

Relevanz-Analyse von Geoobjekten

Masterarbeit (GEO 511)
Geographisches Institut der Universität Zürich
Winterthurerstrasse 190
8057 Zürich

Die Arbeit verfasst hat:

Armin Meier
Weingasse 52
5612 Villmergen
armin.meier@access.unizh.ch
03-716-388

Die Arbeit abgenommen haben:

Dr. Tumasch Reichenbacher Prof. Dr. Sara I. Fabrikant

Abgabedatum: 27. Oktober 2008

Inhaltsverzeichnis

1	Einführung	1
1.1	Einführendes Beispiel	1
1.2	Motivation	2
1.3	Ziel der Arbeit	2
1.4	Bezug zur Geographie	3
1.5	Aufbau der Arbeit	3
2	Hintergrund	4
2.1	Thematischer Hintergrund	4
2.1.1	Kontext	4
2.1.2	Kontext-sensitive Systeme	5
2.1.3	Relevanz	9
2.1.4	fuzzy set Theorie	10
2.1.5	Eingrenzung des Themas	11
2.1.6	Stand der Forschung	12
2.2	Methodischer Hintergrund	12
2.2.1	OpenJUMP	12
2.2.2	Funktionsweise des Plugins	13
3	Modellierung der Relevanz	14
3.1	Räumliche Relevanz	14
3.1.1	Definition	15
3.1.2	Räumlicher Relevanztyp I (rR_I)	16
3.1.3	Räumlicher Relevanztyp II (rR_{II})	17
3.1.4	Räumlicher Relevanztyp III (rR_{III})	18
3.2	Zeitlicher Kontext	18
3.2.1	Definition	20
3.2.2	Zeitlicher Relevanztyp (zR)	21

3.3	Thematischer Kontext	23
3.3.1	Definition	24
3.3.2	Thematischer Relevanztyp (tR)	25
3.4	Abweichungen und Modifikationen	26
3.5	Fazit und Ausblick	29
4	Erweiterte Modellierung der Relevanz	30
4.1	Raum und Nutzer:	
	personalisierte räumliche Relevanz (prR)	30
4.2	Raum und Nutzer: Wegzeit-Relevanz (wzR)	32
4.3	Raum, Zeit und Nutzer:	
	korrigierte zeitliche Relevanz (kzR)	33
4.4	Abweichungen und Modifikationen	34
4.5	Fazit und Ausblick	34
5	Formalisierung der Relevanz	36
5.1	Personalisierte räumliche Relevanz	36
5.2	Wegzeit-Relevanz	39
5.3	Korrigierte zeitliche Relevanz	41
5.3.1	Einfach terminierte Geoobjekte	41
5.3.2	Zweifach terminierte Geoobjekte	42
5.4	Thematische Relevanz	43
5.5	Kombination	45
5.5.1	Durchschnitt	45
5.5.2	Gewichteter Durchschnitt	46
5.5.3	Minimum	46
5.6	Fazit und Ausblick	46
6	Resultate	47
7	Diskussion	51
8	Zusammenfassung und Ausblick	53

Abbildungsverzeichnis

2.1	binäre Zugehörigkeit (links) und fuzzy Zugehörigkeit (rechts)	10
3.1	Relevanz in Abhängigkeit der Distanz zwischen Nutzer und Objekt . .	16
3.2	Relevanz in Abhängigkeit Distanz zwischen gleichartigen Objekten . .	17
3.3	Relevanz in Abhängigkeit der Sichtbarkeit	19
3.4	Relevanz in Abhängigkeit von Δt bis zur Startzeit	22
3.5	Relevanz in Abhängigkeit von Δt bis zur Öffnungszeit	23
3.6	Relevanz in Abhängigkeit von Δt bis zur Endzeit	23
3.7	Relevanz in Abhängigkeit der semantischen Distanz	25
3.8	Räumliche Relevanz für einen Velofahrer	26
3.9	Relevanz in Abhängigkeit der semantischen Distanz	27
3.10	Nicht-linear abnehmende räumliche Relevanz	28
3.11	Diskretisierte räumliche Relevanz	28
3.12	Zusammenfassung	29
4.1	Konkrete Zeitangaben	31
4.2	Relevanz in Abhängigkeit der Wegzeit	33
4.3	Zusammenfassung	35
5.1	Konkrete Zeitangaben	37
5.2	Konkrete Zeitangaben	39
5.3	Konkrete Zeitangaben	41
5.4	Relevanz auf Zeitstrahl	42
5.5	Relevanz in Abhängigkeit von $\Delta t_{effektiv}$ (bis Endzeit)	43
6.1	Interface	47
6.2	Screenshot vom resultierenden layer	48
6.3	Tabelle der Geoobjekte	48
6.4	Tabelle der Geoobjekte	49

6.5	Tabelle der Geoobjekte	49
6.6	Tabelle der Geoobjekte	50
6.7	Tabelle der Geoobjekte	50

Tabellenverzeichnis

5.1	<i>m</i> -Werte	38
5.2	Geschwindigkeit des Nutzers in [m/min]	40
5.3	Ähnlichkeiten zwischen verschiedenen Objektarten	44

Kapitel 1

Einführung

Das folgende Kapitel bietet eine Einführung in die Thematik dieser Arbeit. Es wird beschrieben, inwiefern die Arbeit von Interesse ist indem der Stellenwert der Arbeit in einem grösseren Kontext aufgezeigt wird. Anschliessend werden die Ziele der Arbeit und der Bezug zur Geographie erläutert. Zuletzt folgt noch die Beschreibung des Aufbaus dieser Arbeit.

1.1 Einführendes Beispiel

Stellen Sie sich vor, Sie befinden sich in einer fremden Stadt und müssen sich darin zurecht finden. Dabei haben Sie beispielsweise die Absicht ein Theater zu besuchen. Der PDA, den Sie zur Orientierung in Ihren Händen halten, hat nun mehrere Möglichkeiten, die Geoinformationen in kartographischer Form darzustellen. Beispielsweise könnte er alle Geoobjekte der Stadt, ungeachtet Ihrer Absichten, darstellen und die Karte so unter Umständen zu überladen. Eine ander Möglichkeit wäre, wenn er nur die Geoobjekte, an welchen Sie interessiert sind, darstellen würde, wodurch der Karteninhalt reduziert werden könnte. Eine raffiniertere Art der Informationsübermittlung wäre jedoch, wenn der PDA in der Lage wäre, alle Geoobjekte bezüglich des aktuellen Kontextes (Position in Raum und Zeit, Wetterlage, Interessen etc.) vorgängig zu gewichten und nur die Geoobjekte darstellen würde, welche einen hohen Wert erhielten und daher als besonders relevant gelten. Auf diese Weise könnte der Karteninhalt weiter reduziert und die Informationsübermittlung optimiert werden, ohne dabei wichtigen Informationen zu verlieren. Der Schlüssel für einen solchen Kontext-sensitiven Kartendienst liegt in der Entwicklung eines Formalismus, welcher in der Lage ist, die Relevanz von Geoobjekten bezüglich des Kontextes zu bestimmen.

1.2 Motivation

Internetsuchmaschinen sind schon längere Zeit im Stande, die Anfrage des Nutzers mit dem Inhalt von Dokumenten zu vergleichen und die Dokumente dieser Übereinstimmung entsprechend zu bewerten. Basierend auf dieser Bewertung werden die Dokumente anschliessend aufgelistet, beginnend mit demjenigen Dokument, welches den höchsten Wert erhalten hat und somit die höchste Übereinstimmung mit der Anfrage aufweist. Diese Rangierung ermöglicht dem Nutzer, innert kurzer Zeit das *world wide web* nach den für ihn wichtigsten Dokumenten zu durchsuchen.

Dieses Bedürfnis, den Nutzer von einer Informationsflut zu bewahren, besteht auch im Bereich der mobilen Kartographie ([Reichenbacher et al.(2002)]). Dabei sollen keine Dokumente, sondern Objekte mit Raumbezug, fortan Geoobjekte genannt, bewertet werden. Auf einem Monitor oder Display alle Geoobjekte darzustellen, kann den Nutzer überfordern und besonders die kleinen Displays mobiler Geräte überladen. Ein Relevanz-Wert könnte auch bei Geoobjekten als Kriterium dienen, ob oder wie salient ein Objekt dargestellt werden soll. Dadurch könnte die Lesbarkeit einer Karte erhöht werden, ohne dabei (wichtige) Informationen zu verlieren.

Die Anwendungsmöglichkeiten eines solchen Bewertungs-Formalismus sind vielfach. Mit Hilfe dieses Formalismus könnte im Bereich der mobilen Kartographie die Extraktion von Geoobjekten aus einer geographischen Datenbank (aufgrund ihrer Relevanz) unterstützt werden. Ausserdem könnten Bildschirmkarten, welche auf kleinen Displays angezeigt werden, die Symbolisierung von Geoobjekten ihrer Relevanz anpassen und so die Navigation des Kartenlesers vereinfachen ([Reichenbacher(2004)]). Entscheidungsunterstützungssysteme, wie etwa ein mobiler digitaler Touristenführer, könnten mit Hilfe dieses Formalismus Geoobjekte bewerten und dem Nutzer geeignete Routen- oder Besichtigungsvorschläge unterbreiten [Malaka und Zipf(2000)].

1.3 Ziel der Arbeit

Die vorliegende Arbeit trägt einen Teil zu dieser Entwicklung bei, indem ein Formalismus zur Bewertung solcher Geoobjekte in Bezug auf ihren aktuellen räumlichen, zeitlichen und thematischen Kontext ausgearbeitet wird. Mit Hilfe dieses Bewertungs-Formalismus soll die Relevanz von Geoobjekten bestimmt werden. Nicht verwechselt werden sollte diese Art der Bewertung mit der Bewertung von geographischen Informationen in Form von *Dokumenten*. Diese Thematik fällt in den Bereich der Geographic Information Retrieval (SPIRIT, STORM -> Quellen!).

Die Leitfrage dieser Arbeit lautet:

Wie kann die Relevanz eines Geoobjektes in Bezug auf seinen räumlichen, zeitlichen und Nutzer-spezifischen Kontext formalisiert werden?

Ausserdem soll dieser Formalismus in dem open source GIS-Programm (OpenJUMP) als Plugin implementiert und seine Resultate diskutiert werden.

1.4 Bezug zur Geographie

Einerseits verursacht die voranschreitende Vernetzung elektronischer Geräte (*ubiquitous computing*, Weiser) einen immer grösser werdenden geographischen Datenberg, andererseits jedoch führt diese Vernetzung auch dazu, dass immer mehr Kontext-Informationen abrufbar werden (Position, Wetter, Verkehr, Öffnungszeiten etc.), welche die Bewertung von Informationen ermöglichen. Diese Arbeit ist für die Geographie insofern von Bedeutung, als dass sie einen Formalismus zur Bewertung von *geographischen* Informationen (in Form von Geoobjekten) in Bezug auf ihren aktuellen Kontext erzeugt und so Aussagen über ihre Relevanz ermöglicht.

Daher lässt sich die Thematik dieser Arbeit in die Disziplin der Geographischen Informationsvisualisierung und -analyse (GIVA) einordnen.

1.5 Aufbau der Arbeit

Die Arbeit ist so aufgebaut, dass im Kapitel 2 die im Zusammenhang mit der Arbeit wichtigen Begriffe erklärt und bestehende Kontext-sensitive Ansätze und deren Abgrenzung zur eigenen Arbeit beschrieben werden. Im darauffolgenden Kapitel 3 wird erläutert, wie sich die Kontext-Dimensionen Raum, Zeit und Nutzer auf die Relevanz eines Geoobjektes auswirken können. Ausserdem wird erklärt, wie diese Zusammenhänge zwischen den einzelnen Kontext-Dimensionen und der Relevanz in dieser Arbeit qualitativ modelliert werden. Anschliessend werden im Kapitel 4 die Kontext-Dimensionen miteinander kombiniert und ebenfalls modelliert. Die konkrete Formalisierung dieser Modelle, durch welche die Relevanz eines Geoobjektes quantitativ bestimmt werden soll, folgt im Kapitel 5. Dieser Formalismus wird im open source GIS OpenJUMP implementiert und die entsprechenden Resultate im Kapitel 6 präsentiert, ohne dabei auf den Code einzugehen. Beendet wird die Arbeit durch die Diskussion im Kapitel 7 und die anschliessende Zusammenfassung mitsamt Ausblick im Kapitel 8.

Kapitel 2

Hintergrund

Nachdem im vorangehenden Kapitel die Motivation und die Ziele dieser Arbeit aufgeführt worden sind, beleuchtet dieses Kapitel nun den thematischen und methodischen Hintergrund. Dabei werden einerseits die in dieser Arbeit zentralen Begriffe und verwandte Ansätze vorgestellt. Andererseits soll die Vorgehensweise kurz erklärt werden, wie der Formalismus zur Bewertung der Relevanz von Geoobjekten konkret implementiert wird.

2.1 Thematischer Hintergrund

In diesem Unterkapitel werden die im Zusammenhang mit dieser Arbeit zentralen Begriffe *Kontext*, *Kontext-sensitive Systeme*, *Relevanz* und *fuzzy set* erläutert. Außerdem wird die Thematik dieser Arbeit eingegrenzt und der Stand der Forschung wird aufgezeigt.

2.1.1 Kontext

Der Begriff *Kontext* spielt in dieser Arbeit eine zentrale Rolle, denn die Relevanz eines Geoobjektes hängt stets vom Kontext ab (s. Kapitel 2.1.3). Aus diesem Grund wird an dieser Stelle eine Definition des Begriffs *Kontext*, welcher allgemein ziemlich unscharf festgelegt ist und je nach Zusammenhang anders verstanden werden kann geliefert, geliefert, welche im Rahmen dieser Arbeit gelten soll.

Die im Zusammenhang mit der Informationswissenschaft, wozu auch die mobile Kartographie gezählt werden kann, wohl verbreiteste Definition von *Kontext* stammt von [Abowd et al.(1999)]:

„Context is any information that can be used to characterize the situation

of an entity. An entity is a person, place, or object that is considered relevant to the interaction between a user and an application, including the user and applications themselves.“

Danach kann jede Art von Information, welche für das Verhalten einer Anwendung relevant ist, als Kontext-Information verstanden werden. Anders formuliert gilt Kontext als Input einer Anwendung, welcher den Output steuert oder beeinflusst. Der Kontext ist dabei ständigen Veränderungen unterworfen, weshalb er eher als ein dynamischer Prozess als ein statisches Phänomen zu verstehen ist (Beale 04 -> Quelle!).

Der Terminus des Kontextes ist sehr weitläufig („Context is *any information* ...“) und daher es ist sinnvoll, diese Bedeutungsvielfalt durch eine Klassierung zu strukturieren. Eine Möglichkeit besteht in der Unterscheidung zwischen *Primär-* und *Sekundärkontext* ([Abowd et al.(1999)]). *Ort, Identität, Aktivität* und *Zeit* sind dabei als primäre Kontext-Informationen zu verstehen, welche auf sekundäre Kontext-Informationen verweisen und daher als Referenten zu verstehen sind. Die Identität eines Nutzers (Primärkontext) kann so beispielsweise auf seine Telephonnummer, seine Adresse und sein Geburtstagsdatum (Sekundärkontext) verweisen.

Eine andere Klassifizierung sieht die Unterteilung in low-level und high-level Kontext-Informationen vor ([Chen und Kotz(2000)]). Während die low-level Kontext-Informationen über physikalische Sensoren *gemessen* werden können (Ort, Temperatur, Strahlung etc.), kann beispielsweise das soziale Umfeld eines Nutzers (high-level Kontext) nur aus der Kombination seiner Position und der Position benachbarter Nutzer *abgeleitet* werden kann.

[Reichenbacher(2004)] schlägt im Zusammenhang mit der mobilen Kartographie vor, die *Kontext-Informationen* zu *Kontext-Dimensionen* zusammenzufassen. Dabei schlägt er die Kontext-Dimensionen *Situation* (Position in Raum und Zeit), *Nutzer* (Interessen, Einschränkungen etc.), *Aktivität, Information* (Events, Wetter etc.) und *System* (Displaygrösse, Übertragungsrate etc.) vor. Weitere Kategorisierungsformen siehe ([Schilit et al.(1994)], [Schmidt et al.(1999)], [Schmidt et al.(1998)]).

Die Kontext-Informationen, welche in dieser Arbeit verwendet werden, werden im Kapitel 2.1.5 aufgelistet.

2.1.2 Kontext-sensitive Systeme

Nachdem im vorangehenden Kapitel der Begriff des Kontextes erläutert worden ist, soll nun eine Definition von Kontext-sensitiven Systeme mitsamt Beispielen geliefert

werden.

Nutzt ein System Kontext-Informationen dazu, einen Nutzer mit wichtigen (Kontext-abhängig) Informationen zu versorgen, spricht man im Zusammenhang der Informationswissenschaften von *context-aware systems* ([Abowd et al.(1999)]). Die Autoren liefern dazu folgende Definition:

„A system is context-aware if it uses context to provide relevant information and/or services to the user, where relevancy depends on the user's task.“

Dabei klassieren sie ([Abowd et al.(1999)]) Kontext-sensitive Systeme aufgrund ihrer Funktionen, welche sich allesamt einzelnen Kontext-Informationen anpassen können:

1. *Präsentation* von Informationen oder Diensten
Bsp.1: Bei schlechtem Wetter sollen dem Nutzer keine Informationen über offenen Badeanstalten kommuniziert werden (Selektion von Informationen).
Bsp.2: Die Schriftgröße soll sich nach dem Sehvermögen des Nutzers richten (Darstellung der Informationen).
Bsp.3: Einem Touristen wird die Dienstleistung angeboten, eine persönliche Besichtigungstour berechnen lassen ([Malaka und Zipf(2000)]).
2. automatische *Ausführung* von Diensten
Bsp.: Sobald sich ein Tourist einer Sehenswürdigkeit nähert, soll sich automatisch ein Informationsfenster auf dem Display öffnen ([Schmidt-Belz et al.(2002)]).
3. Daten mit Kontext-Informationen versehen (*tagging*)
Bsp.: Heutige Digitalkameras versehen jede Bilddatei mit dem Zeitpunkt, zu welchem das Photo geschossen wurde.

Jedes Kontext-sensitive System muss eine Auswahl an Kontext-Informationen treffen, an welchen eine Funktion angepasst werden soll. In der Folge werden nun in chronologischer Reihenfolge verschiedene Kontext-sensitive Systeme und ihre berücksichtigten Kontext-Informationen genannt.

ParcTab

Eines der ersten Kontext-sensitiven Systeme wurde im *Xerox PARC* in Palo Alto entwickelt und trägt den Namen **ParcTab** ([Want et al.(1995)]). Neben der Entwicklung eines Gerätes (hardware) für Büro Zwecke, welches drahtlose Kommunika-

tion und mobiles Rechnen unterstützt, ging es darum, Kontext-sensitive Applikationen (software) darauf zu implementieren, welche verschiedene positionsabhängige Informationen abbilden (in welchem Raum man sich befindet, wo der nächste Drucker ist, welche Personen sich in der Nähe befinden etc.).

CYBERGUIDE

Mitten der neunziger Jahre wurde am *Georgia Institute of Technology* das Projekt **Cyberguide** ([Long et al.(1996)]) gestartet, welches als Kontext-sensitiver elektronischer Touristenführer für mobile Nutzer agieren sollte. Ähnlich dem vorgängig vorgestellten System ParcTab wurde die Position, zusätzlich jedoch auch die Orientierung, welche aus dem Bewegungspfad abgeleitet wurde, dazu verwendet, dem Nutzer mit relevanten Informationen zu versorgen (welches Objekt gerade betrachtet wird etc.). Im Gegensatz zu ParcTab bilden Touristen das Zielpublikum und dank Positionsbestimmung durch GPS liess sich das System auch *outdoor* und nicht nur *indoor* durch Infrarot anwenden.

GUIDE

Das Projekt **GUIDE**, ein Gemeinschaftsprojekt der Universität und Stadt Lancaster, zielte darauf ab, einen Kontext-sensitiver Tourenführer für die Stadt Lancaster zu entwickeln ([Cheverst et al.(2000)]). GUIDE ist imstande, das Nutzerprofil und die zur verfügung stehende Zeit (beides musste manuell eingegeben werden) zusammen mit der aktuellen Position dazu zu verwenden, eine Besichtigungstour für die Stadt Lancaster zu berechnen und benutzerspezifische und aktuelle Points of Interest zu ermitteln.

DEEP MAP

Das Projekt **DEEP MAP** war das erste Projekt am *European Media Laboratory* (EML) in Heidelberg und sollte als digitaler persönlicher mobiler Touristenführer für die Stadt Heidelberg fungieren ([Malaka und Zipf(2000)]). Das Ziel dieses interdisziplinären Forschungsprojektes war die Entwicklung eines Systems, welches umfangreiche und heterogene Datensätze mit einer Vielzahl an Funktionalitäten kombinieren soll. Beispielsweise wurde der Dienst an den Nutzer derart angepasst, dass vom System personalisierte Touren durch Heidelberg vorgeschlagen werden konnten, welche sich nach den Interessen und Fähigkeiten des Nutzers richten.

CRUMPET

Ebenso wie das Projekt DEEP MAP war auch das Projekt **CRUMPET** (CReation of User-friendly Mobile services PErsonalised for Tourism) ein Forschungsprojekt, welches am *EML* durchgeführt worden ist ([Schmidt-Belz et al.(2002)]). Wie dem Namen des Projektes zu entnehmen ist, ging es unter anderem darum, benutzerfreundliche und personalisierte Dienste zu implementieren. Dem ist insofern Rechnung getragen worden, als dass das System dem Nutzer in Abhängigkeit von seinem Standort und seinen Interessen einerseits Informationen proaktiv (ohne Aufforderung) übermittelt und andererseits Touren von seinem aktuellen Standort zum gewünschten Ziel berechnet worden sind.

COMPASS

Das System **COMPASS** (COntext-aware Mobile Personal ASSistant), welches vom *Telematica Instituut* in Enschede entwickelt worden ist, kann einem Touristen aufgrund seines aktuellen Kontextes (Position in Raum und Zeit) und seiner persönlichen Interessen Empfehlungen darüber machen, welche Points of Interests besonders relevant sind ([Setten et al.(2004)]). In einem ersten Schritt werden aufgrund harter Kriterien (räumliche Entfernung) die Geoobjekte gefiltert, anschliessend werden sie aufgrund weicher Kriterien (Zeit und Interessen des Nutzers) bewertet.

SAiMotion

Das Projekt **SAiMotion** (Situation Awareness in Motion) zielt darauf ab, ein mobiles Besucherinformationssystem für Messen zu entwickeln, welches erlaubt, Informationen proaktiv und an den Benutzer und seinen aktuellen Standort angepasst zu selektieren und für ihn eine persönliche Führung durch das Messegelände zu bieten ([Hermann und Heidmann(2002)]). Die zeitliche Dimension spielt bei der Routenberechnung insofern eine Rolle, als dass die Route so berechnet wird, dass die vom Benutzer angegebenen Termine eingehalten werden können.

UbiquiTO

Als letzten Ansatz soll an dieser Stelle **UbiquiTO** genannt werden, ein adaptiver Touristenführer für Turin welcher besonders auf die Gestaltung des Abendprogrammes ausgelegt ist ([Amendola et al.(2004)]). Die Anpassung erfolgt derart, dass die räumliche Nähe und die Interessen des Nutzers benutzt werden, um jedem Objekt einen Wert zu verleihen um sie anschliessend in geordneter Reihenfolge dem Nutzer

als relevanteste POIs vorzuschlagen. Ausserdem wird die Grösse des Gerätedisplays und weitere Kontext-Dimensionen, wie etwa Tageszeit oder die Tatsache, ob sich der Nutzer schnell oder langsam fortbewegt, dazu verwendet, die Benutzerschnittstelle über die Schriftgrösse oder die Farbwahl entsprechend zu gestalten.

2.1.3 Relevanz

Der Begriff der *Relevanz* wird in verschiedenen Disziplinen verwendet, beispielsweise in der Logik, in der Philosophie, in den Informations- und Kommunikationswissenschaften ([Saracevic(1996)]). Je nach Disziplin kommt dem Begriff somit eine andere Bedeutung zu, was das Konzept der Relevanz schwer fassbar macht ([Reichenbacher(2004)]). In dieser Arbeit interessiert das Konzept der Relevanz im Zusammenhang mit der Disziplin der Informationsswissenschaften.

[Saracevic(1996)] definiert die Relevanz als ein „... criterion reflecting the effectiveness of exchange of information between people (or between people and objects potentially conveying information) in communicative relation, all within a context.“. Demnach beschreibt die Relevanz den kontext-abhängigen Wirkungsgrad (*effectiveness*), mit welchem der verbale und non-verbale (Karte) Informationsaustausch abläuft. [Sperber und Wilson(1986)] definieren diesen Wirkungsgrad als Verhältnis zwischen *cognitive effects* und *processing effort*. Eine Information ist demnach für den Empfänger relevant, wenn sie in einem gegebenen Kontext möglichst wirkungsvoll seinen Informationsbedarf zu decken vermag, wenn also eine hohe Wirkung der Information mit einem niedrigen Aufwand (Verarbeitung der Information) des Empfängers einhergeht.

Der Kontext ist bei der Bestimmung der Relevanz von zentraler Bedeutung, denn die Relevanz einer Information ist immer von einem oder mehreren Kontext-Informationen abhängig und kann nicht unabhängig davon bestimmt werden. So gilt beispielsweise, dass dieselbe Information für unterschiedliche Empfänger unterschiedlich relevant ist.

Diese Beschreibung der Relevanz mit ihrem wichtigen Merkmal, der Kontext-Abhängigkeit, lässt sich auch auf geographische Informationen anwenden, denn auch eine geographische Information kann für einen Empfänger, abhängig vom Kontext, mehr oder weniger relevant sein.

Grundsätzlich kann zwischen der objektiven und der subjektiven Relevanz unterschieden werden ([Reichenbacher(2004)]). Die objektive Relevanz kommt bei den klassischen Bewertungs- und Rangierungsmethoden im Bereich von Information-Retrieval-Systemen vor (wie z.B. Internet-Suchmaschinen), wenn durch die beiden

Bewertungsparameter *precision* und *recall* unabhängig vom subjektiven Kontext des Nutzers die Trefferqualität von Dokumenten bestimmt wird ([Machill und Beiler(2007)]). Im Bereich der Pragmatik und der Kommunikation macht die objektive Relevanz nicht viel Sinn, denn ob eine Information relevant ist oder nicht, hängt grösstenteils vom Empfänger (Nutzer) ab und kann nicht allgemein bestimmt werden. Die subjektive Relevanz bezeichnet also die Relevanz in Abhängigkeit des Nutzers und weiterer Kontext-Informationen.

Während Dokumente beim Information-Retrieval häufig entweder relevant oder nicht-relevant sind (binär)([Saracevic(2007)]), wird in dieser Arbeit mit Hilfe eines Formalismus die Bewertung von Geoobjekten angestrebt, welche auch einen Relevanzwert dazwischen annehmen können. Auf diese *fuzzy set Theorie* wird im folgenden Abschnitt näher eingegangen.

2.1.4 fuzzy set Theorie

Die *fuzzy set*-Theorie von [Zadeh(1965)] beschreibt, dass sich die Zugehörigkeit eines Elementes zu einer Menge nicht immer mit „0“ oder „1“, sondern mit einer Zahl (zwischen „0“ und „1“) als Zugehörigkeitsgrad beschrieben werden kann (Abbildung 2.1).

Diese Theorie soll in dieser Arbeit zur Anwendung kommen, wonach ein Geoobjekt

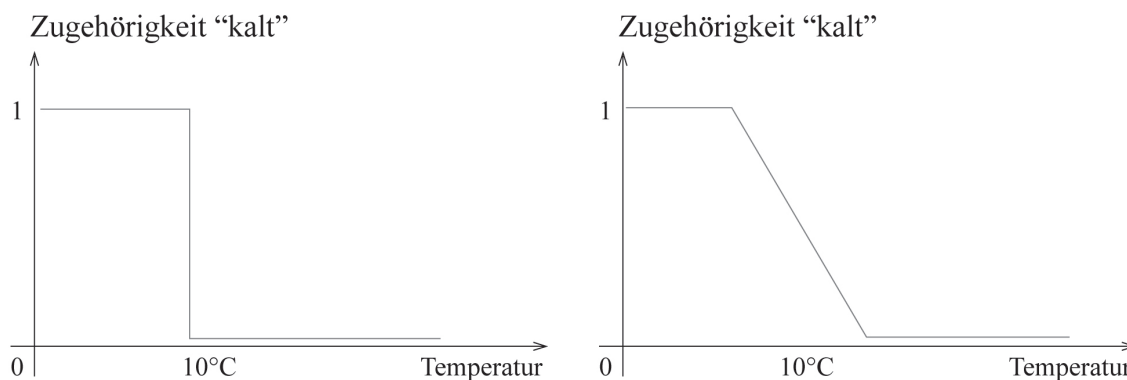


Abbildung 2.1: binäre Zugehörigkeit (links) und fuzzy Zugehörigkeit (rechts)

nicht nur als relevant respektive als nicht relevant betrachtet werden soll (binär), sondern eben auch einen dazwischenliegenden Wert annehmen kann. Über eine Zugehörigkeitsfunktion (*membership function*) soll somit für jedes Objekt der Relevanzgrad, welcher zwischen „0“ und „1“ liegen soll, bestimmt werden.

2.1.5 Eingrenzung des Themas

Bis anhin wurden die zentralen Begriffe und verwandte Ansätze erläutert, sodass nun auf die Eingrenzung des Themas dieser Arbeit, respektive die Abgrenzung gegenüber der erwähnten Ansätze, eingegangen werden kann.

Im Vergleich zu den oben genannten Projekten, welche sich mit kompletten Architekturen solcher Kontext-sensitiver Dienste (Wahrnehmen, Verwalten und Anwenden von Kontext-Informationen) und im Falle von ParcTab gar mit der Entwicklung von hardware beschäftigen, befasst sich die vorliegende Arbeit mit der Entwicklung einer Formel zur Bewertung der Relevanz von Geoobjekten. Diese Formel könnte dann in einer konkreten Anwendung innerhalb einer solchen Architektur zum Einsatz kommen, beispielsweise könnten mit ihr Geoobjekte für eine Karte entsprechend ihrer Relevanz gefiltert oder dargestellt werden. Die Beschaffung der Kontext-Informationen sowie auch deren Verwaltung werden in dieser Arbeit als vorausgesetzt betrachtet.

Wie bereits angesprochen, gibt es unzählige Kontext-Informationen und nicht alle können in dieser Arbeit für die Bestimmung der Relevanz eines Geoobjektes einbezogen werden. In der vorliegenden Arbeit werden folgende Kontext-Dimensionen berücksichtigt:

- *Raum*: Die räumliche Verteilung der Geoobjekte und des Nutzers haben einen Einfluss auf die Relevanz eines Geoobjektes.
- *Zeit*: Die aktuelle Tageszeit und die Öffnungszeiten von Geoobjekten beeinflussen ebenfalls die Relevanz eines Geoobjektes.
- *Nutzer*: Diese Kontext-Dimension ist äusserst vielschichtig und enthält viele, teilweise schwer erfassbare, Kontext-Informationen. In dieser Arbeit werden dabei folgende Kontext-Informationen des Nutzers verwendet:
 - *Interessen* (Thematik): Die Interessen des Nutzers bezüglich der Kategorie eines Geoobjektes sind massgeblich an der Relevanz eines Objektes beteiligt.
 - *Alter* und *Fortbewegungsmittel*: Diese beiden Kontext-Informationen des Nutzers können ebenfalls zur Bestimmung der Relevanz eines Geoobjektes herangezogen werden.

2.1.6 Stand der Forschung

[Zipf und Richter(2002)] verwenden als erste den Begriff der Relevanz im Zusammenhang mit der mobilen Kartographie. Ihr Konzept der Fokuskarte sieht vor, relevante Geoobjekte kartographisch durch einen geringeren Generalisierungsgrad und durch geeignete Farbgebung hervorzuheben. Dieser Ansatz verfolgt als Ziel weniger die Bestimmung der Relevanz von Geoobjekten, denn relevante Objekte werden pragmatisch durch die räumliche Nähe zur Position des Nutzers festgelegt, sondern vielmehr die geeignete kartographische Darstellung. [Alexander Zipf(2003)] lieferte später eine Formalisierung der Relevanz für Geoobjekte in solchen Fokuskarten. Darin adaptiert er das Verfahren aus der Informationsvisualisierung von [Keim und Kriegel(1994)] zur Berechnung eines solchen Relevanz-Wertes, indem er die Relevanz eines Geoobjektes als Funktion der Summe der gewichteten Distanzen zwischen Merkmalen des Geoobjektes und den korrespondierenden Anfrageparametern, zuzüglich der Eignung des Geoobjektes als Landmarke, definiert. Diese Formalisierung ist wenig konkret, denn weder die Kontext-Informationen noch ihr Einfluss auf die Relevanz eines Geoobjektes werden darin beschrieben. Dennoch zeigt dieser Ansatz die Notwendigkeit, die Relevanz von Geoobjekten quantitativ zu bestimmen. Seither ist die Forschung über die Formalisierung der Relevanz für Geoobjekte ins Stocken geraten. Die zuvor beschriebenen Systeme COMPASS ([Setten et al.(2004)]) und UbiquiTO ([Amendola et al.(2004)]) verwenden zwar auch den Kontext, um die Relevanz von Geoobjekten zu bestimmen, jedoch ist unklar, wie genau dieser Einfluss des Kontexts auf die Relevanz formalisiert worden ist. Nach Angaben von Herrn Alexander Zipf wird diese Thematik der Relevanz-Analyse von Geoobjekten zur Zeit am kartographischen Institut der Universität in Bonn wieder aufgenommen.

2.2 Methodischer Hintergrund

In diesem Kapitel soll die Vorgehensweise erklärt werden, mit welcher in dieser Arbeit die Formalisierung von Relevanz in OpenJUMP implementiert und umgesetzt wird. Dazu wird zuerst das open source GIS OpenJUMP und anschliessend die Funktionsweise dieses implementierten Plugins vorgestellt.

2.2.1 OpenJUMP

OpenJUMP ist ein geographisches Informationssystem, welches gratis im *world wide web* heruntergeladen werden kann. Dabei handelt es sich um ein Informationssys-

tem, welches von den Firmen *Vivid Solutions* und *Refractions Research* entwickelt worden ist. OpenJUMP ist eine Abkürzung und steht für *Open Java Unified Mapping Platform*. Open bedeutet, dass das Programm open source (quelloffen) ist und der (JAVA-) Code somit für alle einsichtbar und durch Plugins erweiterbar ist.

Neben OpenJUMP gibt es beispielsweise auch noch QGIS (Quantum GIS) oder GRASS (Geographic Resources Analysis Support System), welche ebenfalls quelloffen und gratis zur Verfügung stünden. Die Wahl fiel auf OpenJUMP, weil die Programmierschnittstelle dieser Plattform über JAVA und nicht über C++, wie bei den anderen beiden Systemen, erweitert werden kann und der Autor bereits über JAVA-Kenntnisse verfügt.

OpenJUMP ist quelloffen und lässt sich, wie oben bereits angedeutet, durch Plugins erweitern. Diese Plugins, welches als kleine Programme in einem übergeordneten Programm verstanden werden können, können entweder heruntergeladen oder selber in JAVA programmiert werden. In dieser Arbeit wird ein Plugin selber programmiert, welches die Relevanz von Geoobjekten bezüglich ihres Kontextes quantifizieren soll. Im Kapitel 6 werden die Resultate vorgestellt, ohne dabei näher auf den Code einzugehen.

2.2.2 Funktionsweise des Plugins

Das Plugin, mit welchem das GIS-Programm OpenJUMP in dieser Arbeit erweitert wird, benötigt bestimmte Kontext-Informationen, um die Relevanz der Geoobjekte bestimmen zu können (s. Kapitel 2.1.5). Einerseits gibt der Nutzer diese Informationen dem Plugin explizit bekannt, indem er mit Hilfe eines Interfaces persönliche Angaben zu seinem Alters, seinem Fortbewegungsmittel und seinen Interessen macht. Andererseits stammen die anderen Kontext-Informationen, nämlich die räumliche Verteilung der Geoobjekte, ihre Kategorie (Objektart) und zeitliche Fixierung von den Geoobjekten selbst und liegen so, zusammen mit der aktuellen (System-) Zeit dem Plugin, implizit vor. Die aktuelle Position des Nutzers wird in dieser Arbeit nicht über GPS bestimmt und dem Plugin automatisch übermittelt, sondern muss manuell eingegeben und verändert werden.

Diese Kontext-Informationen werden anschliessend vom Plugin dazu verwendet, jedem Geoobjekt einen Relevanz-Wert zuzuteilen. Wie genau aus diesen einzelnen Informationen ein Relevanz-Wert abgeleitet werden kann, wird im folgenden Kapitel beschrieben.

Kapitel 3

Modellierung der Relevanz

Wie der Einleitung zu entnehmen ist, wird in dieser Arbeit eine Formalisierung der Relevanz angestrebt, um dadurch die Bedeutung einzelner Geoobjekte für einen Nutzer unter Berücksichtigung des Kontextes quantitativ ausdrücken zu können. Im vorliegenden Kapitel wird für die drei Dimensionen *Raum*, *Zeit* und *Nutzer* einzeln untersucht, wie sich die Relevanz in Abhängigkeit ihres jeweiligen Kontextes verhält. Dieser Zusammenhang zwischen der Kontext-Dimension und entsprechender Relevanz soll in einem Modell veranschaulicht werden.

Das Ziel dieses Kapitels ist die qualitative Modellierung der räumlichen, zeitlichen und thematischen Relevanz in Abhängigkeit der zugehörigen Kontext-Dimension.

3.1 Räumliche Relevanz

Die Entwicklung von GPS ermöglichte schon früh die Ausnutzung der Informationen über die räumliche Verteilung von Nutzer, Geoobjekten und weiterer Personen. Solche Dienste werden unter der Bezeichnung Location Based Services (LBS) zusammengefasst (Quelle:LBS-Paper Wiley Verlag. . .). Bei LBS handelt es sich auch um Kontext-Sensitive Systeme, welche in Abhängigkeit der Position des Nutzers bestimmte Dienste ermöglichen. Beispielsweise sind sie in der Lage, Geoobjekte aufgrund ihrer räumlichen Entfernung zum Nutzer aus einer Datenbank zu filtern und so nur räumlich nahe Objekte darzustellen. Auch das vorher beschriebene System COMPASS ([Setten et al.(2004)]) ist dazu in der Lage. Das System Deep Map ([Malaka und Zipf(2000)]) verwendet die Position des Nutzers, um dem Nutzer proaktiv (ohne Aufforderung) wichtige Informationen anzubieten, sobald er sich

einem Objekt nähert.

Anstatt die räumliche Relevanz mit Hilfe einer Formel zu quantifizieren und (gemäss dieses Wertes) beispielsweise unterschiedlich salient darzustellen, verschneiden solche Dienste einen Buffer, definiert als eine bestimmte Distanz um die Position des Nutzers, mit den Geoobjekten und enden in einer binären Lösung. Der Nachteil solcher Dienste liegt darin, dass keine Aussage darüber möglich ist, welche Geoobjekte von hohem und welche Geoobjekte von niederem Interesse sind. Ein Objekt wird somit entweder als relevant erachtet und dargestellt oder eben nicht.

Dieses Unterkapitel befasst sich mit der Modellierung der räumlichen Relevanz, um eine feinere Unterscheidung der räumlichen Relevanz als die binäre Vorgehensweise zu ermöglichen (vgl. fuzzy set, Kapitel 2.1.4). Zuerst wird auf den Begriff der *räumlichen Relevanz* eingegangen indem drei unterschiedliche Relevanztypen vorgestellt werden, welche sich jeweils alle lediglich auf den räumlichen Kontext beziehen. Anhand eines einfachen Modelles wird dabei gezeigt, wie sich der jeweilige Relevanztyp in Abhängigkeit der räumlichen Verteilung von Geoobjekt und Nutzer verhält.

Das Ziel dieses Unterkapitels ist es, den Einfluss der räumlichen Verteilung von Geoobjekt und Nutzer auf die Relevanz eines Geoobjektes zu beschreiben und zu modellieren.

3.1.1 Definition

Im Kapitel 2.1.3 wurde erläutert, dass sich der Begriff *Relevanz* immer auf einen oder mehrere Kontext-Dimensionen bezieht. Die räumliche Relevanz berücksichtigt somit die Dimension *Raum*, also die Verteilung des Nutzers und der Geoobjekte im geographischen Raum respektive deren Lagebeziehung untereinander. Ergo bezeichnet die räumliche Relevanz die Bedeutung eines Geoobjektes, alleine hinsichtlich seiner Position im geographischen Raum beziehungsweise seiner euklidischen Entfernung zum Nutzer oder zu einem anderen Objekt.

Der Raum kann unterschiedlich grob verstanden werden, [Reichenbacher(2004)] spricht dabei von unterschiedlicher Granularität (z.B. Koordinatenpaar, Ortsbezeichnung, Region). In dieser Arbeit wird die räumliche Information eines Geoobjektes in Koordinatenpaar-Form verwendet. Ausserdem wird der Raum in dieser Arbeit rein geometrisch (in Koordinatenform) verstanden, d.h. Komponenten wie etwa die Bedeutung eines Ortes werden nicht berücksichtigt (vgl. *space* versus *place*, [Edwardes(2007)]).

In der Folge wird zwischen drei verschiedenen Typen der räumlichen Relevanz unterschieden und ihr Verhalten in Bezug auf den geographischen Raum wird modelliert.

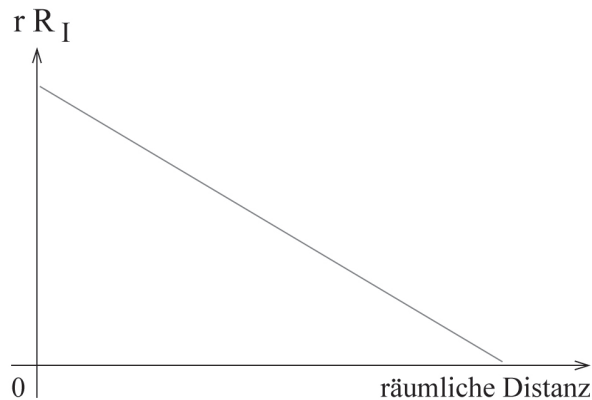
3.1.2 Räumlicher Relevanztyp I (rR_I)

Die erste Art der räumlichen Relevanz ist eine Funktion der räumlichen Entfernung zwischen dem Nutzer und dem Geobjekt (GO).

$$rR_I = f(\text{Distanz}_{\text{Nutzer-GO}}) \quad (3.1)$$

In den meisten Fällen ist die Überwindung von Distanz mit Aufwand oder mit Kosten verbunden und daher verhält es sich mit der räumlichen Relevanz meistens so, dass nahe gelegene Geobjekte für den Nutzer relevanter sind als ferner gelegene. In dieser Arbeit gilt daher die Annahme, dass mit steigender räumlicher Distanz zwischen Nutzer und Geobjekt die räumliche Relevanz rR_I des Geobjektes sinkt (Abbildung 3.1).

Hat der Nutzer beispielsweise Interesse an einem Hotel, so wird er, gemäss die-



Räumlich Nahes ist in den meisten Fällen relevanter als Fernes...

Abbildung 3.1: Relevanz in Abhängigkeit der Distanz zwischen Nutzer und Objekt

ser Annahme, das räumlich nächstgelegene wählen, sofern alle anderen Kontext-Dimensionen ausser Acht gelassen werden (*ceteris paribus*).

Im Abschnitt 3.4 wird anhand eines Beispiels erläutert, dass die räumliche Relevanz rR_I eines Geobjektes mit zunehmender Entfernung vom Nutzer nicht immer abnehmen muss und die obige Annahme nicht immer zutrifft. Auch die folgenden Modelle der räumlichen, zeitlichen und thematischen Relevanz-Typen sind nicht zwingend

allgemeingültig, sondern gelten nur im Rahmen dieser Arbeit.

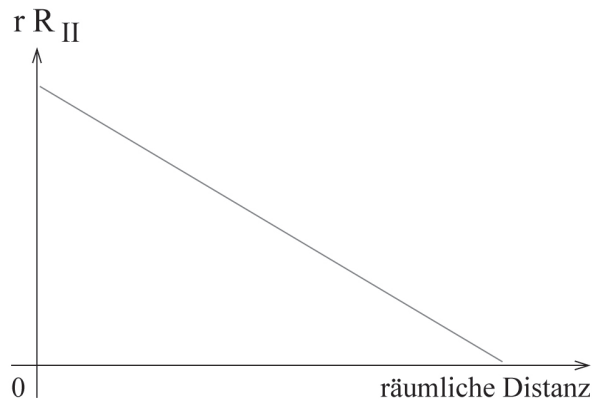
3.1.3 Räumlicher Relevanztyp II (rR_{II})

Die zweite Art der räumlichen Relevanz ist eine Funktion der Entfernung zwischen zwei oder mehreren gleichartigen Geobjekten.

$$R_{II} = f(\text{Distanz}_{GO-GO}) \quad (3.2)$$

Die räumliche Distanz zwischen zwei oder mehreren Geobjekten kann auch unabhängig von der Position des Nutzers im geographischen Raum von Bedeutung sein. Befindet sich ein Objekt in räumlicher Nähe zu einem oder mehreren gleichartigen Objekten, so kann dieser Cluster aufgrund des grösseren Angebotes dem einzelnen Objekt eine höhere Relevanz zukommen lassen. In dieser Arbeit gilt daher die Annahme, dass mit sinkender räumlicher Entfernung zwischen zwei oder mehreren gleichartigen Geobjekten die räumliche Relevanz rR_{II} des Geobjektes steigt (Abbildung 3.2).

Interessiert sich ein Nutzer beispielsweise für eine Bar, so kann eine Ansammlung



Die räumliche Nähe zu benachbarten gleichartigen Objekten hat das Objekt selbst einen positiven Effekt...

Abbildung 3.2: Relevanz in Abhängigkeit Distanz zwischen gleichartigen Objekten

von Bars, wie sie häufig in einem städtischen Vergnügungsviertel anzutreffen ist, einer sich darin befindlichen Bar eine stärkere Anziehungskraft auf den Nutzer verleihen als wenn sie alleinstehend wäre. Die Möglichkeit, eine grössere Auswahl zu haben und gegebenenfalls eine alternative Bar besuchen zu können, kann unter Umständen (eine Bar ist überfüllt oder einer anderen Bar sind kurzfristig und unerwartet die Getränke ausgegangen), nützlich sein. Der durch eine solche Ansammlung

an Geoobjekten resultierende Mehrwert kann durch die räumliche Relevanz rR_{II} ausgedrückt werden.

Nicht zu verwechseln ist dieses Prinzip der Nachbarschaften mit dem Gravitations-Prinzip. Dadurch, dass ein Objekt über mehr Kapazitäten verfügt als ein anderes (beispielsweise mehr Sitzplätze bei einem Kino), kann ein Objekt auch eine höhere Relevanz erhalten als ein Geoobjekt mit geringerer Kapazität. Doch dies ist eine nicht-räumliche Eigenschaft des Objekts und lässt sich nicht hier eingliedern.

Auch die räumliche Nähe verschiedenartiger Geoobjekte kann von Bedeutung sein, denn wenn ein Nutzer mehrere Kommissionen zu erledigen hat, so ist nicht die Distanz zum nächsten Objekt sondern die gesamthaft zurückzulegende Strecke von Interesse. Dieses Optimierungsproblem bedarf aber neben der räumlichen Verteilung auch der Absichten des Nutzers und wird daher nicht zu den räumlichen Relevanztypen gezählt.

3.1.4 Räumlicher Relevanztyp III (rR_{III})

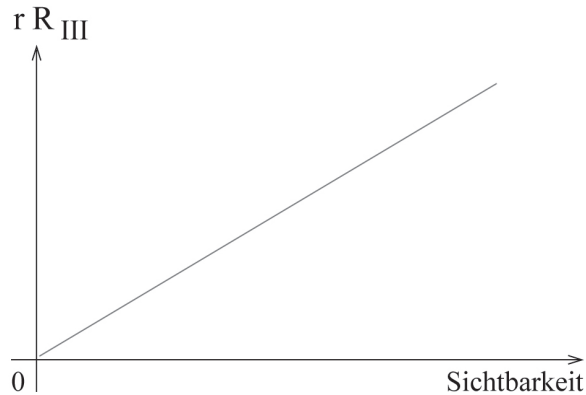
Die dritte Art der räumlichen Relevanz ist eine Funktion der Sichtbarkeit eines Geoobjektes.

$$R_{III} = f(\text{Sichtbarkeit}_{GO}) \quad (3.3)$$

Manche Gebäude vermögen dank ihrer Exposition die Aufmerksamkeit des Nutzers auf sich zu ziehen indem sie aus ihrer Umgebung hervorragen. Solche Orientierungshilfen (Landmarks, [Alexander Zipf(2003)]) sind dann besonders relevant, wenn sich ein Nutzer in einer ihm unbekanntem Gegend orientieren muss. Auch die Sichtbarkeit eines Geoobjektes lässt sich alleine aufgrund der Anordnung des Objektes im Raum bestimmen und ist daher auch eine Art der räumlichen Relevanz. In dieser Arbeit gilt daher die Annahme, dass mit steigender Sichtbarkeit auch die räumliche Relevanz rR_{III} des Geoobjektes steigt (Abbildung 3.3).

3.2 Zeitlicher Kontext

Das kontextsensitive System *Deep Map* verwendet den zeitlichen Kontext eines Touristen, um ihm unter Berücksichtigung von seinem Zeitbudget, Besichtigungsrouten vorzuschlagen ([Malaka und Zipf(2000)]). Das System *COMPASS* verwendet ebenfalls die Zeit, um den Nutzer mit relevanten Informationen zu versorgen ([?]). Dabei wird eine zeitliche Relevanz in Abhängigkeit der Zeit, die zwischen dem letzten Besuch einer Objektart vergangen ist, definiert. Objekte, welche kürzlich besucht



Sichtbare Objekte dienen der Orientierung und können daher die Relevanz eines Objektes erhöhen. . .

Abbildung 3.3: Relevanz in Abhängigkeit der Sichtbarkeit

worden sind, erhalten dabei einen geringeren Relevanz-Wert als zeitlich weiter entfernte Objekte. Die Idee dahinter ist, dass beispielsweise ein Nutzer, welcher gestern in einem griechischen Restaurant das Nachtessen eingenommen hat, heute nicht schon wieder in einem griechischen Restaurant zu Abend essen möchte. Als letztes Beispiel soll das System *SAiMotion* genannt werden ([Hermann und Heidmann(2002)]). Um auf einem Messegelände einem Nutzer eine persönliche Route vorschlagen zu können, bedarf dieses System neben den persönlichen Interessen des Besuchers auch seinen zeitlichen Terminvereinbarungen im Verlaufe des Besuchs. Die Idee dabei ist, dass dem Nutzer nur Stände vorgeschlagen werden sollen, wenn für deren Besichtigung auch genügend Zeit vorhanden ist.

Während in dem vorgängigen Unterkapitel *Räumliche Relevanz* der Einfluss des geographischen Raumes auf die Relevanz von Geoobjekten diskutiert worden ist, wird hier die Bedeutung der Kontext-Dimension *Zeit* behandelt. Dieses Unterkapitel befasst sich also mit der Modellierung der zeitlichen Relevanz. Zuerst wird erläutert, was unter dem Begriff der *zeitlichen Relevanz* verstanden wird und anschliessend wird ein zeitlicher Relevanztypen vorgestellt. Ausserdem wird anhand eines Modelles demonstriert, wie sich dieser Relevanztyp *mit der Zeit* verändert.

Das Ziel dieses Unterkapitels ist es, die Wirkung der Kontext-Dimension *Zeit* auf die zeitliche Relevanz zu beschreiben und zu modellieren.

3.2.1 Definition

Die zeitliche Relevanz bezieht sich analog zur räumlichen Relevanz alleine auf eine Kontext-Dimension, in diesem Falle auf die *Zeit*. Somit beschreibt die zeitliche Relevanz die Bedeutung eines Objektes alleine hinsichtlich seines zeitlichen Kontextes. Die Zeit kann unterschiedlich grob verstanden werden, [Reichenbacher(2004)] spricht in diesem Zusammenhang von unterschiedlicher *Granularität*. So kann zwischen drei Stufen unterschieden werden, wobei jede dieser Stufen einen Einfluss auf die Relevanz eines Objektes haben kann:

- Mikroskala (Tageszeit): Die Tageszeit hat insofern einen Einfluss auf die Relevanz eines Geoobjektes, als dass über sie bestimmt werden kann, ob ein Einkaufsgeschäft geöffnet hat oder ob eine Theateraufführung bereits begonnen hat. Ob ein Geschäft oder ein Theater zeitlich relevant ist, hängt somit von der Tageszeit ab. So ist ein Kino am Morgen zeitlich weniger relevant als ein Einkaufsgeschäft.
- Mesoskala (Wochentag): Manchmal kann auch der Wochentag von Bedeutung sein, denn Öffnungszeiten richten sich nicht nur nach der Tageszeit, sondern auch nach dem Wochentag (z.B. sind Restaurants am Wochenende häufig geöffnet, während gewöhnliche Geschäfte geschlossen sind). Zur Bestimmung der zeitlichen Relevanz eines Restaurants kann somit zusätzlich zur Uhrzeit auch der Wochentag herangezogen werden.
- Makroskala (Jahreszeit): Auch die Jahreszeit kann bei bestimmten Objekten eine Rolle spielen. So hängt die Relevanz eines offenen Schwimmbades oder einer offenen Eisbahn von der Jahreszeit ab. Zur Bestimmung der zeitlichen Relevanz muss bei gewissen saisonalen Objekten neben der Uhrzeit und dem Wochentag ebenfalls die Jahreszeit berücksichtigt werden.

In dieser Arbeit wird lediglich die zeitliche Relevanz in Abhängigkeit der Tageszeit berücksichtigt, als ob der Wochentag und die Jahreszeit keinen Einfluss auf die Relevanz des Geoobjektes hätten (wie beispielsweise bei einem Theater).

Nach [Worboys und Duckham(2004)] lässt sich die Welt der Entitäten in *continuants* (Objekte) und *occurents* (Ereignisse, Prozesse) einteilen. Beide dieser Konzepte lassen sich zeitlich fixieren, beispielsweise über die Ladenöffnung eines Geschäftes (*continuant*) oder die Dauer eines Vulkanausbruches (*occurent*). In dieser Arbeit werden Objekte und keine Prozesse bewertet, in dem über ein zeitliches Attribut

des Geoobjektes auf dessen zeitliche Fixierung zugegriffen wird. So ist eine Theateraufführung zwar ein Ereignis und kein Objekt, jedoch gehört eine solche Aufführung einem Theater an wodurch das entsprechende Theater bewertet werden kann.

In dieser Arbeit wird zwischen Geoobjekten unterschieden, welche durch einen *Zeitpunkt* zeitlich fixiert sind (Start einer Theateraufführung) und Geoobjekten, welche durch eine *Zeitspanne* zeitlich fixiert sind (Öffnungszeiten eines Restaurants). Anders ausgedrückt wird zwischen *einfach terminierten* Objekten, bei welchen somit nur ein Ereignis (*Startzeit*) interessiert und *zweifach terminierten* Objekten, bei welchem zwei Ereignisse (*Öffnungszeit* und *Schlusszeit*) von Interesse sind, unterschieden.

3.2.2 Zeitlicher Relevanztyp (zR)

Die zeitlichen Relevanz zR ist eine Funktion der Zeit, die bis zum Eintreten eines Ereignisses (Startzeit, Öffnungszeit oder Endzeit), welches wie erwähnt einem Objekt angehört, vergeht (Δt). Als *Eintrittszeit* wird in dieser Arbeit der Zeitpunkt bezeichnet, an welchem dieses Ereignis stattfindet.

$$\Delta t = \text{Tageszeit} - \text{Eintrittszeit} \quad (3.4)$$

Nicht anwenden lässt sich diese Art der Relevanz auf Geoobjekte, welche zeitlich nicht fixiert sind, wie beispielsweise ein Naturpark oder ein Aussichtsturm.

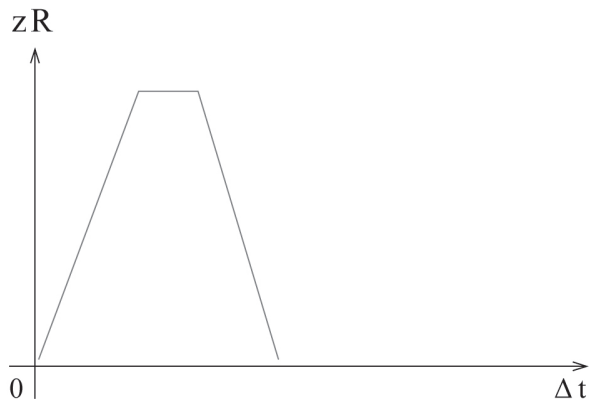
$$zR = f(\Delta t) \quad (3.5)$$

Die zeitliche Relevanz zR von *einfach terminierten* Geoobjekten verhält sich im Gegensatz zu rR_I nicht so, dass mit einer niedrigeren zeitlichen Entfernung der Startzeit eine grössere zeitliche Relevanz zR einhergeht. Vielmehr gilt bei den meisten Terminen, sich eher ein wenig im voraus und weder *à point* noch viel zu früh einzutreffen. In dieser Arbeit gilt daher die Annahme, dass einfach terminierte Geoobjekte kurz vor der Startzeit (kleines Δt) die höchste zeitliche Relevanz zR aufweisen und diese Relevanz anschliessend mit zunehmendem Δt wieder abnimmt (Abbildung 3.4).

Interessiert sich beispielsweise ein Nutzer für ein Theaterbesuch, so wäre es ideal, sich einige Minuten vor der Aufführung im Theater einzutreffen um noch genügend Zeitreserve zu haben (Ticket lösen, sich am Platz einfinden etc.). Je ferner die Aufführung jedoch liegt, desto unbedeutender wird auch das Theater (es wird angenommen, dass der Nutzer in der Wartezeit keine andere Beschäftigung vorsieht). Ebenso unbedeutend ist ein Theater, dessen Aufführung bereits begonnen hat, wenn

also Δt negativ ist (eine negative zeitliche Distanz ist im Gegensatz zur räumlichen Distanz möglich).

Im Gegensatz zu einfach terminierten Geoobjekten, dessen Startzeit es möglichst



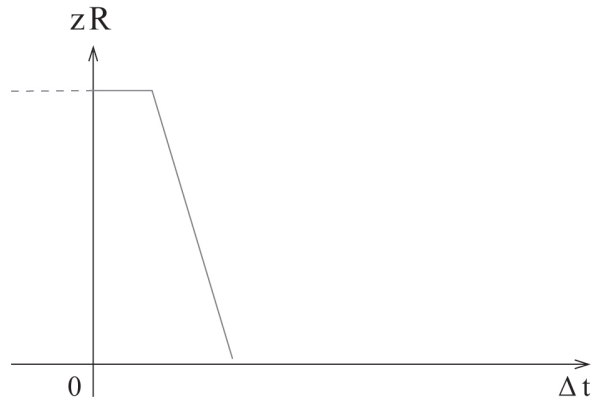
Objekte, deren Ereignis erst in ein paar Minuten startet, sind am relevantesten...

Abbildung 3.4: Relevanz in Abhängigkeit von Δt bis zur Startzeit

ein paar Minuten im voraus zu treffen gilt (s. oben), nimmt die zeitliche Relevanz eines Geschäftes nach Öffnungszeit nicht ab, sondern bleibt konstant bis kurz vor Endzeit. Der Grund liegt darin, dass *zweifach terminierte* Objekte meistens über eine längere *Zeitspanne* geöffnet haben und nicht ein bestimmter *Zeitpunkt* anvisiert werden muss, wie dies bei einfach terminierten Objekten der Fall ist. Ist Δt bis zur Öffnungszeit also negativ, Δt bis Endzeit jedoch positiv, so hat das Geoobjekt geöffnet und ist in zeitlicher Hinsicht relevant. In dieser Arbeit wird ein Geoobjekt, welches in wenigen Minuten seine Türen öffnen wird, zwar zeitlich als weniger relevant erachtet als ein bereits geöffnetes, jedoch zeitlich relevanter eingestuft als eines, welches erst viel später öffnen wird. In dieser Arbeit gilt daher die Annahme, dass zweifach terminierte Objekte dann die höchste zeitliche Relevanz zR aufweisen, wenn Δt bis zur Öffnungszeit negativ oder möglichst klein ist (Abbildung 3.5).

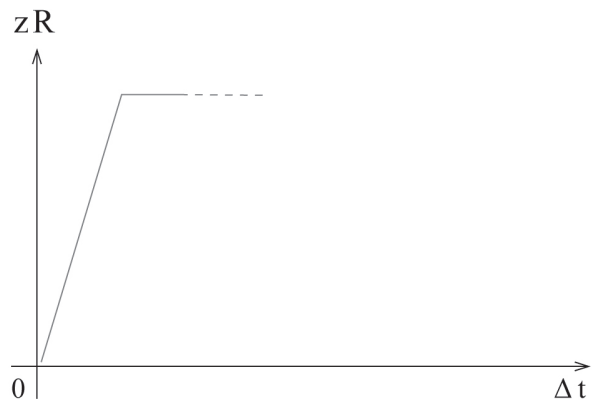
Bei der Endzeit verhält es sich gerade umgekehrt, denn in dieser Arbeit gilt die Annahme, dass ein zweifach terminiertes Objekt dann die höchste zeitliche Relevanz zR aufweist, wenn Δt bis zur Endzeit möglichst gross ist (Abbildung 3.6).

Δt verändert sich permanent und ohne Einwirkung des Nutzers, sodass sich auch die zeitliche Relevanz eines Geoobjektes fortlaufend ändert. Dies hat zur Folge, dass die Relevanzberechnung ein fortlaufender Prozess ist und dies auch bei der Implementierung in OpenJUMP berücksichtigt werden sollte.



Je näher die Eintrittszeit bis zur Ladenöffnung liegt, desto höher die zeitliche Relevanz...

Abbildung 3.5: Relevanz in Abhängigkeit von Δt bis zur Öffnungszeit



Je ferner die Eintrittszeit bis zur Ladenschliessung liegt, desto höher die zeitliche Relevanz...

Abbildung 3.6: Relevanz in Abhängigkeit von Δt bis zur Endzeit

3.3 Thematischer Kontext

Während in den vorangehenden Unterkapiteln 3.1 und 3.2 die Auswirkung der Position eines Geoobjektes in Raum und Zeit auf seine Relevanz diskutiert worden ist, wird in diesem Kapitel der Einfluss der Interessen des Nutzers (bezüglich der Kategorie eines Geoobjektes) behandelt.

Das System COMPASS verwendet ebenfalls die Information der Interessen des Nutzers bezüglich der Kategorie eines Geoobjektes, um ein Geoobjekte diesbezüglich zu bewerten ([?]).

Dieses Kapitel befasst sich mit der Modellierung der thematischen Relevanz. Zuerst wird erläutert, was unter dem Begriff der *thematischen Relevanz* verstanden wird und es wird ein Relevanztyp vorgestellt, welcher sich auf die Kontext-Dimension

Thematik bezieht. Anhand eines Modelles wird veranschaulicht, wie sich der dieser Relevanztyp in Abhängigkeit der *Thematik* verhält.

Das Ziel dieses Unterkapitels ist es, die Wirkung der Kontext-Dimension *Thematik* auf die thematische Relevanz zu beschreiben und zu modellieren.

3.3.1 Definition

Die thematische Relevanz bezieht sich analog zur räumlichen oder zeitlichen Relevanz alleine auf eine Kontext-Dimension, in diesem Falle auf die *Thematik*. Unter der *Thematik* wird in dieser Arbeit die Kategorie eines Geobjektes bezüglich der Kategorie (Hotel, Restaurant, Theater...) und nicht anderer Attribute (wie die Preiskategorie oder die Kapazität) verstanden. Somit beschreibt die thematische Relevanz in dieser Arbeit die Bedeutung eines Geobjektes alleine aufgrund dem Verhältnis zwischen der Kategorie und dem Interesse des Nutzers. Diese Ähnlichkeit (Verwandtschaft) zwischen zwei Begriffen, wobei der eine für die Kategorie des Objektes und der andere für die Interessen des Nutzers steht, lässt sich durch die semantische Distanz ausdrücken ([Rodríguez und Egenhofer(2004)]).

Wie der räumliche oder zeitliche Kontext kann auch der thematische Kontext eines Geobjektes unterschiedlich grob verstanden werden. So lässt sich die Kategorie eines Geobjektes unterschiedlich detailliert abbilden, denn ein *Hotel*, eine *Jugendherberge* oder eine *Pension* sind allesamt Elemente von der Kategorie *Unterkunft*, welche wiederum ein Element der Kategorie *Gebäude* ist. Im Rahmen dieser Arbeit gelten folgende Kategorien: *Hotel*, *Jugendherberge*, *Pension*, *Restaurant*, *Imbiss*, *Bar*, *Kino*, *Theater* und *Disco*. Ein Geobjekt in dieser Arbeit muss also einer dieser Kategorien angehören.

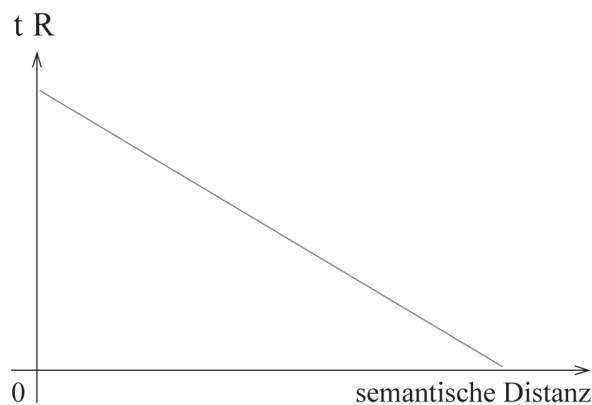
Die Interessen des Nutzers lassen sich durch seine explizite Äusserung oder durch implizit, von diesen Äusserungen abgeleitet, gewinnen. Bei der expliziten Bekanntgabe seiner Interessen deklariert der Nutzer direkt und ausdrücklich, worin seine Interessen bestehen, während bei der impliziten Äusserung seine Interessen aus seinem Nutzerprofil, dem *user model* (-> Quelle!), abgeleitet werden. In dieser Arbeit wird die explizite Vorgehensweise angewendet.

3.3.2 Thematischer Relevanztyp (tR)

Die thematische Relevanz ist eine Funktion der semantischen Distanz zwischen der Kategorie des Geoobjektes (Begriff 1) und der Interessen des Nutzers (Begriff 2).

$$tR = f(\text{semantische Distanz Begriff 1, Begriff 2}) \quad (3.6)$$

Mit zunehmender semantischer Distanz nimmt auch die Relevanz eines Geoobjektes ab (Abbildung 3.9) und dabei gibt es wohl kaum eine Ausnahme. Der Grund liegt darin, dass sich die thematische Relevanz im Gegensatz zur räumlichen beziehungsweise zeitlichen Relevanz direkt nach dem Interesse des Nutzers richtet. Der Nutzer gibt also bekannt, an welcher Kategorie er interessiert ist und anschliessend werden die Geoobjekte entsprechend bewertet. Somit ist es sinnlos, ein Objekt als relevant einzustufen, an welchem der Nutzer gar nicht interessiert ist. Interessiert sich ein



Mit zunehmender Entfernung zwischen den Eigenschaften eines Geoobjektes und den Interessen des Nutzers nimmt die Relevanz ab...

Abbildung 3.7: Relevanz in Abhängigkeit der semantischen Distanz

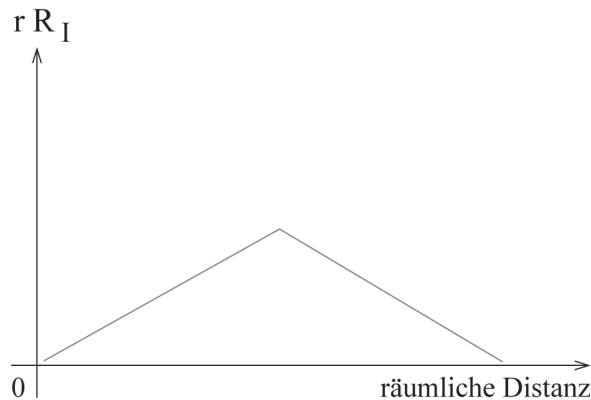
Nutzer beispielsweise für ein Kino, so bekommen alle Kinos unabhängig von ihrer räumlichen oder zeitlichen Entfernung die Höchstnote bezüglich des thematischen Kontextes. Wenn der Nutzer seine Interessen nur bezüglich der Kinos geäussert hat, so dürfte ein Theater zwar für die meisten Nutzer weniger relevant sein als ein Kino, jedoch immer noch relevanter als beispielsweise ein Freibad, weil ein Kino kaum etwas mit einem Freibad zu tun hat. Folglich ist die semantische Distanz zwischen den Begriffen Kino und Freibad meistens grösser als zwischen Kino und Theater.

3.4 Abweichungen und Modifikationen

Die Zusammenhänge zwischen den Kontext-Dimensionen und den entsprechenden Relevanztypen könnten, wie oben bereits erwähnt, auch anders modelliert werden und sind daher nicht als allgemeingültige Gesetzmässigkeiten zu verstehen. Beispielsweise gilt nicht für jeden Nutzer, dass mit zunehmender Entfernung eines Geoobjektes dessen räumliche Relevanz rR_I abnehmen muss. Das folgende Beispiel soll dies verdeutlichen:

Für einen Fahrradfahrer, welcher eine zwei-stündige Tour unternehmen will und auf halbem Weg in einem Restaurant den Durst löschen will, sind die Restaurants ab einer Entfernung von einer Stunde Fahrzeit wichtiger als jene, die dem Startpunkt näher liegen, denn zu Beginn seiner Fahrt wird er weder durstig noch erschöpft sein. Auch die Restaurants, welche weiter entfernt sind als der halbe Weg, sind wiederum weniger stark zu gewichten, denn ab dieser Entfernung wird der Durst und die Erschöpfung des Nutzers wiederum zu gross sein. Aus diesem Grund sieht auch das Modell dieses Beispiels (Abbildung 3.8) entsprechend anders aus als in Abbildung 3.1.

Auch die zeitliche Relevanz zR muss nicht für jeden Nutzer gleich modelliert wer-



Es gibt auch Umstände, unter welchen die räumliche Relevanz mit zunehmender Entfernung nicht linear sinkt...

Abbildung 3.8: Räumliche Relevanz für einen Velofahrer

den, wie folgendes Beispiel zeigen soll:

Ist aus den Absichten des Nutzers bekannt, dass er vor dem Einkauf noch einen Kaffee trinken möchte, so muss dafür auch Zeit eingeplant werden, wodurch auch Einkaufsgelegenheiten wieder relevant werden, welche nicht in unmittelbarer zeitlicher Umgebung liegen. Solche Aktivitätssequenzen bedingen somit einen grösseren Zeitmasstab, weil nicht nur zeitlich naheliegende Geoobjekte von hoher Bedeutung sind, und sind ausserdem als Optimierungsproblem äusserst komplex zu berücksich-

tigen.

Für jeden Nutzer können sich die einzelnen Relevanztypen somit unterschiedlich

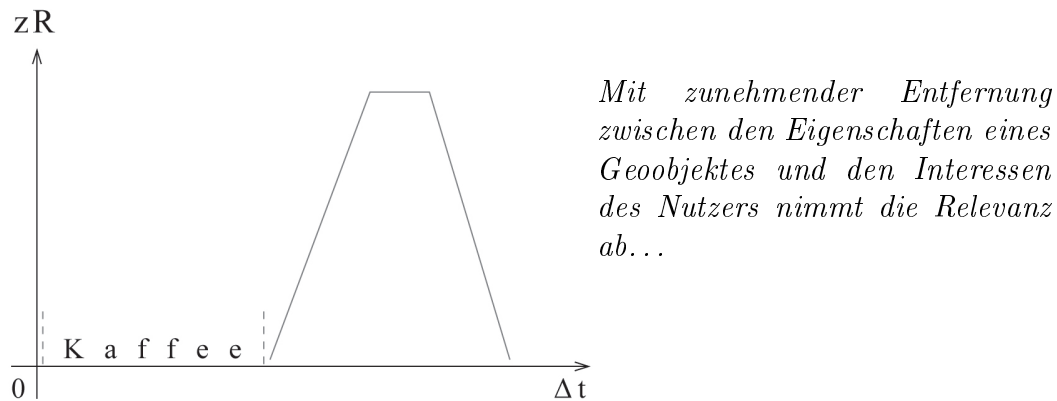
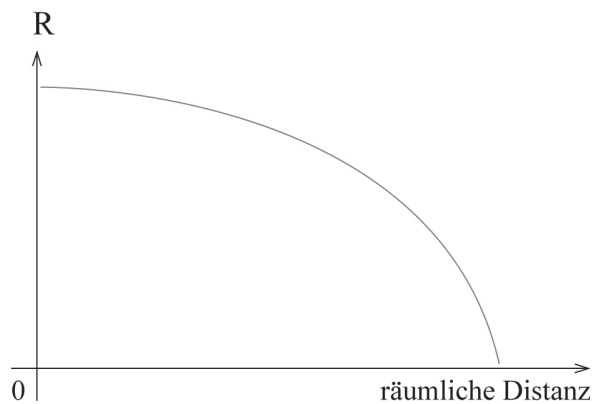


Abbildung 3.9: Relevanz in Abhängigkeit der semantischen Distanz

verhalten. Es ist klar, dass die in dieser Arbeit auf Annahmen basierenden Modellen den Nutzer nicht in seiner komplexen Individualität berücksichtigen können und Annahmen diese fehlenden Kontext-Informationen (z.B. seine Gefühlslage) kompensieren müssen. Dennoch kann davon ausgegangen werden, dass diese Annahmen für die meisten Nutzer und unter gewöhnlichen Umständen zutreffen.

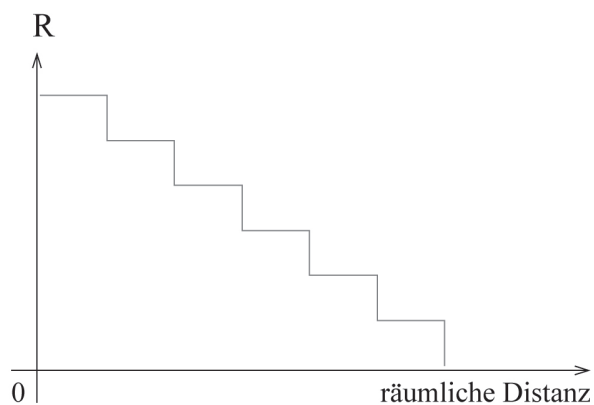
Die obig vorgestellten Modelle der räumlichen Relevanz waren stets linear, denn die lineare Form beschreibt den einfachsten Fall der Modellierung der räumlichen Relevanz, nämlich dass sich die Relevanz mit zunehmender Distanz linear und stetig verändert. Dies macht dann Sinn, wenn sich der Aufwand zur Überwindung von Distanz ebenfalls linear verhält. Dieser Verlauf muss jedoch nicht immer linear sein (Abbildung 3.10). Falls die Kosten für die Distanzüberwindung oder die damit verbundenen Mühen des Nutzers nicht linear sind, würde eine exponentielle oder quadratische Funktion die Entwicklung der räumlichen Relevanz in Abhängigkeit der Distanz besser abbilden.

Ausserdem ist auch eine Diskretisierung des Verlaufes möglich (Abbildung 3.11). Dies bedeutet, dass die Relevanz-Werte gemäss Modell gerechnet und anschliessend klassiert werden, wodurch die Werte zwar ungenauer dafür aber übersichtlicher werden.



Nehmen die Kosten für die Distanzüberwindung nicht-linear zu, so nimmt die Relevanz entsprechend ab. . .

Abbildung 3.10: Nicht-linear abnehmende räumliche Relevanz



Eine Klassierung der Relevanz ist auch denkbar. . .

Abbildung 3.11: Diskretisierte räumliche Relevanz

3.5 Fazit und Ausblick

In diesem Kapitel wurde gezeigt, wie sich die drei Kontext-Dimensionen Raum, Zeit und Thematik isoliert auf die Relevanz eines Geoobjektes auswirken können. Die modellierten Zusammenhänge zwischen den Kontext-Dimensionen und der entsprechenden Relevanz basieren dabei auf Annahmen und sind daher nicht als allgemeingültige Gesetzmässigkeiten zu verstehen.

Im folgenden Kapitel werden diese drei Kontext-Dimensionen miteinander kom-

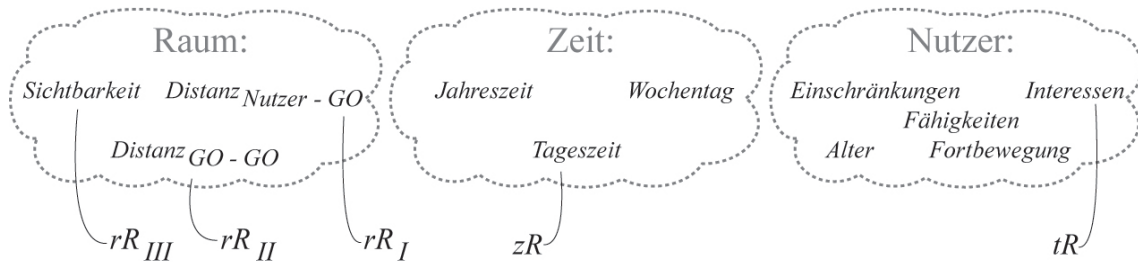


Abbildung 3.12: Zusammenfassung

biniert werden, um komplexere Relevanztypen zu erhalten, welche die individuelle Ausprägung des Nutzers stärker integrieren können.

Kapitel 4

Erweiterte Modellierung der Relevanz

Im vorgängigen Kapitel ist für die drei Kontext-Dimensionen Raum, Zeit und Nutzer einzeln untersucht worden, wie sich die Relevanz in Abhängigkeit des jeweiligen Kontextes verhält. In diesem Kapitel werden diese Kontext-Dimensionen miteinander kombiniert, um übergeordnete Relevanztypen zu erhalten, welche sich stärker den individuellen Eigenschaften des Nutzers anpassen. Dieser Zusammenhang zwischen den (kombinierten) Kontext-Dimensionen und der entsprechenden Relevanz soll in einem Modell veranschaulicht werden.

Das Ziel dieses Kapitels ist die qualitative Modellierung der Relevanz in Abhängigkeit der kombinierten Kontext-Dimensionen, um den Zusammenhang zwischen Relevanz und Kontext aufzuzeigen.

4.1 Raum und Nutzer:

personalisierte räumliche Relevanz (prR)

Bereits im Kapitel 3.1.2 wird die Annahme getroffen, dass die räumliche Relevanz rR_I eines Geoobjektes mit zunehmender Entfernung zum Nutzer linear und für jeden Nutzer in gleichem Masse abnehmen soll. Bis zu welcher Entfernung jedoch ein Geoobjekt für den Nutzer relevant ist oder wie schnell die räumliche Relevanz mit zunehmender Entfernung abnimmt, ist vom Nutzer abhängig und soll in diesem Abschnitt berücksichtigt werden.

Die Kontext-Dimension *Nutzer* umfasst äusserst viele Komponenten, unter anderem seine Interessen, Fähigkeiten oder Einschränkungen. In dieser Arbeit und im Zusammenhang mit der personalisierten räumlichen Relevanz beschränkt sich die Kontext-Dimension des Nutzers alleine auf das Alter und Fortbewegungsmittel des Nutzers.

Auch die Kontext-Dimension Raum hat unterschiedliche Ausprägungen (s. Kapitel 3.1). In dieser Arbeit und im Zusammenhang mit *prR* wird unter dem Raum die räumliche Entfernung zwischen dem Geobjekt und dem Nutzer verstanden.

In dieser Arbeit wird die personalisierte räumliche Relevanz *prR* somit in Abhängigkeit der räumlichen Entfernung vom Geobjekt zum Nutzer und seines Alters respektive seines Fortbewegungsmittels gestellt. Mit anderen Worten werden das Alter und das Fortbewegungsmittel des Nutzers herangezogen, um *rR_I* um die individuellen Eigenschaften des Nutzers zu erweitern.

$$prR = f(Distanz_{Nutzer-GO}, Alter_{Nutzer}, Fortbewegungsmittel_{Nutzer}) \quad (4.1)$$

In Abbildung 4.1 ist abgebildet, wie in dieser Arbeit das Verhältnis zwischen der personalisierten räumlichen Relevanz und des Alters des Nutzers respektive der räumlichen Entfernung angenommen wird. In dieser Arbeit gilt die Annahme, dass ein

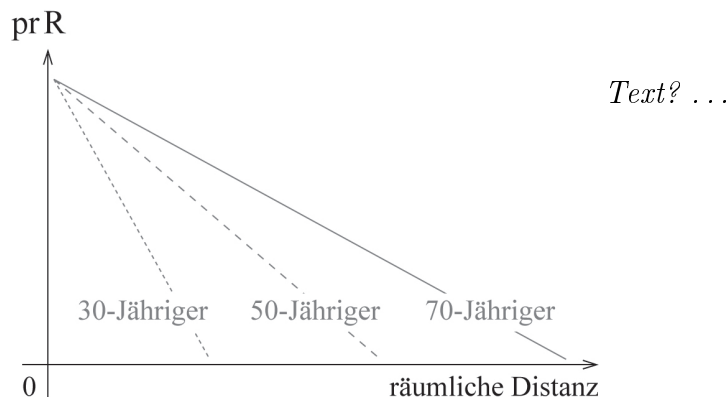


Abbildung 4.1: Konkrete Zeitan-gaben

jüngerer Nutzer weniger Mühe bekundet um eine gleiche Distanz zu überwinden wie ein älterer Nutzer. Dieser geringerer Aufwand widerspiegelt sich dementsprechend auch in der Relevanz eines Geobjektes, weil für einen jüngeren Nutzer Geobjekte noch relevant sein können, welche für einen älteren Nutzer bereits zu weit weg und somit irrelevant sind (Abbildung 4.1). Ebenso gilt die Annahme, dass das Fortbewegungsmittel des Nutzers die maximale Entfernung zwischen Nutzer und Geobjekt

insofern verändert, als dass ein Automobilist für dieselbe Distanz weniger Aufwand aufbringen muss als ein Fussgänger (es wird angenommen, dass der Aufwand der Benzinkosten die körperliche Anstrengung nicht übertrifft). Daher sind für einen Automobilisten auch noch Objekte relevant, welche für einen Fussgänger aufgrund der zu grossen räumlichen Entfernung bereits nicht mehr von Bedeutung sind.

4.2 Raum und Nutzer: Wegzeit-Relevanz (wzR)

Die Überwindung von räumlicher Distanz benötigt neben körperlicher Anstrengung und Kosten auch Zeit (*Reisezeit* oder *Wegzeit*). Je nachdem, wie schnell der Nutzer unterwegs ist, kann er den geographischen Raum mehr oder weniger schnell überwinden. Die Wegzeit ist somit eine Funktion der räumlichen Entfernung zwischen Geoobjekt und Nutzer und der Geschwindigkeit des Nutzers, wobei sich die Geschwindigkeit in dieser Arbeit nach seinem Alter und Fortbewegungsmittel richtet.

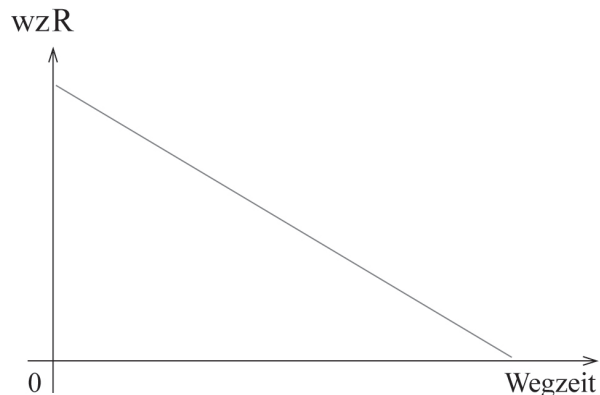
$$Wegzeit = f(\text{räumlicher Entfernung, Geschwindigkeit}_{\text{Nutzer}}) \quad (4.2)$$

Die Wegzeit-Relevanz in dieser Arbeit ist eine Funktion der Wegzeit und somit abhängig von den Kontext-Dimensionen *Raum* und *Nutzer*.

$$wzR = f(Wegzeit) \quad (4.3)$$

In den allermeisten Fällen ist die Zeit ein beschränktes Gut und sollte sinnvoll eingeteilt werden, weil die meisten Nutzer keine Zeit im Überfluss haben (getreu dem Motto: „Zeit ist Geld“). Aus diesem Grund sind in der Regel Geoobjekte, welche zeitlich näher liegen, relevanter als fernere. In dieser Arbeit gilt somit die Annahme, dass mit steigender zeitlicher Entfernung eines Geoobjektes auch seine Wegzeit-Relevanz sinkt (Abbildung 4.2). So wird in dieser Arbeit angenommen, dass sich ein Nutzer, der sich beispielsweise für ein Theater interessiert, für das zeitlich naheliegendste Theater entscheiden wird, sofern alle anderen Kontext-Dimensionen ausser Acht gelassen werden (*ceteris paribus*).

In dieser Arbeit leitet sich die Wegzeit aus der räumlichen Entfernung zwischen Geoobjekt und Nutzer, seinem Fortbewegungsmittel und seinem Alter ab (wie auch schon prR). Über eine Strassenattributierung (zulässige Geschwindigkeit, Baustellen etc.) könnte die Wegzeit auch davon abhängig gemacht werden, wie schnell sich der Nutzer auf einem Strassenabschnitt bewegen könnte (vgl. Autobahn versus Quar-



Zeitlich Ferneres ist in den meisten Fällen weniger relevant als Näheres. . .

Abbildung 4.2: Relevanz in Abhängigkeit der Wegzeit

tierstrasse). Auch wenn dies der Wegzeit und somit auch der Wegzeit-Relevanz mehr Sinn verleihen würde, wird dieser Ansatz hier nicht umgesetzt, da in dieser Arbeit lediglich die Kontext-Dimensionen Raum, Zeit und Nutzer verwendet werden sollen.

4.3 Raum, Zeit und Nutzer: korrigierte zeitliche Relevanz (kzR)

Ein zeitlich fixiertes Geoobjekt weist einerseits eine Relevanz bezüglich seiner zeitlichen Entfernung bis zur Eintrittszeit auf (zR , s. Kapitel 3.2.2), andererseits eine Relevanz bezüglich seiner Wegzeit (wzR , s. Kapitel 4.2) auf. Diese beiden Relevanztypen sollen in dieser Arbeit nicht isoliert betrachtet, sondern miteinander kombiniert werden. Diese Kombination ist sinnvoll, denn ist ein Restaurant noch 20 Minuten lang geöffnet, die Wegzeit dahin beträgt jedoch 30 Minuten, so ist das Restaurant nicht relevant, weil es bis zum Eintreffen des Nutzers bereits geschlossen sein wird (s. Abbildung 3.6). Somit sollte zuerst die Wegzeit von Δt subtrahiert werden, denn während der Distanzüberwindung hat sich auch Δt entsprechend verändert. Diese Differenz kann als $\Delta t_{effektiv}$, respektive als Wartezeit (positive Differenz) oder Verspätung (negative Differenz) übersetzt werden.

$$\Delta t_{effektiv} = \Delta t - \text{Wegzeit} \quad (4.4)$$

Die um die Wegzeit korrigierte zeitliche Relevanz kzR ist somit eine Funktion der Tageszeit und der Wegzeit und kombiniert die drei Kontext-Dimensionen Raum, Zeit und Nutzer miteinander. Mit anderen Worten wird die zeitliche Relevanz zR

um die Wegzeit erweitert, um die korrigierte zeitliche Relevanz stärker dem Nutzer anzupassen.

$$kzR = f(\Delta t_{\text{effektiv}}) \quad (4.5)$$

$\Delta t_{\text{effektiv}}$ kann nun erneut in die Modelle, welche in den Abbildungen 3.4 bis 3.6 veranschaulicht sind, eingesetzt werden, um die korrigierte zeitliche Relevanz abzubilden. Dabei gelten dieselben Annahmen wie bei der zeitlichen Relevanz zR .

4.4 Abweichungen und Modifikationen

Wie auch schon im vorhergehenden Kapitel gilt in diesem Kapitel, dass die Zusammenhänge zwischen den Kontext-Dimensionen und den Relevanztypen auch anders modelliert werden könnten (nicht-linear, diskret etc.) und dass es unzählige Ausnahmefälle gibt, wo die Modelle nicht zutreffen. In dieser Arbeit allerdings sollen die gezeigten Modelle gelten und im folgenden Kapitel auch formalisiert werden.

4.5 Fazit und Ausblick

Durch die Kombinationen der Kontext-Dimensionen *Raum* und *Zeit* mit der Kontext-Dimension *Nutzer* können die Relevanztypen prR , wzR und kzR mehr auf die individuellen Eigenschaften des Nutzers (Alter und Fortbewegungsmittel) abgestimmt werden. Dennoch basieren auch diese Relevanztypen auf Annahmen, welche im Rahmen dieser Arbeit gegeben sind. Könnten alle Eigenschaften des Nutzers und der Geoobjekte exakt erfasst und berücksichtigt werden, könnte für jeden Nutzer die Relevanz jedes Objektes genau bestimmt werden und es bräuchte keine Annahmen mehr. Durch die Bildung von prR , wzR und kzR nähert man sich, trotz der bestehenden Abhängigkeit von Annahmen, ein wenig diesem Zukunftsszenario.

In der Abbildung 4.3 sind alle Relevanztypen vorgestellt, welche in diesem und im vorangehenden Kapitel beschrieben worden sind.

Die rot markierten Relevanztypen in der Abbildung 4.3 werden im kommenden Kapitel *Formalisierung* formalisiert und integriert, um die Relevanz eines Geoobjektes in Bezug auf seinen Kontext quantitativ bestimmen zu können.

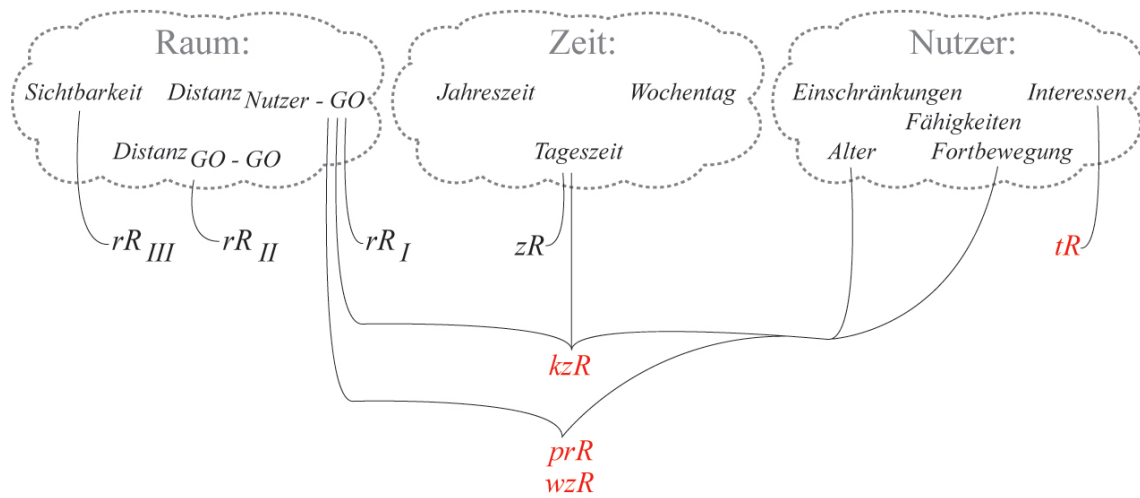


Abbildung 4.3: Zusammenfassung

Kapitel 5

Formalisierung der Relevanz

Während das vorgängige Kapitel den Zusammenhang zwischen den drei Kontext-Dimensionen Raum, Zeit und Nutzer und den entsprechenden Relevanzarten modellierte, geht es in diesem Kapitel nun darum, diese qualitativen Modelle konkret zu formalisieren. Formalisiert werden sollen tR , prR , wzR , und kzR . Durch diese Formalisierung kann jedem Objekt ein Relevanz-Wert zugeordnet werden, wodurch das Konzept der *unscharfen Mengen* oder von *fuzzy sets* unterstützt wird ([Zadeh(1965)]). Dabei wird erneut jede der vier Relevanztypen zuerst separat formalisiert um anschliessend mit den anderen Arten kombiniert zu werden.

Ziel dieses Kapitels ist die Herleitung einer Funktion, welche alle vier Relevanzarten berücksichtigt und so die Relevanz jedes Geoobjekts bezüglich aller drei Kontext-Dimensionen quantifizieren lässt. Die entsprechende Formel soll später implementiert werden.

5.1 Personalisierte räumliche Relevanz

Wie im vorherigen Kapitel *Modellierung* diskutiert, gilt in dieser Arbeit die Annahme, dass mit steigender räumlicher Entfernung zwischen Nutzer und Geoobjekt auch die räumliche Relevanz des entsprechenden Geoobjektes sinkt. Dieser Zusammenhang lässt sich gemäss Modell (s. Abbildung 3.1) mit linearen Funktion veranschaulichen. Somit lässt sich die räumliche Relevanz rR_I durch die allgemeine Geradengleichung ausdrücken. Die allgemeine Gleichung einer Geraden lautet:

$$y = m * x + q \tag{5.1}$$

Die Ordinate beschreibt den Y-Wert der Geraden, welche mit dem räumlichen Relevanz-Wert R gleichgesetzt werden kann. Wie stark die Relevanz mit steigender räumlicher Distanz abnimmt, kann mit der Steigung m der Geraden ausgedrückt werden. Die Abszisse steht für die räumliche Entfernung zwischen Nutzer und Geoobjekt und q beschreibt den den Maximalwert, den R für den Fall, dass die räumliche Distanz null Meter beträgt, annehmen kann. R ist die Unbekannte und soll höchstens den Wert eins und mindestens den Wert null aufweisen, woraus sich ergibt, dass die Konstante q den Wert eins annimmt. Nicht ganz so trivial ist die Bestimmung der Konstanten m .

Wie stark die Relevanz in Abhängigkeit von der Entfernung zwischen Nutzer und Geoobjekt abnimmt, ausgedrückt durch die Steigung m , könnte theoretisch allgemein bestimmt werden, indem angenommen wird, dass die maximale Entfernung 500 Meter betragen dürfe und alle weiter entfernten Geoobjekte für R den Wert null erhielten (Abbildung 5.1).

Dadurch liesse sich m leicht bestimmen und mit der räumlichen Distanz als einziger

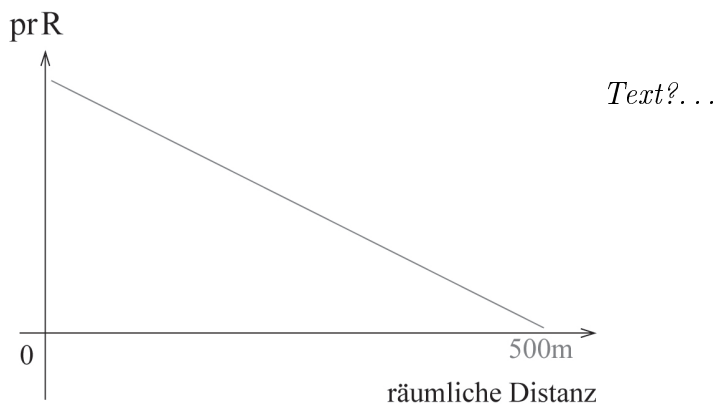


Abbildung 5.1: Konkrete Zeitangaben

Variablen liesse sich gemäss Formel 5.2 schliesslich auch R quantifizieren.

$$\text{räumliche Relevanz} = -\frac{1}{500} * \text{räumliche Distanz} + 1 \quad (5.2)$$

Bei der personalisierten räumlichen Relevanz allerdings beträgt die maximale Entfernung nicht allgemein 500 Meter, sondern in dieser Arbeit wird angenommen, dass sie sich nach dem Alter und dem Fortbewegungsmittel des Nutzers richtet. Beim Alter des Nutzers wird in dieser Arbeit zwischen drei Kategorien unterschieden, nämlich *jünger als 50*, *zwischen 50 und 65* und *älter als 65*. Das Fortbewegungsmöglichkeit wird in die Kategorien *zu Fuss*, *mit dem Fahrrad* und *mit dem Auto*

unterteilt. Somit lautet die Funktion der Steigung m wie folgt:

$$\frac{\Delta y}{\Delta x} = -\frac{1}{\text{maximale Entfernung}} = m \quad (5.3)$$

Dabei wird für jede Kombination der beiden Kategorien entschieden, welches die maximale Entfernung (in Metern) zwischen Nutzer und Geoobjekt sein soll (Δx). Δy entspricht der Konstante q , welche den maximalen Relevanz-Wert ($= 1$) ausdrückt. Das Minuszeichen rührt daher, dass die Gerade im Modell eine negative Steigung aufweist. Sobald sich der Nutzer dem System anmeldet und sein Alter und seine Aktivität bekannt gibt, kann m berechnet und als Konstante in der Formel 5.4 eingesetzt werden. Die Werte für m , welche in dieser Arbeit angenommen werden und empirisch noch nicht geprüft sind, sind der Tabelle 5.1 zu entnehmen.

Die Bestimmung der personalisierten räumlichen Relevanz (Formel 5.4), welche im

		Fortbewegung		
		zu Fuss	Fahrrad	Auto
Alter	jünger/gleich 50	$-\frac{1}{500}$	$-\frac{1}{1600}$	$-\frac{1}{3000}$
	zwischen 50 und jünger/gleich 65	$-\frac{1}{400}$	$-\frac{1}{1500}$	$-\frac{1}{3000}$
	älter als 65	$-\frac{1}{300}$	$-\frac{1}{1400}$	$-\frac{1}{3000}$

Tabelle 5.1: m -Werte

kommenden Kapitel implementiert werden soll, ist somit von den Variablen m und $rdist$ abhängig. Dabei wird die Entfernung in Metern und entlang von Netzwerken (Strassennetz) gemessen. Die Idee dahinter ist, dass sich ein Nutzer auch nur entlang eines Netzwerkes bewegen kann und nicht über Hindernisse wie Gebäude oder Flüsse hinweg. Die Parameter m und $q (= 1)$ sind Konstanten.

$$\text{personalisierte räumliche Relevanz} = m * \text{räumliche Distanz} + 1 \quad (5.4)$$

Somit erweitern das Alter und das Fortbewegungsmittel des Nutzers die räumliche Relevanz insofern, als dass sie zur Bestimmung von m herangezogen werden. Dementsprechend wird für die personalisierte räumliche Relevanz neben dem Raum auch die Kontext-Dimension *Nutzer* (Alter und Fortbewegungsmittel) verwendet wodurch die Formel 5.4 zwei Variablen aufweist.

5.2 Wegzeit-Relevanz

Die Wegzeit-Relevanz wzR kann im Gegensatz zur korrigierten zeitlichen Relevanz kzR sowohl für zeitlich nicht fixierte Geoobjekte als auch für zeitlich fixierte Objekte bestimmt werden. Dabei wird wzR analog zur personalisierten räumlichen Relevanz prR bestimmt, indem die Geradengleichung des Modells (wzR in Abbildung 4.2) bestimmt wird. Die Abszisse steht für die Wegzeit zwischen Nutzer und Geoobjekt. Die Konstante q beschreibt den Relevanz-Wert für den Fall, dass die Wegzeit null beträgt und ist folglich gleich eins. Mit der Steigung m wird beschrieben, wie stark die Relevanz mit steigender zeitlicher Distanz abnimmt. Während R die Unbekannte, die Wegzeit die Variable und q die eine Konstante ist, muss nun lediglich noch m als zweite Variable bestimmt werden.

Während in dieser Arbeit für prR gilt, dass eine bestimmte Strecke jungen Nutzern nicht dieselbe Mühe bereitet wie älteren Nutzern, gilt bei der zeitlichen Distanz, dass eine bestimmte Zeitdauer bei allen Nutzern gleich empfunden wird. Demzufolge wird die Konstante m unabhängig vom Alter und Fortbewegungsmittel des Nutzers bestimmt, indem Δy gleich eins und Δx gleich 5 Minuten gesetzt werden (Formel 5.13). Das heisst, dass in dieser Arbeit Geoobjekte ab einer zeitlichen Entfernung von 5 Minuten, für den Nutzer in dieser Arbeit als nicht relevant erachtet werden, was empirisch noch zu belegen ist und in dieser Arbeit als Annahme gilt (Abbildung 5.2).

Die 5 Minuten entsprechen dabei auch der Zeit, die der Nutzer benötigt, um gemä-

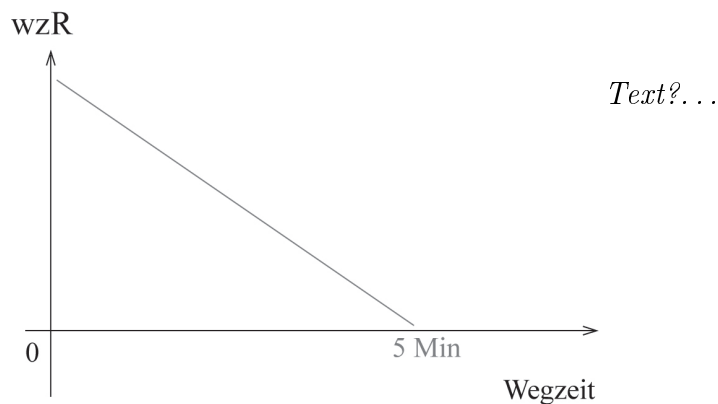


Abbildung 5.2: Konkrete Zeitangaben

ss seiner Geschwindigkeit (Tabelle 5.2) die maximalen Entfernungen (Tabelle 5.1) zu erreichen. Anders ausgedrückt lassen sich die 10 Minuten aus dem Quotienten aus der maximalen Entfernung und der Geschwindigkeit des entsprechenden Nutzers

ableiten (Formel 5.5).

$$Wegzeit = \frac{\text{räumliche Entfernung}}{\text{Geschwindigkeit}} \quad (5.5)$$

Wenn also bestimmt wird, dass für einen Fussgänger mit einer Geschwindigkeit von 100 Metern pro Minute nur Objekte bis zu einer maximalen Entfernung von 600 Metern relevant sind, so liesse sich dies auch (über seine Geschwindigkeit) durch die maximale zeitliche Entfernung von 6 Minuten ausdrücken. Dies bedeutet, dass bei die personalisierte räumliche Relevanz und die Wegzeitliche Relevanz identisch sind, denn ob ein Objekt nun 6 Minuten oder 600 Meter entfernt ist soll in dieser Arbeit keine Rolle spielen. Erst wenn beispielsweise über eine Strassenattributierung (Baustellen etc.) die zulässige Geschwindigkeit eines Strassenabschnittes berücksichtigt wird, würden sich kzR und wzR voneinander unterscheiden.

Während bei der personalisierten räumlichen Relevanz beide Parameter des Nutzers über die Variable m direkt in die Formel 5.4 einfließen, werden sie bei der Formalisierung der zeitlichen Relevanz indirekt über die Bestimmung der Wegzeit ($zdist$) miteinbezogen (Formel 5.6).

Somit lautet die Formel für die Wegzeit-Relevanz für zeitlich nicht fixierte Geoobjekte wie folgt:

$$wzR = -\frac{1}{10} * Wegzeit + 1 \quad (5.6)$$

Die Geschwindigkeit des Nutzers wird in Abhängigkeit des Alters und des Fortbewegungsmittels des Nutzers gestellt und ist der Tabelle 5.2 zu entnehmen. Dabei gelten die Geschwindigkeiten in dieser Arbeit lediglich als Annahmen und sind empirisch nicht belegt.

		Fortbewegung		
		zu Fuss	Fahrrad	Auto
Alter	jünger/gleich 50	100	320	600
	zwischen 50 und jünger/gleich 65	80	300	600
	älter als 65	60	280	600

Tabelle 5.2: Geschwindigkeit des Nutzers in [m/min]

5.3 Korrigierte zeitliche Relevanz

Bei den zeitlich fixierten Geobjekten verhält es sich ein wenig komplizierter als bei den zeitlich nicht fixierten Objekten, denn bei ihnen muss zusätzlich zur Wegzeit auch die zeitliche Entfernung zum Eintreten eines Ereignisses (Δt) berücksichtigt werden (s. Kapitel 4.3). Das heisst, dass sich dieser Relevanztyp nur auf zeitlich fixierte Geobjekte anwenden lässt.

Bei der Eintrittszeit wird zwischen einfach und zweifach terminierten Geobjekten unterschieden.

5.3.1 Einfach terminierte Geobjekte

Wie sich das Modell der Relevanz eines Objektes mit Zeitpunkt in Abhängigkeit von $\Delta t_{\text{effektiv}}$ verhält, ist im vorgängigen Kapitel 3 beschrieben. Die konkreten Zeitangaben, welche in dieser Arbeit angenommen werden und empirisch nicht belegt sind, sind dem Modell 5.3 zu entnehmen.

Die Formalisierung dieses Modells geht analog zu den vorangegangenen Formali-

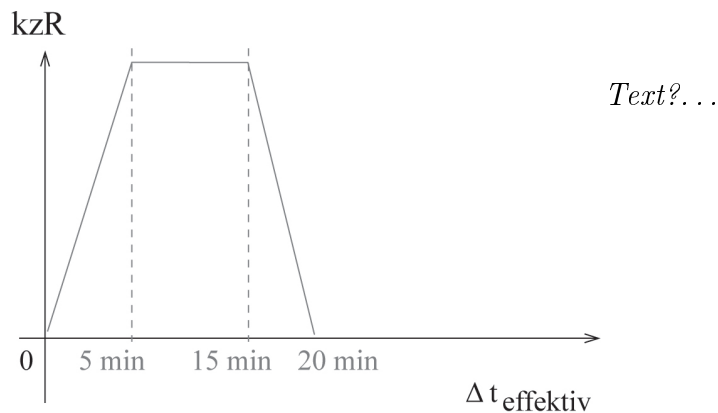


Abbildung 5.3: Konkrete Zeitangaben

sierungen, nämlich über die Geradengleichung, von statten, nur dass es hier fünf

Gültigkeitsbereiche gibt.

$$kzR = \begin{cases} \text{für } eEZ \leq 0 : & 0 \\ \text{für } 0 < eEZ \leq 5 : & \frac{1}{5} * eEZ \\ \text{für } 5 < eEZ \leq 15 : & 1 \\ \text{für } 15 < eEZ \leq 20 : & -\frac{1}{5} * eEZ + 4 \\ \text{für } eEZ > 20 : & 0 \end{cases} \quad (5.7)$$

5.3.2 Zweifach terminierte Geoobjekte

Während einfach terminierte Geoobjekte nur über eine Eintrittszeit verfügen (Startzeit), weisen zweifach terminierte Objekte deren zwei auf (Öffnungszeit und Endzeit). Aus den illustrierten Modellen im Kapitel 3.2.2 ist ersichtlich, wie sich die zeitliche Relevanz in Abhängigkeit von $\Delta t_{effektiv}$ bis zur Öffnungszeit einerseits (Abbildung 3.5) und von $\Delta t_{effektiv}$ bis zur Endzeit andererseits (Abbildung 3.6) verhält. Diese zeitliche Relevanz eines Objektes, welches von 08:00 Uhr morgens bis 18:00 Uhr abends geöffnet ist, lässt sich aber einfacher auf einem Zeitstrahl ablesen (Abbildung 5.4). In dieser Arbeit gilt die Annahme (empirisch nicht belegt), dass

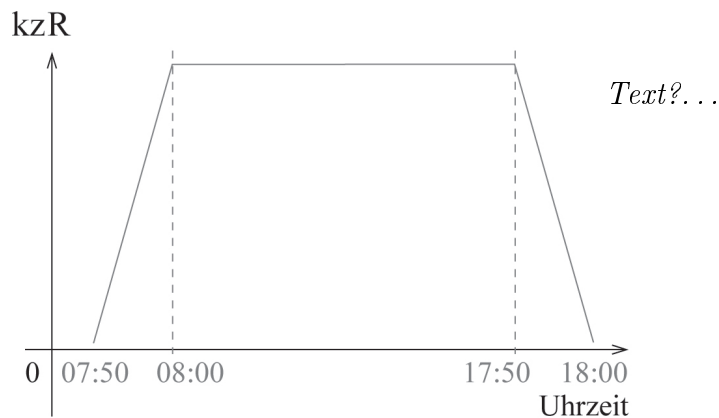


Abbildung 5.4: Relevanz auf Zeitstrahl

die maximale zeitliche Entfernung bis zur Öffnungszeit 10 Minuten betragen und kzR 10 Minuten vor Ende sinken soll. Die Idee dahinter ist, dass erstens eine kurze Wartezeit meistens in Kauf genommen und zweitens für den Besuch eines Objektes eine bestimmte Zeit benötigt wird (5 Minuten vor Endzeit kann in den meisten Restaurants keine Bestellungen mehr aufgegeben werden).

Die mathematische Formalisierung eines solchen Zeitstrahls ist insofern schwierig,

als dass die Uhrzeit keine mathematische Grösse ist. Aus diesem Grund wird ein Modell gesucht, welches die zeitliche Relevanz einzig von $\Delta t_{effektiv}$ abhängig macht.

Die Abbildung 5.5 zeigt die korrigierte zeitliche Relevanz für ein Geoobjekt, welches

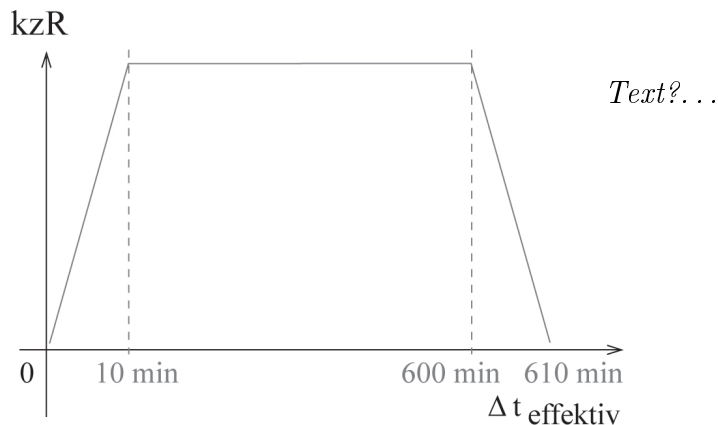


Abbildung 5.5: Relevanz in Abhängigkeit von $\Delta t_{effektiv}$ (bis Endzeit)

10 Stunden geöffnet hat. Beträgt $\Delta t_{effektiv}$ bis zur Endzeit 10 Stunden und 5 Minuten ($eEZ = 605 \text{ min}$), so bedeutet dies auch, dass das Objekt in 5 Minuten öffnen wird. Durch die Verwendung der Öffnungsdauer (ÖD) eines Objektes lässt sich somit die zeitliche Relevanz so modellieren, dass $\Delta t_{effektiv}$ bis zur Öffnungszeit und $\Delta t_{effektiv}$ bis zur Endzeit gleichzeitig berücksichtigt werden. Die Formalisierung dieses Modells ist in Formel 5.8 festgehalten.

$$kzR = \begin{cases} \text{für } eEZ < 0 : & 0 \\ \text{für } 0 < eEZ < 10 : & \frac{1}{10} * eEZ \\ \text{für } 10 < eEZ < \text{ÖD} : & 1 \\ \text{für } \text{ÖD} < eEZ < \text{ÖD} + 10 : & -\frac{1}{10} * eEZ + \frac{\text{ÖD}+10}{10} \\ \text{für } eEZ > \text{ÖD} + 10 : & 0 \end{cases} \quad (5.8)$$

5.4 Thematische Relevanz

Wie in Kapitel 3.3.2 erwähnt, nimmt die thematische Relevanz mit zunehmender semantischer Distanz zwischen der Anfrage des Nutzers und der Kategorie des Geoobjektes ab. Während bei den metrischen Datentypen (Raum und Zeit) die entsprechende Distanz als euklidische Entfernung (Raum) respektive als Differenz (Zeit)

bestimmt werden konnten, wird dies beim nichtmetrischen Datentyp der Thematik schwieriger.

[Jones et al.(2001)] beschreiben die semantische Distanz zwischen zwei Begriffen als kürzeste Distanz, um in einer Hierarchie (Baumdiagramm) vom einen Begriff zum anderen zu gelangen. Eine solche Hierarchie könnte in dieser Arbeit eingesetzt werden, um beispielsweise die semantische Distanz zwischen den Kategorien *Hotel* und *Restaurant* zu bestimmen. [Rodríguez und Egenhofer(2004)] verwenden zur Bestimmung der Ähnlichkeit zweier Geoobjekte nicht nur diese hierarchische Distanz, sondern auch die Attribute oder die Funktionen der Objekte. Dieser Ansatz wird hier jedoch nicht umgesetzt, weil in dieser Arbeit lediglich die Ähnlichkeit zwischen den Kategorien der Objekte untersucht werden soll (s. Kapitel 3.3.1).

In dieser Arbeit wird die Ähnlichkeit zwischen der Kategorie, für welche sich der Nutzer interessiert (*Anfrage*) und der Kategorie eines Geoobjektes, im voraus aufgrund der geschätzten Ähnlichkeit bestimmt (Tabelle 5.3). Es wird somit keine Hierarchie festgelegt, sondern es wird für jedes mögliche Paar (zweier Kategorien) die Ähnlichkeit geschätzt. Diese Schätzung gilt bloss im Rahmen dieser Arbeit und ist empirisch nicht belegt.

Eine Ähnlichkeit von null bedeutet, dass die Anfrage des Nutzers mit der Objektart

		Obejektart des Geoobjektes								
		Hotel	Pension	Jugendherberge	Restaurant	Imbiss	Bar	Kino	Theater	Disco
Anfrage	Hotel	1	0.9	0.8	0.2	0.1	0.2	0	0	0
	Pension	0.9	1	0.9	0.2	0.1	0	0	0	0
	Juendherberge	0.8	0.9	1	0	0	0	0	0	0
	Restaurant	0.2	0.2	0	1	0.8	0.5	0	0	0
	Imbiss	0.1	0.1	0	0.8	1	0.6	0	0	0
	Bar	0.2	0	0	0.5	0.6	1	0.4	0.4	0.6
	Kino	0	0	0	0	0	0.4	1	0.9	0.8
	Theater	0	0	0	0	0	0.4	0.9	1	0.4
	Disco	0	0	0	0	0	0.6	0.8	0.4	1

Tabelle 5.3: Ähnlichkeiten zwischen verschiedenen Objektarten

des Geoobjektes überhaupt nichts gemeinsam hat, während der Wert eins für maximale Übereinstimmung steht. Die semantische Distanz und die Ähnlichkeit verhalten sich dabei komplementär zu eins, je grösser also die Ähnlichkeit, desto geringer ist

die semantische Distanz (Formel 5.9).

$$\textit{semantische Distanz} = 1 - \textit{Ähnlichkeit} \quad (5.9)$$

Bei einer semantischen Distanz von null soll die thematische Relevanz $tR = 1$ sein, während bei einer semantischen Distanz von eins die thematische Relevanz $tR = 0$ sein soll. Die Ähnlichkeitswerte aus der Tabelle 5.3 werden in dieser Arbeit somit auch als thematische Relevanz-Werte interpretiert, weil sich auch die thematische Relevanz und die semantische Distanz komplementär zu eins verhalten (Formel 5.10).

$$\textit{thematische Relevanz} = \textit{Ähnlichkeit} \textit{ (gemäss Tabelle 5.3)} \quad (5.10)$$

5.5 Kombination

Bis jetzt sind alle vier Relevanztypen formalisiert worden, sodass ein Geoobjekt auch vier einzelne Relevanz-Werte aufweist. Nun folgt deren Kombination, damit ein Wert resultiert, welcher die gesamthafte Relevanz ($totR$) ausdrückt. Dabei gilt es zu beachten, dass die Wegzeit-Relevanz wzR in dieser Arbeit stets identisch ist mit prR (vgl. Kapitel 5.2) und wzR daher nicht in die Formalisierung einbezogen wird, um die räumliche Distanz respektive die dafür benötigte Zeit nicht doppelt zu verrechnen. Es gibt verschiedene Möglichkeiten, aus vier einzelnen Werten einen einzelnen Wert abzuleiten. An dieser Stelle werden drei Varianten, mitsamt ihren Vor- und Nachteilen vorgestellt.

5.5.1 Durchschnitt

Die wohl einfachste Variante, die Werte der vier Relevanztypen zu einem gesamtheitlichen Relevanz-Wert $totR$ zu vereinen, ist es, den Durchschnitt zu bilden. Auf diese Weise werden alle Relevanztypen in gleichem Masse berücksichtigt. Allerdings kann auf diese Weise keinem Relevanztypen ein besonderes Gewicht verliehen werden, wie das beim gewichteten Durchschnitt der Fall ist. Ein weiterer Nachteil ist, dass beispielsweise ein Restaurant, welches zwar räumlich nah liegt, jedoch bereits geschlossen hat und daher nicht mehr besucht werden kann, dennoch für $totR$ einen Wert \neq null erhält.

$$totR = \frac{prR + kzR + tR}{3} \quad (5.11)$$

5.5.2 Gewichteter Durchschnitt

Ähnlich der zuvor vorgestellten Variante berücksichtigt auch diese Variante sämtliche Relevanztypen, jedoch unterschiedlich stark. Die Berechnung wird durch die Bestimmung dieser Gewichte ($g1 - g3$) zwar ein wenig erschwert, jedoch kann dadurch beispielsweise die Tatsache einbezogen werden, dass die Interessen des Nutzers den höchsten Stellenwert erhalten sollen und somit tR stärker gewichtet werden soll als die anderen Relevanztypen. Wie auch schon beim einfachen Durchschnitt liegt hier der Nachteil vor, dass Geoobjekte, welche zu weit weg oder bereits geschlossen haben, trotzdem für $totR$ einen Wert \neq null erhalten.

$$totR = \frac{g1 * prR + g2 * kzR + g3 * tR}{3} \quad (5.12)$$

5.5.3 Minimum

Diese Variante basiert auf der Annahme, dass der kleinste Relevanz-Wert über die Relevanz eines Geoobjektes entscheidet, denn auch wenn ein Objekt räumlich nahe liegen würde, wäre es trotzdem nicht relevant, falls sich der Nutzer nicht für diese Kategorie von Objekten interessieren würde. In dieser Arbeit gilt diese dritte Variante als geeignetste Form der Kombination und wird daher anschliessend implementiert. Die Relevanz eines Geoobjektes wird in dieser Arbeit daher wie folgt definiert:

$$totR = MIN(prR, kzR, tR) \quad (5.13)$$

5.6 Fazit und Ausblick

In diesem Kapitel wurden die einzelnen Relevanztypen prR , wzR , kzR und tR , welche in den vorangehenden Kapiteln 3 und 4 beschrieben und modelliert worden sind, zuerst einzeln formalisiert und anschliessend miteinander kombiniert. Bei der Kombination der einzelnen Relevanztypen zur gesamtheitlichen Relevanz $totR$ eines Geoobjektes wird wzR nicht integriert, weil prR und wzR in dieser Arbeit identisch sind und nicht doppelt verrechnet werden sollten. Durch diese Formalisierung kann nun die Relevanz eines Geoobjektes unter Berücksichtigung der drei Kontext-Dimensionen Raum, Zeit und Nutzer quantifiziert werden.

Diese Formel kann nun in OpenJUMP implementiert werden. Die Präsentation der Resultate dieser Implementierung folgt im nächsten Kapitel.

Kapitel 6

Resultate

In diesem Kapitel wird aufgezeigt, welche Resultate das in OpenJUMP erweiterte Plugin (s. Kapitel 2.2.2) hervorbringt, indem es den implementierten Formalismus auf die Geoobjekte anwendet und so jedem dieser Objekte einen Relevanz-Wert zuteilt. Über das Interface (Abb. 6.1) werden dem Plugin anschliessend drei unterschiedliche Ausgangslagen vorgegeben. Die Resultate werden sowohl kartographisch als auch tabellarisch gezeigt.

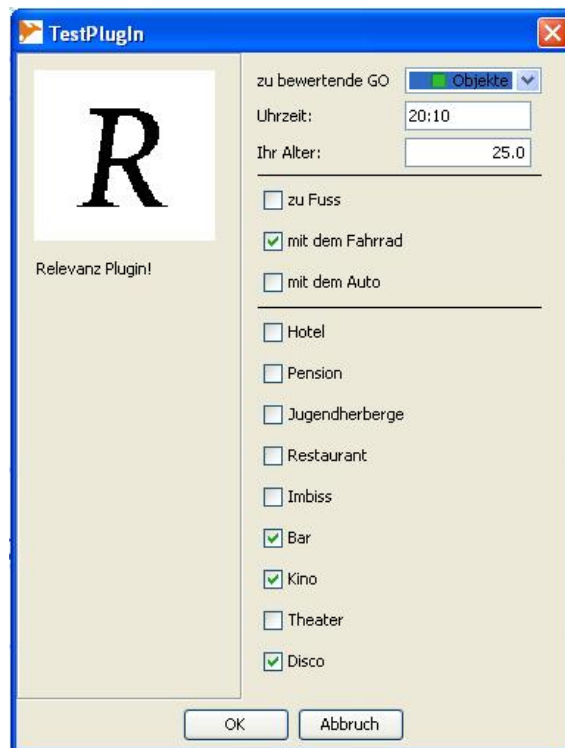


Abbildung 6.1: Interface

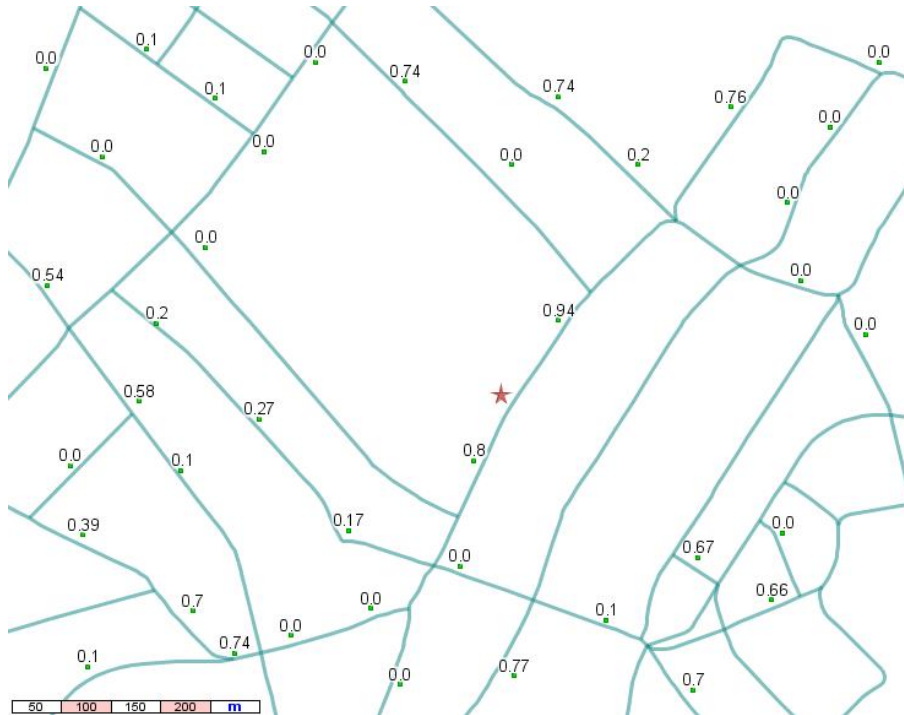


Abbildung 6.2: Screenshot vom resultierenden layer

Diese beiden Abbildungen (6.2 und 6.3) zeigen das Resultat für einen 25-jährigen Nutzer, welcher sich für Ausgang (*Bar, Kino, Disco*) interessiert. Er ist mit dem *Fahrrad* unterwegs und es ist 20:10 Uhr.

FID	Objektart	Startzeit	Endzeit	Räuml. Dist	Wegzeit	prR	kzR	wzR	tR	totR
41	Bar	19:30	24:00	94.0	0.3	0.94	1.0	0.94	1.0	0.94
28	Bar	19:00	24:00	365.0	1.1	0.77	1.0	0.77	1.0	0.77
40	Theater	20:15	0	379.0	1.2	0.76	0.76	0.76	0.9	0.76
3	Bar	18:00	24:00	424.0	1.3	0.74	1.0	0.74	1.0	0.74
20	Disco	20:00	24:00	477.0	1.5	0.7	1.0	0.7	1.0	0.7
36	Bar	18:30	23:30	479.0	1.5	0.7	1.0	0.7	1.0	0.7
12	Bar	18:30	23:30	728.0	2.3	0.54	1.0	0.54	1.0	0.54
8	Restaurant	11:30	23:00	783.0	2.4	0.51	1.0	0.51	0.5	0.5
17	Restaurant	11:00	22:00	521.0	1.6	0.67	1.0	0.67	0.5	0.5
19	Restaurant	10:30	23:00	551.0	1.7	0.66	1.0	0.66	0.5	0.5
32	Restaurant	15:00	23:30	420.0	1.3	0.74	1.0	0.74	0.5	0.5
35	Restaurant	11:30	23:00	415.0	1.3	0.74	1.0	0.74	0.5	0.5
42	Restaurant	11:00	22:45	69.0	0.2	0.96	1.0	0.96	0.5	0.5
10	Kino	20:30	0	617.0	1.9	0.61	0.39	0.61	1.0	0.39
14	Theater	20:30	0	427.0	1.3	0.73	0.27	0.73	0.9	0.27
4	Hotel	08:00	22:00	319.0	1.0	0.8	1.0	0.8	0.2	0.2
13	Hotel	09:00	22:30	567.0	1.8	0.65	1.0	0.65	0.2	0.2
15	Kino	20:30	0	276.0	0.9	0.83	0.17	0.83	1.0	0.17
1	Imbiss	11:00	19:00	642.0	2.0	0.6	0.0	0.6	0.6	0.0
2	Pension	09:00	21:00	703.0	2.2	0.56	1.0	0.56	0.0	0.0

Abbildung 6.3: Tabelle der Geoobjekte

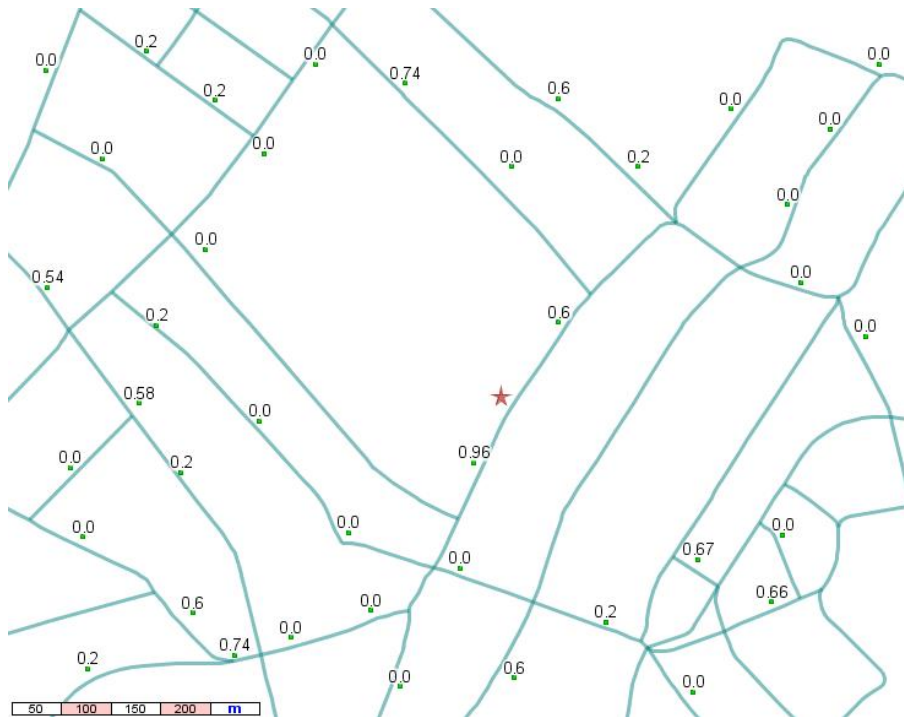


Abbildung 6.4: Tabelle der Geobjekte

Diese beiden Abbildungen (6.4 und 6.5) zeigen das Resultat für einen 25-jährigen Nutzer, welcher sich für eine Verpflegung (*Restaurant, Imbiss*) interessiert. Er ist mit dem *Fahrrad* unterwegs und es ist 20:10 Uhr. Vergleicht man diese beiden Abbildungen mit den beiden vorangehenden Abbildungen (6.2 und 6.3) lässt sich gut erkennen, welchen Einfluss unterschiedliche Interessen der Nutzer (*ceteris paribus*) auf die Relevanz-Werte der Geobjekte haben.

FID	Objektart	Startzeit	Endzeit	Räuml. Dist	Wegzeit	prR	kzR	wzR	tR	totR
42	Restaurant	11:00	22:45	69.0	0.2	0.96	1.0	0.96	1.0	0.96
32	Restaurant	15:00	23:30	420.0	1.3	0.74	1.0	0.74	1.0	0.74
35	Restaurant	11:30	23:00	415.0	1.3	0.74	1.0	0.74	1.0	0.74
17	Restaurant	11:00	22:00	521.0	1.6	0.67	1.0	0.67	1.0	0.67
19	Restaurant	10:30	23:00	551.0	1.7	0.66	1.0	0.66	1.0	0.66
3	Bar	18:00	24:00	424.0	1.3	0.74	1.0	0.74	0.6	0.6
28	Bar	19:00	24:00	365.0	1.1	0.77	1.0	0.77	0.6	0.6
36	Bar	18:30	23:30	479.0	1.5	0.7	1.0	0.7	0.6	0.6
41	Bar	19:30	24:00	94.0	0.3	0.94	1.0	0.94	0.6	0.6
8	Restaurant	11:30	23:00	674.0	2.1	0.58	1.0	0.58	1.0	0.58
12	Bar	18:30	23:30	728.0	2.3	0.54	1.0	0.54	0.6	0.54
2	Pension	09:00	21:00	703.0	2.2	0.56	1.0	0.56	0.2	0.2
4	Hotel	08:00	22:00	319.0	1.0	0.8	1.0	0.8	0.2	0.2
9	Pension	09:30	21:00	592.0	1.9	0.63	1.0	0.63	0.2	0.2
13	Hotel	09:00	22:30	567.0	1.8	0.65	1.0	0.65	0.2	0.2
21	Pension	09:00	21:30	365.0	1.1	0.77	1.0	0.77	0.2	0.2
30	Pension	08:30	22:15	787.0	2.5	0.51	1.0	0.51	0.2	0.2
37	Pension	09:00	22:00	564.0	1.8	0.65	1.0	0.65	0.2	0.2
1	Imbiss	11:00	19:00	642.0	2.0	0.6	0.0	0.6	1.0	0.0
5	Theater	20:00	0	286.0	0.9	0.82	0.0	0.82	0.0	0.0

Abbildung 6.5: Tabelle der Geobjekte

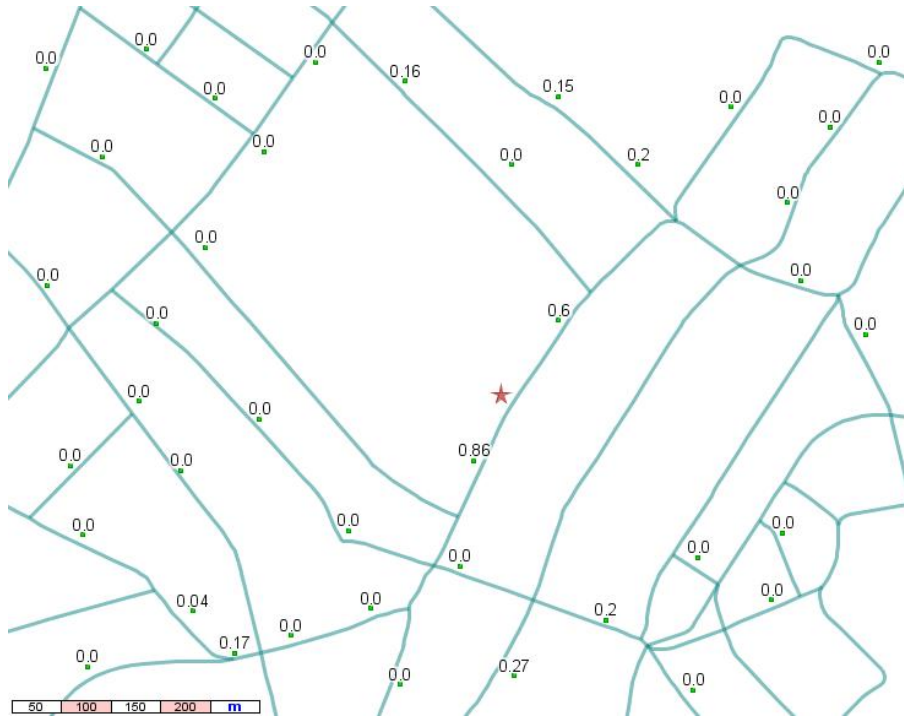


Abbildung 6.6: Tabelle der Geobjekte

Diese beiden Abbildungen (6.6 und 6.7) zeigen das Resultat für einen 25-jährigen Nutzer, welcher sich für eine Verpflegung (*Restaurant, Imbiss*) interessiert. Er ist zu *Fuss* unterwegs und es ist 20:10 Uhr. Hier lässt sich gut erkennen, dass die Relevanzwerte der Geobjekte im Gegensatz zu den beiden obigen Abbildungen (6.4 und 6.5) tiefer sind, weil der Nutzer zu *Fuss* unterwegs ist.

FID	Objektart	Startzeit	Endzeit	Räuml. Dist	Wegzeit	prR	kzR	wzR	tR	totR ▼
42	Restaurant	11:00	22:45	70.0	0.7	0.86	1.0	0.86	1.0	0.86
41	Bar	19:30	24:00	94.0	0.9	0.81	1.0	0.81	0.6	0.6
28	Bar	19:00	24:00	365.0	3.7	0.27	1.0	0.27	0.6	0.27
4	Hotel	08:00	22:00	319.0	3.2	0.36	1.0	0.36	0.2	0.2
21	Pension	09:00	21:30	365.0	3.7	0.27	1.0	0.27	0.2	0.2
35	Restaurant	11:30	23:00	415.0	4.2	0.17	1.0	0.17	1.0	0.17
32	Restaurant	15:00	23:30	420.0	4.2	0.16	1.0	0.16	1.0	0.16
3	Bar	18:00	24:00	424.0	4.2	0.15	1.0	0.15	0.6	0.15
36	Bar	18:30	23:30	479.0	4.8	0.04	1.0	0.04	0.6	0.04
1	Imbiss	11:00	19:00	642.0	6.4	0.0	0.0	0.0	1.0	0.0
2	Pension	09:00	21:00	703.0	7.0	0.0	1.0	0.0	0.2	0.0

Abbildung 6.7: Tabelle der Geobjekte

Kapitel 7

Diskussion

In diesem Kapitel werden die soeben vorgestellten Resultate diskutiert und kritisch beurteilt. Dabei sollen die Stärken und die Grenzen des vorgestellten Relevanz-Formalismus herausgearbeitet und Verbesserungsvorschläge gemacht werden.

Ziel der vorliegenden Arbeit ist es, die Relevanz von Geoobjekten zu bestimmen, um die Informationsübermittlung von der Karte hin zum Nutzer zu optimieren. Demzufolge ist die Relevanz subjektiv und kann je nach Nutzer einen anderen Wert annehmen. Die Kontext-Dimension des Nutzers ist somit zentral. In dieser Arbeit wird der Nutzer alleine auf seine Interessen bezüglich der Kategorie der Geoobjekte, sein Alter und sein Fortbewegungsmittel reduziert. Seine Individualität kann daher auch nur in sehr beschränktem Masse berücksichtigt werden. Es ist klar, dass weitere Kontext-Informationen des Nutzers (wie beispielsweise seine aktuelle Gefühlslage) die Relevanz eines Geoobjektes, wie sie im vorangehenden Kapitel präsentiert worden sind, weiter beeinflussen können. Die Annahmen über solch fehlende Kontext-Informationen, wie sie in dieser Arbeit der Modellierung und Formalisierung zugrunde liegen, können also nur bedingt befriedigen, weil diese Annahmen für alle Nutzer generelle Bestimmungen vorsehen.

Eine weitere Einschränkung der Relevanz-Bestimmung in dieser Arbeit betrifft die Auswahl der Kontext-Informationen, welche berücksichtigt worden sind. Nicht nur die räumliche oder zeitliche Entfernung zwischen Geoobjekt und Nutzer, die (tages-)zeitliche Fixierung oder die semantische Distanz zwischen der Kategorie des Geoobjekt und der Interessen des Nutzers diesbezüglich können die Relevanz eines Geoobjektes beeinflussen. So gibt es zahlreiche andere Kontext-Informationen, welche ebenfalls eine Auswirkung haben, wie beispielsweise die Sichtbarkeit eines Geoobjektes oder die aktuelle Wetterlage. Je mehr Informationen jedoch berücksichtigt

werden sollen, desto mehr Kontext-Informationen müssen gesammelt werden und desto komplexer wird die Relevanz-Analyse.

Trotz diesen Grenzen, welche dem vorgeschlagenen Relevanz-Formalismus gesetzt sind, ist das eigentliche Ziel dieser Arbeit erreicht worden. Nämlich eine nicht nur binäre Unterscheidung zwischen relevanten respektive nicht relevanten Geoobjekten alleine in Abhängigkeit des Raumes (klassische LBS), sondern erstens eine feinere Abstufung des Relevanz-Wertes und zweitens in Abhängigkeit mehrerer Kontext-Dimensionen, zu ermöglichen.

Kapitel 8

Zusammenfassung und Ausblick

Die Relevanz-Analyse von Informationen im Allgemeinen und von geographischen Informationen im Speziellen werden seit längerer Zeit im Bereich (Geographic) Information Retrieval untersucht, um die Informationssuche für den Nutzer zu erleichtern. Dabei werden Dokumente bezüglich ihrer Übereinstimmung mit dem Informationsbedürfnis des Nutzers bewertet.

Dieser Bedarf der Relevanz-Analyse besteht auch im Bereich der mobilen Kartographie, jedoch ist dieser Ansatz noch weitestgehend unerforscht. Hier sollen jedoch keine Dokumente, sondern Geoobjekten bewertet werden. Durch eine solche Bewertung von Geoobjekten könnte beispielsweise ihre Visualisierung gesteuert werden oder es könnte vorgängig bestimmt werden, welche Geoobjekte relevant sind um nur die relevantesten aus einer Datenbank zu extrahieren. In beiden Fällen könnte die Informationsvermittlung einer mobilen Karte zum Nutzer optimiert werden.

Die Relevanz eines Geoobjekts hängt dabei von vielen Kontext-Dimensionen ab, je nach Anwendung. In der vorliegenden Arbeit allerdings wurde die Relevanz eines Objektes in Abhängigkeit der Kontext-Dimensionen des Raumes, der Zeit und des Nutzers bestimmt. Es ist klar, dass es weitere Kontext-Informationen gibt, welche die Relevanz ebenfalls beeinflusse und dass insbesondere die subjektive Kontext-Dimension des Nutzers, welche von zentraler Bedeutung ist, nicht alleine durch die seine Interessen bezüglich der Kategorie eines Geoobjektes, sein Alter und Fortbewegungsmittel abgedeckt werden kann. Aus diesem Grund basieren die Modellierung und Formalisierung der Relevanz in dieser Arbeit auf Annahmen, welche diese fehlenden Kontext-Informationen des Nutzers kompensieren sollen.

Der in dieser Arbeit vorgeschlagene und implementierte Formalismus ist somit als Annäherung an eine Relevanz-Bestimmung von Geoobjekten, welche von mehreren Kontext-Dimensionen abhängt und (nicht-binär) quantifiziert werden kann, zu ver-

stehen.

In vielen Bereichen besteht allerdings noch Forschungsbedarf, um Relevanz-Analysen effektiver umsetzen zu können. So muss abgeklärt werden, welche Informationen über den Nutzer die Relevanz eines Geoobjektes in welchem Mass beeinflussen und auf welche Weise solche Informationen aus seinem Verhalten abgeleitet werden können, ohne ihn dabei ständig anzufragen und dadurch zu belästigen. Weiterer Forschungsbedarf besteht auch in der Gewinnung von low-level und high-level Kontext-Informationen. Während für die low-level Informationen weiter untersucht werden muss, welche relevanten Informationen mit welchen Sensoren gemessen werden können, muss für die high-level Informationen weiter abgeklärt werden, wie sie aus der Kombination von diesen low-level Informationen gewonnen werden können. Bei all diesen Kontext-Informationen muss ausserdem die Sicherheit und der Datenschutz gewährleistet werden (gutevaria 5.pdf -> Quelle!).

Nicht nur die Gewinnung von Kontext-Informationen, sondern auch die Beurteilung ihrer Bedeutung für die Relevanz eines Geoobjektes bedarf weiterer Untersuchungen.

Selbstständigkeitserklärung

Hiermit bestätige ich, dass ich die vorliegende Arbeit mit dem Titel Relevanz-Analyse, selbstständig verfasst habe und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt wurden, sowie Zitate kenntlich gemacht habe.

Datum: 27. Oktober 2008

Unterschrift:

Literaturverzeichnis

- [Abowd et al.(1999)] ABOWD, G. D., DEY, A. K., BROWN, P. J., DAVIES, N., SMITH, M. UND STEGGLES, P. (1999): Towards a Better Understanding of Context and Context-Awareness. In: HUC '99: Proceedings of the 1st international symposium on Handheld and Ubiquitous Computing. Springer-Verlag, S. 304–307.
- [Alexander Zipf(2003)] ZIPF, A. (2003): Zur Bestimmung von Funktionen für die personen- und kontextsensitive Bewertung der Bedeutung von Geoobjekten für Fokuskarten. In: Symposium für Angewandte Geographische Informationstechnologie.
- [Amendola et al.(2004)] AMENDOLA, I., CENA, F., CONSOLE, L., CREVOLA, A., GOY, A., MODEO, S., PERRERO, M., TORRE, I. UND TOSO, A. (2004): UbiquiTO: a Multi-Device Adaptive Guide. In: Proceedings of Mobile Human-Computer Interaction – MobileHCI 2004: 6th International Symposium. Springer, S. 13–16.
- [Chen und Kotz(2000)] CHEN, G. UND KOTZ, D. (2000): A Survey of Context-Aware Mobile Computing Research. Technischer Bericht TR2000-381, Dept. of Computer Science, Dartmouth College.
- [Cheverst et al.(2000)] CHEVERST, K., DAVIES, N., MITCHELL, K., FRIDAY, A. UND EFSTRATIOU, C. (2000): Developing a Context-aware Electronic Tourist Guide: Some Issues and Experiences. 17–24. ACM Press.
- [Edwardes(2007)] EDWARDES, A. J. (2007): Re-placing Location: Geographic Perspectives in Location Based Services. Dissertation, Universität Zürich.
- [Hermann und Heidmann(2002)] HERMANN, F. UND HEIDMANN, F. (2002): User Requirement Analysis and Interface Conception for a Mobile, Location-Based

- Fair Guide. In: In Paterno F. (Ed.) Human Computer Interaction with Mobile Devices. Springer, S. 388–392.
- [Jones et al.(2001)] JONES, C. B., ALANI, H. UND TUDHOPE, D. (2001): Geographical Information Retrieval with Ontologies of Place. In: COSIT 2001: Proceedings of the International Conference on Spatial Information Theory. Springer, London, S. 322 – 335.
- [Keim und Kriegel(1994)] KEIM, D. A. UND KRIEGEL, H.-P. (1994): Visdb: Database exploration using multidimensional visualization. In: IEEE Computer Graphics and Applications, 14, S. 40 – 49.
- [Long et al.(1996)] LONG, S., KOOPER, R., ABOWD, G. D. UND ATKESON, C. G. (1996): Rapid Prototyping of Mobile Context-Aware Applications: The Cyberguide Case Study. ACM Press, S. 97–107.
- [Machill und Beiler(2007)] MACHILL, M. UND BEILER, M. (2007): Die Macht der Suchmaschinen. Herbert von Halem Verlag, Köln.
- [Malaka und Zipf(2000)] MALAKA, R. UND ZIPF, A. (2000): Deep Map - challenging IT research in the framework of a tourist information system. In: Information and Communication Technologies in Tourism 2000. Springer, S. 15–27.
- [Reichenbacher(2004)] REICHENBACHER, T. (2004): Mobile Cartography – Adaptive Visualisation of Geographic Information on Mobile Devices. Dissertation, Technische Universität München.
- [Reichenbacher et al.(2002)] REICHENBACHER, T., ANGSÜSSER, S. UND MENG, L. (2002): Mobile Kartographie – eine offene Diskussion. In: Kartographische Nachrichten 52. Jg., Heft 4, 4, S. 164–166.
- [Rodríguez und Egenhofer(2004)] RODRÍGUEZ, M. A. UND EGENHOFER, M. J. (2004): Comparing Geospatial Entity Classes: An Asymmetric and Context-Dependent Similarity Measure. In: International Journal of Geographical Information Science, 18, S. 229 – 256.
- [Saracevic(1996)] SARACEVIC, T. (1996): Relevance reconsidered. In: Information science: Integration in perspectives.Proceedings of the Second Conference on Conceptions of Library and Information Science.

- [Saracevic(2007)] SARACEVIC, T. (2007): Relevance: A review of the literature and a framework for thinking on the notion in information science. Part II: nature and manifestations of relevance. In: Journal of the American Society for Information Science and Technology, 58(13), S. 1915 – 1933.
- [Schilit et al.(1994)] SCHILIT, B. N., ADAMS, N. UND WANT, R. (1994): Context-aware computing applications. In: In Proceedings of the Workshop on Mobile Computing Systems and Applications. IEEE Computer Society, S. 85–90.
- [Schmidt et al.(1999)] SCHMIDT, A., AIDOO, K. A., TAKALUOMA, A., TUOMELA, U., LAERHOVEN, K. V. UND VELDE, W. V. D. (1999): Advanced interaction in context. In: In Proceedings of First International Symposium on Handheld and Ubiquitous Computing. Springer Verlag, S. 89–101.
- [Schmidt et al.(1998)] SCHMIDT, A., BEIGL, M. UND W. GELLERSEN, H. (1998): There is more to context than location: Environment sensing technologies for adaptive mobile user interfaces. In: In Workshop on Interactive Applications of Mobile Computing IMC'98.
- [Schmidt-Belz et al.(2002)] SCHMIDT-BELZ, B., NICK, A., POSLAD, S. UND ZIPF, A. (2002): Personalized and location-based mobile tourism services. In: Proc. of Mobile-HCI.
- [Setten et al.(2004)] SETTEN, M. V., POKRAEV, S. UND KOOLWAAIJ, J. (2004): Context-aware recommendations in the mobile tourist application COMPASS. In: In Nejd, W. & De Bra, P. (Eds.). AH 2004, LNCS 3137. Springer, S. 235–244.
- [Sperber und Wilson(1986)] SPERBER, D. UND WILSON, D. (1986): Relevance: communication and cognition. Harvard University Press.
- [Want et al.(1995)] WANT, R., SCHILIT, B. N., ADAMS, N. I., GOLD, R., PETERSEN, K., GOLDBERG, D., ELLIS, J. R. UND WEISER, M. (1995): An overview of the ParcTab ubiquitous computing experiment. In: IEEE Personal Communications, 2, S. 28–43.
- [Worboys und Duckham(2004)] WORBOYS, M. UND DUCKHAM, M. (2004): GIS: A Computing Perspective, 2nd Edition. CRC Press.
- [Zadeh(1965)] ZADEH, L. A. (1965): Fuzzy Sets. In: Information and Control, 8(3), S. 338 – 353.

[Zipf und Richter(2002)] ZIPF, A. UND RICHTER, K.-F. (2002): Using Focus Maps to Ease Map Reading – Developing Smart Applications for Mobile Devices. In: Künstliche Intelligenz, 16(4), S. 35 – 37.