUR% <= 240 &then
&do
&sv actdur_cat = 3 + [truncate [calc %:CL.med_ACT_DUR% / 30 ] ]
&ty actdur unter 4 Stunden
&end
&else

```



```

        &do
            &sv actdur_cat = 7 + [truncate [calc %:CL.med_ACT_DUR% / 60 ] ]
            &ty actdur laenger
        &end
    &end
    &ty actdur_cat = %actdur_cat%
    &ty

    &ty Bestimmung des moeglichen Zwecks aus oben bestimmten Parametern
    &ty ANGABEN PARAMETER TESTFAHRER:
    &ty sex = %sex% .status = %.status% pv_avail = %pv_avail% dow = %dow%
    &ty arrtime_cat = %arrtime_cat% actdur_cat = %actdur_cat%

    cursor RES declare res_zwecke.lut info ro sex = %sex% AND status = %.status% AND pv_avail = %pv_avail% ~
        AND d_o_w = %dow% AND a_start = %arrtime_cat% AND a_dur = %actdur_cat%
    cursor RES open
        &if %:RES.aml$next% &then
            &do
                &sv :CL.time_kat = 1
                &ty alles o.k.
                &sv match = 1
            &end
        &else
            &do
                &sv l = %l% + 1
                &ty oh, oh...
                cursor RES close
                cursor RES remove
                cursor RES declare res_zwecke.lut info ro sex = %sex% AND status = %.status% AND ~
                    pv_avail = %pv_avail% AND d_o_w = %dow% AND a_start = [ calc %arrtime_cat% -
                        + 1 ] AND a_dur = %actdur_cat%
                cursor RES open
                    &if %:RES.aml$next% &then
                        &do
                            &ty alles o.k.
                            &sv :CL.time_kat = 2
                            &sv match = 1
                        &end
                    &else
                        &do
                            &sv k = %k% + 1
                            &ty oh, oh...
                            cursor RES close
                            cursor RES remove
                            cursor RES declare res_zwecke.lut info ro sex =
                                = %sex% AND status = %.status% AND pv_avail = ~
                                    %pv_avail% AND d_o_w = %dow% AND a_start = [ calc ~
                                        %arrtime_cat% - 1 ] AND a_dur = %actdur_cat%
                            cursor RES open
                                &if %:RES.aml$next% &then
                                    &do
                                        &ty o.k.
                                        &sv :CL.time_kat = 3
                                        &sv match = 1
                                    &end
                                &else
                                    &do
                                        &sv j = %j% + 1
                                        &ty oh, oh...
                                    &end
                                &end
                            &end
                        &end
                    &end
            &end
        &end

    &ty ANGABEN CURSOR RES:
    &ty sex = %:RES.sex% status = %:RES.status% pv_avail = %:RES.pv_avail% dow = %:RES.d_o_w%
    &ty a_start = %:RES.a_start% a_dur = %:RES.a_dur% modepur = %:RES.modepur% cellfreq = ~
        %:RES.cellfreq% modepur2 = %:RES.modepur2%
    &ty jetzt time_aim setzen:
    &if %match% = 0 &then &sv &sv :CL.time_aim = 0 /* d.h. falls kein passender Wert in res_zwecke.lut gefunden
        /* wurde dann bleibt time_aim = 0

        &else &sv :CL.time_aim = %:RES.modepur% /* ansonsten wird der passende Wert aus res_zwecke.lut
        /* übernommen

    &if %match% = 1 AND %:CL.time_aim% = 0 &then
        &do
            &ty modepur2 notw. /* falls es eine passende Zeile in
            &sv :CL.time_aim = %:RES.modepur2% /* res_zwecke.lut gibt, aber keinen
            &sv :CL.time_kat = 4 /* Wert in modepur heisst es, es gab
            &end /* keinen Median in dieser Kategorie
            /* und wert aus modepur2 wird übernommen

    cursor RES close
    cursor RES remove
    cursor CL next
    &end
    cursor CL close
    cursor CL remove

    &if [exists %clzweck% -cover] &then kill %clzweck% all
    copy cltmp %clzweck%

    &return

    &routine Zweck Ermitteln
    &ty Zweck_Ermitteln laeuft

    &if [exists cltmp -cover] &then kill cltmp all
    copy %clzweck% cltmp

    &sv selhome = 0
    &sv selwork = 0
    tables

    /* sichere Zwecke

    select cltmp.pat
    reselect home >= 100
    &sv selhome = [show number select]
    &ty selhome = %selhome%
    calc aim = 10
    calc aim_s = 1

    nselect
    reselect anz_index = 2 and time_aim = 5

```

```

&if %status% = 3 and %dow% = 1 &then
    &do
        &sv selwork = [show number select]
        &ty selwork = %selwork%
        calc aim = 5
        calc aim_s = 1
    &end
&if %selwork% > 0 &then
    &do
        &ty selwork --
        select cltmp.pat
        reselect aim = 0
        reselect fn_aim = 5
        calc fn_aim = 0
        calc fn_kat = 0
        select cltmp.pat
        reselect aim = 0
        reselect time_aim = 5
        calc time_aim = 0
        calc time_kat = 0
    &end
    &if %selhome% > 0 &then
        /* entsprechend arbeit für zuhause, d.h., wenn ein
        /* Zuhause aus addr bekannt, dann kein anderes mgl.
        &do
            &ty selhome --
            select cltmp.pat
            reselect aim = 0
            reselect time_aim in { 10, 101, 102}
            calc time_aim = 0
            calc time_kat = 0
        &end
        select cltmp.pat
        reselect aim = 0
        reselect fn_aim = time_aim /* egal, ob kat 1 oder 2, weil, wenn übereinstimmung schon mal grosse WK
        calc aim = fn_aim
        calc aim_s = 1
    /* unsichere Zwecke
    select cltmp.pat
    reselect aim = 0
    reselect time_aim = 1
    calc aim = 1
    calc aim_s = 2
    select cltmp.pat
    reselect aim = 0
    reselect fn_aim ne time_aim
    reselect fn_aim > 0 and time_aim > 0
    calc aim = fn_aim
    calc aim_s = 2
    select cltmp.pat
    reselect aim = 0
    reselect fn_aim = 0
    calc aim = time_aim
    calc aim_s = 3
    select cltmp.pat
    reselect aim = 0
    reselect time_aim = 0
    calc aim = fn_aim
    calc aim_s = 3
    select cltmp.pat
    reselect fn_aim = 0 AND time_aim = 0 or aim = 0
    calc aim = 0
    calc aim_s = 0
q
&if [exists %clzweek2% -cover] &then kill %clzweek2% all
copy cltmp %clzweek2%
&return
&routine aufräumen
&ty AUFRÄUMEN lauft
&if [exists %cov%-cl -cover] &then kill %cov%-cl all
&if [exists %clzweek% -cover] &then kill %clzweek% all
    dropitem %clzweek2%.pat %clzweek2%.pat home pick_drop priv_bus wrel_bus school work shop_daily-
        shop lterm leisure other
    dropitem %clzweek2%.pat %clzweek2%.pat fn_aim fn_kat
    dropitem %clzweek2%.pat %clzweek2%.pat time_aim time_kat anz_index
export cover %clzweek2% %clzweek2%
export cover %outputpoints% %outputpoints%
&return

```

Anhang E: Kartenbeispiele