

Universität Stuttgart
Fakultät Informatik

Adressierung und Semantik von geographischen Nachrichten

Authors:

Peter Coschurba

Institute of Parallel and Distributed
High-Performance Systems (IPVR)
Department of Computer Science
University of Stuttgart
Breitwiesenstr. 20 - 22
D-70565 Stuttgart
Germany

Adressierung und Semantik von geographischen Nachrichten

Peter Coschurba

Technical Report 2001/04

Mai 2001

1	Einleitung	1
2	Verwandte Arbeiten	3
2.1	Protokolle aus Dataman-Projekt	3
2.1.1	Adressierung	3
2.1.2	Geographisches Routing	4
2.1.3	Geographische Multicast-Gruppen	5
2.1.4	Domain Name Server	6
2.2	Spezialisierte Protokolle	6
2.2.1	GeoCast in Ad-Hoc Netzen	6
3	Model	7
3.1	Netzwerk	7
3.2	Sender	8
3.3	Klient	8
4	Semantik	9
4.1	Einfaches Model	9
4.2	Beispiel	10
4.3	Gültigkeitsdauer einer Nachricht	11
4.3.1	Definition der Lifetime	11
4.3.2	Mindestaufenthaltsdauer eines Empfängers	12
4.4	Wer ist Empfänger	12
4.4.1	Zeitpunkt	12
4.4.2	Genauigkeit	14
4.5	QoS	15
4.5.1	Prioritäten	15
4.5.2	Verlust	16
4.5.3	Reihenfolge	17
4.5.4	Nachrichtenlaufzeit	17
5	Adressierung	19
5.1	Einführung	19
5.2	Weltmodelle	19
5.2.1	Koordinatensysteme	20
5.2.2	Symbolische Identifier	21
5.3	Sensorik	22
5.3.1	GPS oder DGPS	23
5.3.2	Active Badge	23
5.3.3	Active Bat	24
5.3.4	PinPoint	24
5.3.5	Locust	25
5.3.6	Fazit und Klassifizierung	25
5.4	Gebietsdefinition mit Koordinaten	25
5.4.1	Anforderungen	26
5.4.2	2-D Figuren	26
5.4.3	3-D Figuren	27
5.5	Geschachtelte Adressen	28

5.6	Scope und Architektur	29
5.7	Nachrichtenfilterung	30
5.7.1	Homogene Filterung	31
5.7.2	Heterogene Filterung	31
5.8	Aliase	31
5.8.1	Lokale Aliase	32
5.8.2	Globale Aliase	32
5.8.3	Namensschema für globale Aliase	32
5.8.4	Verwendung von Aliasen	32
5.9	Mobile Ziele	34
5.10	Verknüpfungen	34
6	Ausblick	35
7	Literatur	37

1 Einleitung

In den letzten Jahren haben mobile Computer eine große Verbreitung gefunden. Zur gleichen Zeit wurden Systeme zur Ortsbestimmung (z.B. GPS) auch für normale Anwender verfügbar. Schnell wurde klar, daß eine Kombination dieser beiden Technologien einen deutlichen Mehrwert für die Nutzer bieten kann. Durch die Verwendung von mobilen Rechnern, die ihren Aufenthaltsort kennen, kann dem Nutzer die Information präsentiert werden, die gerade relevant für ihn ist. Diese sogenannten ortsbezogenen Anwendungen ("location aware applications") sind derzeit ein wichtiges Forschungsgebiet mit wachsender Bedeutung. Ein Beispiel einer solchen Anwendung sind die Stick-e-Notes der University of Kent [1]. Diese Anwendung erlaubt es dem Nutzer ein virtuelles Objekt (z.B. eine Notiz) an einen Kontext (z.B. eine Position) zu heften. Sobald dieser Kontext wahr wird, wird eine Aktion (z.B. Anzeigen der Nachricht) ausgelöst. Weitere Beispiele für solche Forschungsprojekte und Anwendungen sind interaktive Stadtführer, wie das Guide-Projekt [2] der Lancaster University. Hier stellt das System selbständig fest, wo sich der Nutzer befindet und informiert ihn über die Attraktionen in seiner Nähe. Auch eine Führung, also eine Wegbeschreibung zur nächsten Attraktion ist möglich. An der Universität Stuttgart wird derzeit mit dem Projekt Nexus [3,4] an einer Plattform für solche ortsbezogene Anwendungen gearbeitet.

Wenn die Information über den aktuellen Aufenthaltsort des Nutzers vorhanden ist, liegt es nahe diese auch für die Kommunikation zu verwenden. Ähnlich wie beim Multicast ist es auch beim sogenannten GeoCast möglich, mehreren Empfängern eine Nachricht zukommen zu lassen. Nur müssen die Empfänger beim GeoCast keiner Gruppe beitreten. Der Sender spezifiziert viel mehr das Zielgebiet der Nachricht, und alle potentiellen Empfänger, die sich in diesem Gebiet befinden, erhalten die Nachricht.

Im Bereich GeoCast sind vor allem zwei Bereiche noch nicht intensiv betrachtet worden. Zum einen ist die Frage der Adressierung des Zielgebietes noch offen. Also wie sehen die Adressierungskonzepte aus, mit denen ein Sender das Zielgebiet spezifizieren kann. Die andere Frage die noch nicht sauber betrachtet wurde, ist die Semantik. Welche Semantik hat eine geographische Nachricht. Was sind Alternativen, und warum sind welche Alternativen am sinnvollsten. Diese Fragen sollen in dieser Arbeit näher betrachtet werden.

2 Verwandte Arbeiten

2.1 Protokolle aus Dataman-Projekt

Die allgemeinen Konzepte für GeoCast wurden hauptsächlich von Imielinski und Navas im Rahmen des Dataman-Projektes [5] an der Rutgers University entwickelt.

Im folgenden werden diese Konzepte erläutert. Da die Adressierung bei allen Ansätzen gleich aussieht, wird diese zuerst gesondert betrachtet.

2.1.1 Adressierung

Alle bisherigen Lösungen für GeoCast nutzen die gleiche Art der Adressierung. Die *Adresse* setzt sich aus den Eckpunkten des kleinsten geschlossenen Polygons zusammen, das das Zielgebiet umschließt [6].

Für jeden der Eckpunkte wird die geographische Position angegeben, wobei für Länge und Breite je 4 Byte vorgesehen sind. Als Koordinatensystem wird das weit verbreitete WGS84 benutzt, das auch für GPS eingesetzt wird. Diese Art der Adressierung erlaubt nahezu beliebig geformte Zielgebiete. Es ist aber nicht möglich, ein dreidimensionales Zielgebiet zu definieren. Für häufig genutzte Zielgebiete erlauben Navas und Imielinski auch die Vergabe von Namen, wie z.B. Rathaus.Esslingen.B-W.de.



Abbildung 2-1. Beispiel für ein mögliches Zielgebiet: Die Fußgängerzone von Stuttgart, dargestellt als weißes Polygon.

Im folgenden werden nun die einzelnen Protokolle betrachtet.

2.1.2 Geographisches Routing

Das erste Verfahren funktioniert analog zum “normalen” Routing. Nur werden die Pakete hier nicht nach IP-Adressen, sondern nach geographischen Regionen weitergeleitet. “Geographic Routing” [7,8] benötigt die drei Komponenten *GeoRouter*, *GeoHost* und *GeoNode*.

Als *GeoRouter* wird die Komponente bezeichnet, die das eigentliche Routing durchführt. Jeder *GeoRouter* kennt seine Service-Area. Als Service-Area bezeichnet man das Gebiet, welches von den an den *GeoRouter* direkt angeschlossenen (Kommunikations-) Netzen bedient wird, also das Gebiet, in das der Router Nachrichten direkt weiterleiten kann. Die Information über die Service-Areas wird mit Hilfe eines modifizierten “Routing Information Protocol” (RIP), dem sogenannten *GeoRIP*-Protokoll an alle *GeoRouter* verteilt. Diese erstellen auf Basis dieser Informationen ihre Routingtabelle. In dieser wird für jeden *GeoRouter* festgehalten, welche Service-Area er bedient und über welchen *GeoRouter* er am besten erreicht werden kann.

Beim Eingang eines Paketes wird nun in der Routing-Tabelle nachgeschaut, welche *GeoRouter* das Zielgebiet bedienen und welches die nächsten *GeoRouter* auf dem Weg zu diesen sind. Die Nachricht wird dann entsprechend weitergeleitet. Sollte der Router selbst für das Zielgebiet zuständig sein, so wird die Nachricht auf den entsprechenden Netzwerken an alle interessierten Empfänger und vor allem an die entsprechenden *GeoNodes* gesendet.

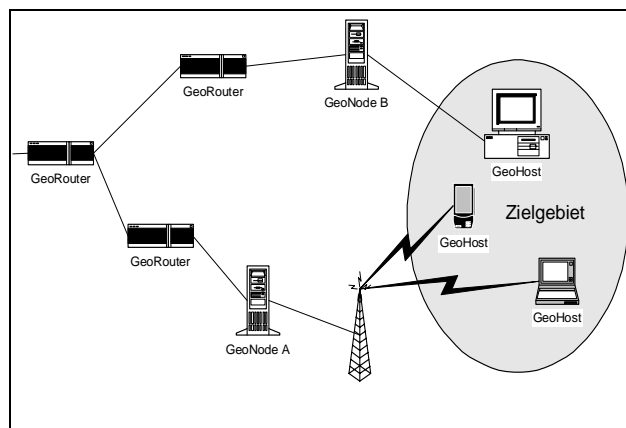


Abbildung 2-2. Geographisches Routing

Damit sichergestellt ist, daß auch neu in das Gebiet kommende Empfänger die Nachricht noch erhalten, hat jede Nachricht eine bestimmte, vom Sender spezifizierbare, Lebensdauer.

Die *GeoNode* ist ein spezifischer Rechner im Zielgebiet, der die Nachricht speichert und regelmäßig wieder an alle Empfänger in den angeschlossenen Netzwerken aussendet, bis die Lebensdauer der Nachricht abgelaufen ist.

Der *GeoHost* ist das Endgerät des Empfängers. Da der Sendebereich des Netzes (z.B. ein Funk-LAN) im Normalfall nicht genau das Zielgebiet einer Nachricht abdeckt, sondern die Nachricht auch noch andere Empfänger erreichen wird, muß der *GeoHost* noch überprüfen, ob er sich im Zielgebiet befindet oder nicht.

2.1.3 Geographische Multicast-Gruppen

Ein anderer Ansatz ist die geographische Adressierung mittels Multicast-Gruppen [6,7]. Hierfür werden sogenannte Partitionen und Atome definiert. Unter einem *Atom* versteht man die kleinste geographische Einheit, die noch eine Adresse hat. *Partitionen* sind größere Gebiete, die sich aus mehreren Atomen und/oder Partitionen zusammensetzen.

Jedem Atom und jeder Partition wird eine Multicast-Adresse zugewiesen. Mitglieder dieser Multicast-Gruppen sind die entsprechenden GeoNodes, die in dem jeweiligen Atom bzw. der jeweiligen Partition liegen. Für Partitionen und Atome, die in einer anderen Partition enthalten sind, werden die entsprechenden GeoNodes auch Mitglieder der übergeordneten Multicast-Gruppe. So wäre z.B. eine GeoNode für den Raum Esslingen, nicht nur Mitglied der Multicast-Gruppe für den mittleren Neckarraum, sondern auch noch der Gruppen für Württemberg, Baden-Württemberg und Deutschland (siehe Abbildung 4).

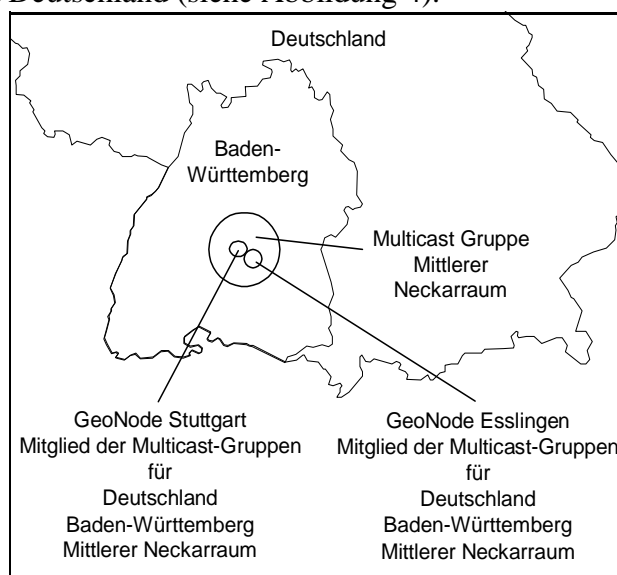


Abbildung 2-3. Beispiel für Gruppenzugehörigkeit

Die GeoNodes senden auch hier, analog zum Geo-Routing, die Nachricht solange auf den entsprechenden lokalen Netzen, bis die Lebenszeit der Nachricht abgelaufen ist.

Wenn nun eine Nachricht gesendet werden soll, dann wird die kleinste Partition gesucht, die das Zielpolygon enthält. Es gibt einen Algorithmus, mit dem die zu einem bestimmten Polygon passende, Multicast-Adresse berechnet werden kann. Die Nachricht wird dann an diese Multicast-Adresse gesendet. Die exakten Informationen über das eigentliche Zielgebiet werden in den Nachrichtenkörper eingefügt. Die entsprechenden GeoNodes erhalten dann die Nachricht und verteilen sie weiter. Wird die Nachricht von einem GeoHost empfangen, so überprüft dieser, ob er sich momentan innerhalb des eigentlichen Zielgebietes befindet. Wenn ja, wird die Nachricht entsprechend verarbeitet, ansonsten wird sie einfach verworfen.

Für die Weiterleitung der Nachrichten wird das Multicast-Routing-Protokoll "Protocol Independent Multicast-Sparse Mode" (PIM-SM) verwendet [9]. Bei diesem Protokoll gibt es für jede Multicast-Gruppe sogenannte Rendezvous-Punkte, an denen die Sender und Empfänger zusammenkommen, d.h. Nachrichten werden vom Sender an diese Rendezvous-Punkte gesendet, und von dort an die Empfänger weiterverteilt. Diese Rendezvous Punkte werden allen PIM-SM-Routern bekanntgemacht. Da die Routing-Tabellen zu groß werden würden, wenn die Mul-

unicast-Adressen und Rendezvous-Punkte für alle Atome und Partionen global bekannt wären, empfehlen Navas und Imielinski hier ein modifiziertes Verfahren. Je größer eine Partition ist, desto weiter werden die Informationen, wie Rendezvous-Punkte, verteilt, bzw. je kleiner eine Partition ist, umso lokaler wird diese Information gehalten.

2.1.4 Domain Name Server

Als einfachste Methode, GeoCast zu realisieren, wird eine Modifikation des Domain Name Server (DNS) vorgeschlagen [7]. Eine neue Domain (.geo) soll dem Namensraum hinzugefügt werden. Eine Anfrage nach Adressauflösung gibt dann eine Liste aller GeoNodes zurück, die das entsprechende Gebiet bedienen. Eine Nachricht könnte dann mittels Unicast an die GeoNodes übermittelt werden. Für größere Nachrichten ist es denkbar, daß die GeoNodes aufgefordert werden, einer temporären Multicastgruppe beizutreten.

2.2 Spezialisierte Protokolle

2.2.1 GeoCast in Ad-Hoc Netzen

Ad-Hoc-Netze, also Netze die spontan und ohne feste Infrastruktur gebildet werden, sind ein relativ neues Forschungsthema. Eine Eigenheit von Ad-Hoc-Netzen ist, daß es keine dedizierte Infrastruktur gibt. Die Infrastruktur wird "ad-hoc" aus den beteiligten Rechnern gebildet. Da die Teilnehmer meist mobil sind ändert sich die Struktur des Netzes also kontinuierlich. Auf der Suche nach geeigneten Routing-Protokollen für Ad-Hoc Netze haben unter anderen Young-Bae Ko und Nitin H. Vaidya vorgeschlagen, die Information über die Lokation der einzelnen Rechner für das Routing zu verwenden [10].

Da in diesem Ansatz die Lokation schon für das Routing von Nachrichten verwendet wird, ist der Schritt zum GeoCast naheliegend. So haben die gleichen Autoren wenig später auch eine Arbeit vorgestellt, in der GeoCast als Multicast an alle Empfänger in einer Region realisiert wird [11].

Da dieser Ansatz aber als Grundlage ein Routing-Protokoll voraussetzt, das es erlaubt Nachrichten nach Lokation weiterzuleiten, ist es nicht einfach auf das Internet übertragbar. Gerade das Routing nach Lokation ist im Kontext des Internets eines der Hauptprobleme für GeoCast.,

3 Model

In diesem Kapitel werden die grundlegenden Annahmen für diese Arbeit definiert.

3.1 Netzwerk

Um GeoCast betrachten zu können benötigt man eine Vorstellung, wie das zugrundeliegende Netzwerk aufgebaut ist.

Wie in Abbildung 3-1 zu sehen, wird davon ausgegangen, dass es verschiedene lokale Netzwerke

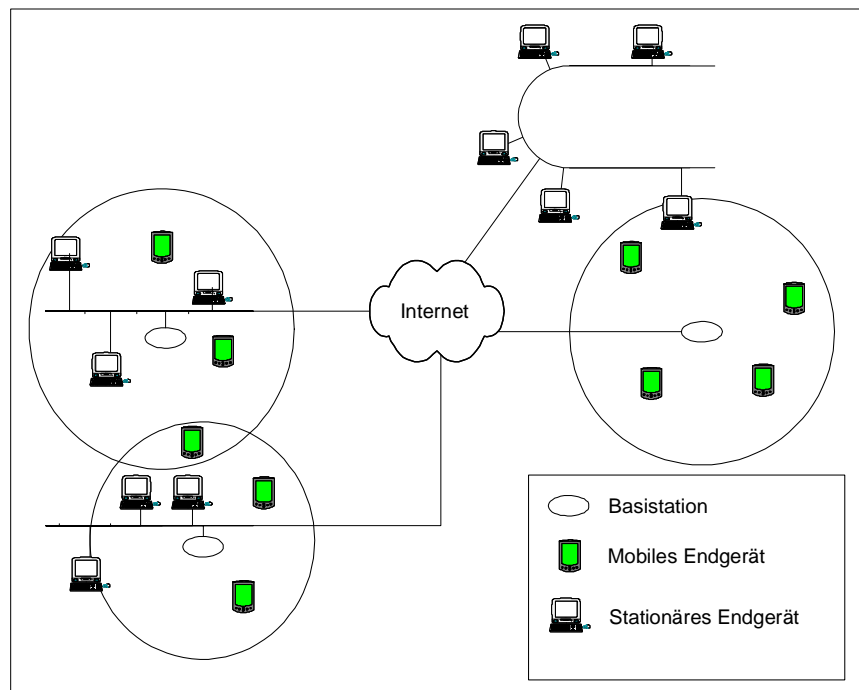


Abbildung 3-1. Beispielnetzwerk

gibt, die alle über das Internet miteinander verbunden sind. Die lokalen Netze können kabelgebunden, funkbasiert oder eine Mischung aus beidem sein. Jedes lokale Netz entspricht auch einem eigenen IP-Subnetz.

Es wird davon ausgegangen, daß jeder Rechner jeweils nur in ein Netz auf einmal eingebucht ist. Mobile Teilnehmer können je nach Ort unterschiedliche Netze nutzen. Stationäre Rechner sind ständig in das gleiche Netz eingebucht.

Weiterhin wird davon ausgegangen, daß die Rechner, die über ein Netzwerk verbunden werden, sich auch in einem begrenzten geographischen Gebiet befinden. Der Fall eines Dial-Ins wird explizit ausgeschlossen.

3.2 Sender

Jedes Gerät das als Sender fungieren soll, benötigt zwingend eine Verbindung zum Netz. Darüberhinaus ist spezielle Software auf dem Sender installiert. Weitere Vorgaben werden an den Sender nicht gestellt.

3.3 Klient

Für den Klienten gelten im Prinzip die gleichen Voraussetzungen wie für den Sender. Hinzu kommt noch, daß der Klient seine augenblickliche Position kennen muß um am GeoCast partizipieren zu können.

4 Semantik

Die Forschung im Bereich GeoCast war bisher sehr technologiegetrieben. Im Fokus stand die Frage, inwieweit die lokationsabhängige Verteilung von Nachrichten überhaupt möglich ist, und wenn das der Fall ist, inwieweit diese Funktionalität effizient realisiert werden kann. Die Frage nach der Semantik wurde bisher noch nicht gestellt. Das heißt derzeit ist die Semantik von GeoCast nur implizit durch die Implementierung bzw. das Design der Protokolle definiert. In diesem Kapitel wird betrachtet wie die Semantik von GeoCast bisher definiert ist und systematisch die sinnvollen Möglichkeiten ausgelotet wie die Semantik sinnvollerweise aussehen sollte.

4.1 Einfaches Model

Wenn man GeoCast nun auf die für die Semantik relevanten Vorgänge reduziert, dann hat man folgendes Bild (siehe Abbildung 4-1):

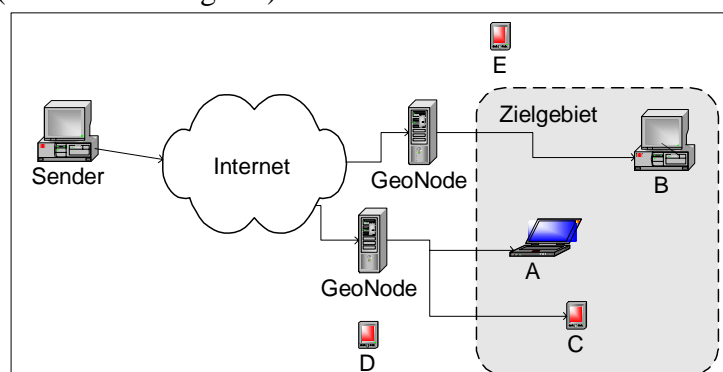


Abbildung 4-1. Schematische Abbildung eines GeoCast

Ein Sender (s) adressiert eine Nachricht (n) an ein geographisches Gebiet. Er hat kein Wissen darüber, wer sich in diesem Gebiet aufhält. Die Nachricht wird vom Sender an die nächste Komponente in Richtung Zielgebiet weitergegeben. Wie diese Komponente aussieht, unterscheidet sich von Protokoll zu Protokoll. Wichtig für die weiteren Betrachtungen ist nur, daß die Nachricht (n) zum Zeitpunkt t_{sn} vom Sender komplett weitergegeben wurde. Die Nachricht wird dann weitergeleitet und erreicht schließlich die Netzwerke im Zielgebiet. Um eine effiziente Nachrichtenweiterleitung im Zielgebiet zu erlauben ist eine Komponente innerhalb des jeweiligen Netzes nötig, die für die lokale Verteilung der Nachrichten sorgt. Im folgenden wird diese Komponente als GeoNode bezeichnet. Da ein Zielgebiet häufig von mehreren Netzen überdeckt wird und die Nachricht jeweils einen anderen Pfad nimmt, wird die Nachricht meist zu unterschiedlichen Zeiten bei den GeoNodes (g) eintreffen. Wir bezeichnen diesen Zeitpunkt im folgenden als t_{gin} , wobei i die unterschiedlichen GeoNodes bezeichnet. Die Nachricht wird dann

von den GeoNodes je nach Protokoll sofort oder etwas verzögert ausgeliefert. Den Zeitpunkt, zu dem die Nachricht auf dem lokalen Netz gesendet wird, bezeichnen wir im folgenden als t_{ej} . Dieser Zeitpunkt ist annäherungsweise identisch mit dem Zeitpunkt zu dem eventuell anwesende Empfänger die Nachricht empfangen haben

4.2 Beispiel

Für die weiteren Erläuterungen ist es hilfreich, ein detailliertes Beispiel zu haben. Darum wird im folgenden die Beispielsituation in Abbildung 4-2 näher beschrieben.

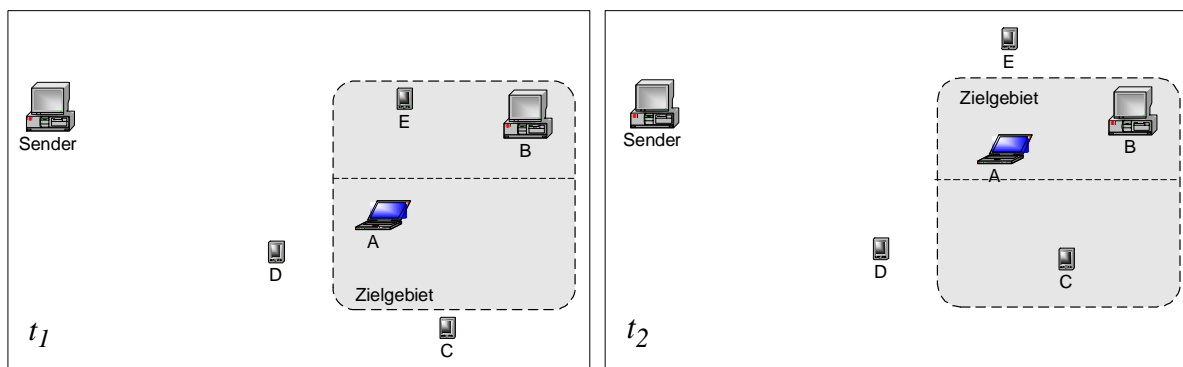


Abbildung 4-2. Beispielsituation für GeoCast.

Die Bilder zeigen die Situation zu den Zeitpunkten t_1 und t_2 . Dazwischen ist der Klient C in das Gebiet hineingekommen und der Klient E hat das Gebiet verlassen. Das genaue Verhalten wird im folgenden beschrieben:

Klient E verläßt das Zielgebiet:
zum Zeitpunkt t_3 ($t_1 < t_3 < t_2$)

Klient C betritt das Gebiet:
zum Zeitpunkt t_4 ($t_1 < t_4 < t_2$)

Wir nehmen weiterhin an, daß E das Gebiet erst verläßt, nachdem C es betreten hat, also $t_3 > t_4$, und somit $t_1 < t_4 < t_3 < t_2$. Dies wird auch in Abbildung 4-3 dargestellt. Die schwarzen Balken bezeichnen die Zeit, in denen die Klienten innerhalb des Zielgebietes sind. Zudem wird das Zielgebiet von 2 LAN's versorgt. Die Trennung wird durch die gestrichelte Linie dargestellt. Zu einem Zeitpunkt t_5 wechselt A das Netzwerk.

Dieses Beispiel wird im folgenden benutzt um die einzelnen Semantiken besser beschreiben zu können.

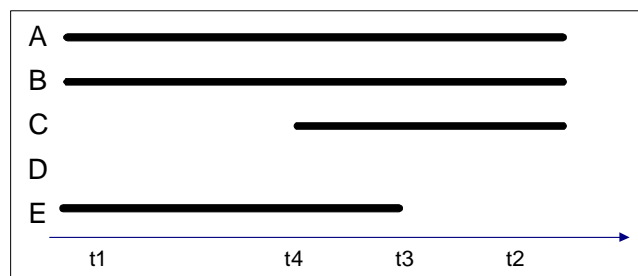


Abbildung 4-3. Klienten im Zielgebiet

4.3 Gültigkeitsdauer einer Nachricht

Eine der speziellen Eigenschaften von Nachrichten in GeoCast ist die sogenannte Lifetime, also die Lebensdauer einer Nachricht. Viele der Nachrichten, die per GeoCast verschickt werden sind nicht nur für die Leute interessant die sich gerade im Gebiet aufhalten, sondern auch für die, die in Kürze das Gebiet betreten. Der Grund hierfür ist die Art der Adressierung. Während bei Unicast der Sender die Nachricht an den interessierten Empfänger direkt schickt, wird bei GeoCast eben nur das Gebiet spezifiziert. Interessierter Empfänger kann dann aber auch die Person sein, die wenig später das Gebiet betritt. Ein Beispiel wäre die Warnung vor einem Stau. Auch die Fahrer, die erst in wenigen Augenblicken in die Gefahrenzone kommen, sollen die Nachricht erhalten. Die Lifetime sorgt also dafür, daß die Nachricht bis zu diesem Zeitpunkt mehrfach gesendet wird, bzw. auch an neu in das Zielgebiet kommende Empfänger verteilt wird.

4.3.1 Definition der Lifetime

Alternativen für die Definition einer "Lifetime" sind:

Die Nachricht

1. hat keine Lebensdauer; sie wird gesendet, sobald dies möglich ist, danach "vergessen". Diese Semantik ist analog zu den Semantiken die aus Uni- oder Multicast bekannt sind. Diese Form der Nachricht reicht für veränderliche Daten wie einen Uhrzeitservice oder den Spielstand bei einem Tennismatch völlig aus. Dieser Fall braucht keinen zusätzlichen Aufwand. Weniger veränderliche Daten wie zum Beispiel Staumeldungen oder Wasserstände müßten dann wiederholt gesendet werden.
2. hat eine absolute Lebensdauer. Bei Staumeldungen könnte dies die Zeit bis zur nächsten Meldung sein; hier wird eine Uhrzeit mitgeschickt. Der Sender oder eine Komponente auf dem Weg zum Empfänger lagert die Nachricht und sendet sie wiederholt, bis der Zeitpunkt gekommen ist, an dem die Nachricht verfällt.
3. hat eine vorgegebene Lebensdauer relativ zum Sendezeitpunkt, innerhalb dieses Zeitraumes wird sie mehrfach gesendet. Nach Ablauf der Zeit, wird die Nachricht verworfen. Wenn ein Unfall passiert, könnte ein Unfallmeldegerät diesen Vorfall an nachfolgende Fahrzeuge melden. Für einen bestimmten Zeitraum wird die Strecke sehr wahrscheinlich nicht geräumt sein. Dieser Zeitraum könnte dann der Nachricht als Lebensdauer mitgegeben werden.
4. wird bis auf weiteres gesendet. Die Nachricht wird solange gesendet, bis eine Einstellungsmeldung kommt. Bei einer Baustelle, die den Verkehr auf einer Straße sperrt, würde bis auf weiteres dieser Sachverhalt an Fahrzeuge im Einzugsbereich der Baustelle gesendet. Durch eine maximale Lebensdauer könnte verhindert werden, das eine Nachricht für immer gesendet wird, nur weil die Abbruch-Nachricht verloren gegangen ist.

BSicherlich wird es auch Nachrichten geben, die ohne Lifetime auskommen. Aber ein Großteil der Nachrichten wird zumindest eine kurze Lifetime haben. Ob man eine absolute Zeit vorgibt, oder aber einen Offset zur Sendezeit angibt ist egal. Mit beiden Methoden kann genau das gleiche erreicht werden, bzw. die eine kann direkt auf die andere abgebildet werden. Um den zu-

ständigen Komponenten das addieren des Offsets auf die Sendezeit zu ersparen sollte man hier die absolute Zeit wählen.

Das letzte Konstrukt, zu senden bis ein Widerruf erfolgt, ist ebenfalls sehr nützlich. Gerade im Bereich Warnungen ist selten bekannt, wie lange eine Gefahrensituation besteht. Allerdings sollten die Werte für die maximale Sendezeit gering gehalten werden. Ansonsten können je nach Realisierung der Funktionalität "Nachrichtenleichen" recht häufig auftreten.

4.3.2 Mindestaufenthaltsdauer eines Empfängers

Eine Bereich der Lifetime, der eher dem Bereich Quality of Service zuzuordnen ist, ist die Fragen nach der Mindestaufenthaltsdauer eines Klienten bis er die Nachrichten bekommt, die noch eine gültige Lifetime haben. Im allgemeinen hängt dieser Parameter sicherlich von der Implementierung und vor allem vom Nachrichtenaufkommen ab. Daher ist es schwer, eine Höchstgrenze für alle Nachrichten anzugeben.

Es ist aber durchaus denkbar, daß für eine begrenzte Zahl von Nachrichten eine solche Dauer angegeben werden kann. So könnten zum Beispiel Warnmeldungen eine höhere Priorität bekommen und damit auch beim Planen der Wiederholungen bevorzugt werden. Dadurch kann sichergestellt werden, das jeder der sich x Sekunden im Bereich aufhält, die entsprechenden Nachrichten erhält. Dies ist aber, im allgemeinen, nur für eine begrenzte Anzahl von Nachrichten möglich.

4.4 Wer ist Empfänger

Eine der zentralen Fragen in jedem Kommunikationsprotokoll ist die Frage wer der Empfänger der Nachricht ist. Bei Unicast wird der Empfänger vom Sender direkt angegeben. Im Falle von Multicast teilt der Empfänger explizit seinen Empfangswunsch mit. Bei GeoCast teilt der Empfänger nur seinen Wunsch mit, GeoCast Nachrichten zu erhalten. Er spezifiziert aber nicht welche. Der Sender adressiert ein Gebiet, und alle in diesem Gebiet sollen die Nachricht bekommen. Aber zu welchem Zeitpunkt muß ein Empfänger nun genau im Zielgebiet sein um sich als Empfänger zu qualifizieren?

4.4.1 Zeitpunkt

Eine offensichtliche Möglichkeit wäre es, die Auswertung zum **Sendezeitpunkt** t_{sn} durchzuführen. D.h. jeder der sich zu diesem Zeitpunkt im Zielgebiet befindet, wird die Nachricht erhalten. In unserem Beispiel wären bei einem Sendezeitpunkt zwischen t_4 und t_3 A,B,C und E Empfänger der Nachricht. Da allerdings für die Übertragung der Nachricht immer etwas Zeit benötigt wird, können potentielle Empfänger das Gebiet verlassen haben und andere das Gebiet betreten haben bis die Nachricht im Zielgebiet ankommt. Wenn die Nachricht also an alle Empfänger gehen soll, die sich zum Sendezeitpunkt im Zielgebiet befunden haben, so muß es eine Möglichkeit geben, festzustellen wer sich zum entsprechenden Zeitpunkt im Gebiet befunden hat.

Eine andere Möglichkeit wäre der **Empfangszeitpunkt**. Doch hier stellt sich die Frage was eigentlich der Empfangszeitpunkt ist. Wie bereits gezeigt, gibt es keinen einheitlichen Empfangs-

zeitpunkt für alle Teilnehmer. Es können die folgenden Zeiten mit Bezug zum Empfang der Nachrichten ausgemacht werden:

1. **Empfangsstart:** Der Empfang hat begonnen, wenn der erste Empfänger die Nachricht erhalten hat, bzw. wenn die Nachricht im ersten Zielnetz angelangt ist.
D.h. alle Klienten empfangen die Nachricht die zum Zeitpunkt $\min(t_{ei})$ über alle i im Zielgebiet sind.
2. **Empfangsende:** Der Empfang endet, wenn die Nachricht beim letzten Empfänger eingegangen ist, bzw. wenn die Nachricht im letzten Zielnetz angelangt ist.
Hier empfangen die Nachricht alle die zum Zeitpunkt $\max(L_{ei})$ über alle i im Zielgebiet sind.
3. **Empfangszeitraum:** Der Empfangszeitraum beginnt mit dem Empfangsstart und endet mit dem Empfangsende.
Die Nachricht wird von denen empfangen, die sich während des Intervalls $[\min(L_{mi}), \max(L_{mi})]$ über alle i im Zielgebiet aufhalten.
4. **Lokale Bereitstellungszeit**
Die Klienten erhalten die Nachricht, die sich zum Zeitpunkt L_{mi} im Netzwerk i befinden.

Wenn die Nachricht an alle ausgeliefert werden soll, die sich zur Empfangszeit im Zielgebiet befinden, dann muß man eine der obigen Möglichkeiten auswählen. Alle Zeiten bis auf die lokale Bereitstellungszeit haben dasselbe Problem. Zum einen muß der Zeitpunkt erst einmal festgestellt werden. Und dann muß festgestellt werden, wer sich zu diesem Zeitpunkt im Zielgebiet aufgehalten hat. Zumindest wenn Ansätze herangezogen werden, bei den die mobilen Endgeräte sich nicht an- und abmelden müssen, ist es schwer diese Forderung zu erfüllen.

Die Semantik die am einfachsten zu realisieren ist, ist die Bestimmung der Empfänger lokal zur lokalen Bereitstellungszeit. D.h. wenn die Nachricht in einem Ziel-LAN ankommt, wird sie dort gesendet. Und wer zu diesem Zeitpunkt Nachrichten aus diesem LAN empfängt, der bekommt die Nachricht. Dies ist die einzige Semantik, die in der "related work" genutzt wird. Das Problem hierbei ist, daß es passieren kann, daß ein Empfänger sich die ganze Zeit im Zielgebiet aufhält, aber die Nachricht nicht erhält, da er in einem ungünstigen Zeitpunkt von einem LAN zum anderen wechselt. Wenn zum Beispiel Klient A im oben eingeführten Beispiel die Netze wechselt nachdem das obere Netz die Nachricht bereits gesendet hat, aber bevor das untere dieses tut, dann wird er die Nachricht nie erhalten, obwohl er die ganze Zeit im Zielgebiet war.

Es gibt natürlich auch noch die Möglichkeit, einen beliebigen Zeitpunkt zur Bestimmung der Empfänger heranzuziehen. Der Sender könnte beim Senden angeben, welche Zeit zur Bestimmung der Empfänger genutzt werden soll. Insbesondere die Möglichkeit einen Zeitpunkt in der Vergangenheit wählen zu können eröffnet interessante Anwendungsmöglichkeiten. So könnte z.B. die Polizei im Falle das Zeugen gesucht werden, einfach einen GeoCast aussenden, und als Zeitpunkt für die Empfängerbestimmung die Zeit des Verbrechens angeben. Um dies Bereitstellen zu können, muß bekannt sein, wer zu welchem Zeitpunkt an welchem Ort war. Dies wird nicht nur schwer zu realisieren sein, sondern wirft auch Probleme im Bereich "Privacy" auf.

Um festzustellen welche Semantik Sinn macht, muß man betrachten was der Sender mit seiner Nachricht bezweckt. Für viele Nachrichten (Werbung, Warnungen, etc.) soll die Nachricht so schnell wie möglich an alle ausgeliefert werden, die sich zum Zeitpunktes des Empfanges im Zielgebiet befinden. Personen die das Zielgebiet verlassen haben, brauchen die Nachricht nicht mehr erhalten, aber Personen die das Ziel betreten sollen die Nachricht bekommen. Das ent-

spricht im Prinzip dem Fall “Lokale Bereitstellungszeit”. Nur kann es hier passieren, daß Personen durch einen Netzwechsel im Zielgebiet die Nachricht nicht erhalten. Dies ist selbstverständlich nicht gewünscht. Alle Empfänger im Zielgebiet sollen die Nachricht sobald als möglich erhalten. Wer das Ziel verläßt soll die Nachricht nicht bekommen. Wer im Zielgebiet bleibt soll die Nachricht auf jeden Fall erhalten.

Daher wird folgende Semantik vorgeschlagen: Analog zum Fall “lokale Bereitstellungszeit” wird die Nachricht in jedem Netz sofort gesendet. Zudem erhält jeder, der in ein Netz kommt in dem die Nachricht bereits gesendet wurde ebenfalls die Nachricht. Somit ist gewährleistet, daß jeder der im Zielgebiet bleibt auf jeden Fall die Nachricht erhält, egal wie die Netzwechsel ausfallen. Zudem läßt sich diese Semantik sehr einfach implementieren, indem man einfach eine kurze Lifetime für jede Nachricht vergibt.

4.4.2 Genauigkeit

Eines der Probleme, wenn GeoCast auch für kleine Gebiete möglich sein soll, ist wie man mit der Ungenauigkeit der Sensorsysteme umgeht. Der Klient muß am Ende immer noch überprüfen, ob er sich wirklich im Zielgebiet aufhält. Wenn das Zielgebiet nun ein Raum ist, dann hat eine Ungenauigkeit von 2m schon eine große Bedeutung. Aber wie kann man nun mit dieser Ungenauigkeit umgehen? Wir gehen davon aus das der Klient die Ungenauigkeit seines Sensorsystems kennt (z.B. GPS 15 m, DGPS 5 m). Es gibt verschiedenen Möglichkeiten wie man vorgehen kann:

Ignorieren der Ungenauigkeit: Die Position die das Sensorsystem liefert, wird als aktuelle Position des Klienten angenommen.

Nachrichtenannahme für Ungenauigkeitsgebiet: Eine Alternative dazu wäre, daß der Klient nicht die Position heranzieht um zu überprüfen, ob er Empfänger ist, sondern daß er die Ungenauigkeit mitberücksichtigt. In Abbildung 4-4 wird ein solches Ungenauigkeitsgebiet gezeigt. Der Klient akzeptiert also Nachrichten für alle Positionen, an denen er sich momentan befinden könnte. Er akzeptiert also alle Nachrichten, bei denen sich Zielgebiet und Ungenauigkeitsgebiet zumindest teilweise überdecken.

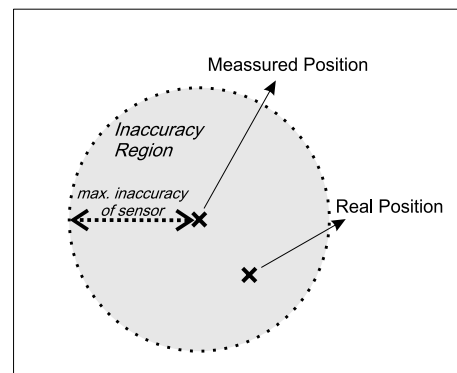


Abbildung 4-4. Ungenauigkeitsgebiet

Empfänger spezifiziert Überdeckungsgrad: Nicht der gesamte Ungenauigkeitsradius wird genutzt, sondern der Empfänger kann einen Prozentsatz angeben, um den sich das Ungenauigkeitsgebiet mit dem Zielgebiet mindestens überdecken muß. Ist dies der Fall wird die Nachricht angenommen, ansonsten wird sie verworfen.

Sender spezifiziert Überdeckungsgrad: Dieser Fall ist analog zum obigen, nur wird die Toleranz hier vom Sender spezifiziert. Ein Beispiel hierfür ist in Abbildung 4-5 zu sehen. Hier hat der Sender einen Überdeckungsgrad von 30% angegeben. Demzufolge werden alle Klienten die Nachricht akzeptieren bei denen sich mindestens 30% des Ungenauigkeitsgebietes mit dem Zielgebiet überdeckt.

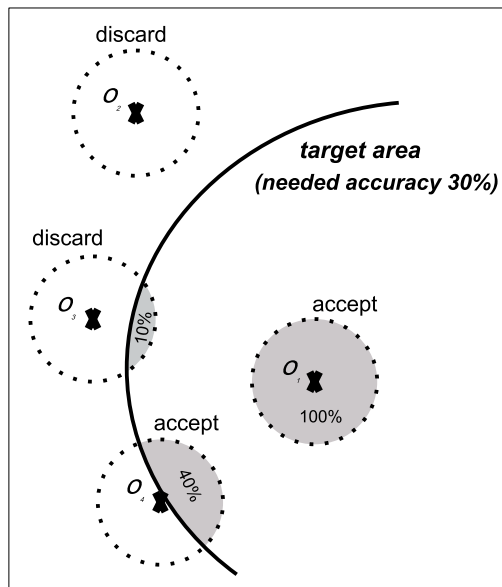


Abbildung 4-5. Senderspezifischer Überdeckungsgrad

Es bleibt die Frage welcher dieser Lösungen am sinnvollsten ist. Sicherlich sollte der Klient jede Warnung für das Gebiet bekommen, in dem er sich befindet. Demzufolge ist die erste Lösung, einfach die Ungenauigkeit zu ignorieren, nicht sinnvoll. Die zweite Alternative, alle Nachrichten zu akzeptieren, die das Ungenauigkeitsgebiet zumindest teilweise überdecken scheint hier besser geeignet.

Allerdings soll der Klient aber auch nicht durch Nachrichten für angrenzende Gebiete abgelenkt werden. So würde z.B. einer der die Straße herunterläuft alle Nachrichten für die angrenzenden Gebäude erhalten. Für Warnungen mag das ja noch akzeptabel sein, nicht aber für Notizen über interne Veranstaltungen oder gar Werbung. Die Person, die weiß, wie wichtig die Nachricht ist, ist der Sender. Deshalb scheint es sinnvoll, dem Sender die Möglichkeit zu geben den Überdeckungsgrad zu spezifizieren. Wenn eine Warnung gesendet wird, dann wird der Überdeckungsgrad klein gewählt und ansonsten groß. Das Ganze kann dann auch noch mit einem empfangerspezifischen Mindestüberdeckungsgrad kombiniert werden. Dadurch kann der Empfänger das System an seine spezifischen Wünsche und Bedürfnisse anpassen. So könnte ein Empfänger angeben, daß der Überdeckungsgrad mindestens 30% betragen muß, und würde gegebenenfalls eine Nachricht verwerfen, obwohl der Sender 1% Überdeckung angegeben hat.

kungsgrad klein gewählt und ansonsten groß. Das Ganze kann dann auch noch mit einem empfangerspezifischen Mindestüberdeckungsgrad kombiniert werden. Dadurch kann der Empfänger das System an seine spezifischen Wünsche und Bedürfnisse anpassen. So könnte ein Empfänger angeben, daß der Überdeckungsgrad mindestens 30% betragen muß, und würde gegebenenfalls eine Nachricht verwerfen, obwohl der Sender 1% Überdeckung angegeben hat.

4.5 QoS

4.5.1 Prioritäten

Prioritäten können ausgehend vom Sender oder vom Empfänger vergeben werden. Wenn es unterschiedliche Nachrichtenarten gibt, kann sich aus der Art der Nachricht eine Priorität ergeben. Prioritäten sind ein Ausdruck des Wertes, der einer Nachricht beigemessen wird und können sich auf die Behandlung dieser Nachricht in Form von bevorzugter Weiterleitung auswirken. Wenn eine Nachricht mit einer Priorität verschickt wird, erfordert dies einen weiteren Parameter.

- Es werden keine Prioritäten vergeben, das heißt, alle Nachrichten werden gleich behandelt. Es gibt keinen zusätzlichen Aufwand. Wenn man bei GeoCast bestimmte Nachrichten vorrangig behandeln möchte, reicht dieser Fall nicht aus.
- Der Sender bekommt eine Priorität zugewiesen. Wenn ein Gerät an GeoCast angeschlossen ist, das auf die Polizei oder den Katastrophenschutz zugelassen ist, würde diesem Gerät zum Beispiel höchste Priorität zugeordnet. Hier kommt ein Verwaltungsaufwand für die Vergabe der Prioritäten hinzu. Man bräuchte eine Stelle, die diese Prioritäten vergibt und möglichst

überwacht. Nachrichten von einer “wichtigen” Quelle könnten diese Wichtigkeit in Form einer Priorität zugeordnet bekommen.

- Der Sender vergibt Prioritäten. Bei der Heimfahrt fällt jemandem zum Beispiel auf, daß er seine Brieftasche im Biergarten vergessen hat. Er möchte jetzt eine Nachricht mit sehr hoher Priorität an Personen in diesem Biergarten schicken, ob seine Brieftasche gefunden wurde. Die Einschätzung der Wichtigkeit einer Nachricht läge in diesem Fall beim Sender. Die Priorität würde der jeweiligen Nachricht entsprechend ihrer Wichtigkeit zugeordnet. Diese Möglichkeit läßt den Sender flexibler Arbeiten, als bei einer Zuordnung der Priorität zum Sender. Diese Möglichkeit läßt allerdings auch Mißbrauch zu. Daher müßte ein Schutz eingebaut oder das System überwacht werden.
- Die Nachricht erhält implizit eine Priorität. Wenn es Nachrichtengruppen gibt, kann zum Beispiel eine Gefahrenwarnung automatisch mit der höchsten Priorität weitergereicht werden. Diese Priorität hätte sie aufgrund ihrer Zugehörigkeit zu den Gefahrenwarnungen. Dieser Fall ist der Vergabe von Prioritäten durch den Sender sehr ähnlich. Die Nachricht erhält eine Priorität vom Sender (hier implizit durch die Nachrichtenart). Es bedarf auch in diesem Fall einer Kontrolle, welcher Sender welche Prioritäten vergeben darf. Dies könnte durch die Vergabe von Rechten geregelt werden. Auch müßten Nachrichten in Gruppen eingeteilt sein, was einen weiteren Parameter mit sich bringen würde.

Welche Auswirkungen die Priorität auf die Nachricht hat, wird hier nicht weiter beschrieben. Im Prinzip können dadurch Variationen an den verschiedenen Parametern vorgenommen werden. So könnte z.B. eine Nachricht mit höherer Priorität beim Ausliefern während der Lifetime bevorzugt behandelt werden. (Siehe „Mindestaufenthaltsdauer eines Empfängers“ auf Seite 12.)

4.5.2 Verlust

Die prinzipiellen Alternativen sind hier:

- keine Garantie (analog zu IP/UDP)
- mindestens ein Empfänger erhält die Nachricht
- die Nachricht wird von allen Empfängern empfangen (analog zu TCP)
- Die Nachricht wurde an alle LAN's ausgeliefert und dort gesendet

Zu Garantieren, daß es keinen Nachrichtenverlust gibt, bzw. das Nachrichtenverlust erkannt wird, ist schwierig wenn sich die Teilnehmer nicht anmelden müssen. Da nicht bekannt ist, welche Endgeräte die Nachricht erhalten sollen, kann auch nicht geprüft werden, ob sie sie erhalten haben. Anders sieht es mit dem letzten Fall aus. Je nach Realisierung des GeoCasts sollte es möglich sein, diese Garantie geben zu können. Insbesondere wenn man dann noch durch eine entsprechende Lifetime dafür sorgt, daß die Nachricht auf dem LAN mehrmals gesendet wurde, kann man davon ausgehen, daß die meisten Empfänger die Nachricht erhalten haben. Dieses sollte für die meisten Anwendungen ausreichen. Mehr zu garantieren ist nur möglich wenn sich die Endgeräte anmelden bzw. wenn es einen Service gibt, der weiß, welche Endgeräte sich momentan im Zielgebiet befinden.

4.5.3 Reihenfolge

Hier gibt es im Prinzip vier Alternativen:

- Es wird keine Aussage über die Reihenfolge von Nachrichten gemacht (analog zu IP/UDP)
- Nachrichten von einem Sender kommen beim Empfänger in der gleichen Reihenfolge wie sie gesendet wurden. (analog zu TCP)
- Es wird erkannt, wenn eine Nachricht von einem Sender empfangen wurde, von dem bereits eine später gesendete Nachricht empfangen wurde.
- Nachrichten kommen in der Reihenfolge an, in der sie gesendet wurden, auch wenn sie von verschiedenen Sendern kommen.

Vor allem wenn man Nachrichten hat, die sich auf eine Zustandsänderung beziehen, ist die Reihenfolge der Nachrichten wichtig. So sollte die Reihenfolge von Warnungen und Entwarnungen nicht vertauscht werden. Da es aber nicht trivial ist, die Reihenfolge zu garantieren, und viele Anwendungen dies auch nicht benötigen, sollte die Anwendung entsprechende Maßnahmen treffen wenn nötig.

4.5.4 Nachrichtenlaufzeit

Ein weiterer Parameter für die Quality of Service ist die Nachrichtenlaufzeit. Es ist denkbar, dass Nachrichten mit höherer Priorität schneller befördert werden als andere. Hierbei würde keine harte Aussage über die maximale Laufzeit gemacht, sondern nur festgelegt dass die Nachricht bevorzugt weitergeleitet wird. Dies entspricht dem Differentiated Service im Internet.

Es ist natürlich auch denkbar, daß es Implementierungen gibt, die es dem Sender erlauben eine harte Maximallaufzeit anzugeben. Allerdings müßte hier vor dem Sendevorgang ausgehandelt werden, ob diese Bedingung erfüllt werden kann. Hierzu ist es nötig, daß eine Nachricht den kompletten Weg zum Ziel geht und die Ressourcen reserviert. Da dieses logischerweise mindestens so lang dauert wie der eigentliche Versand der Nachricht, lohnt sich dies nur im Falle, daß viele Nachrichten versendet werden sollen.

5 Adressierung

5.1 Einführung

Ein zentraler Aspekt bei GeoCast ist die Frage wie das Ziel für die Nachricht spezifiziert wird. Also die Frage nach der Adressierung. Auch mit dieser Frage hat sich die Forschung bisher kaum beschäftigt. Bisher werden zur Angabe des Zielgebietes Polygone verwendet, deren Eckpunkte durch Angabe der Länge und Breite im WGS84 Format definiert sind. Die Auflösung der Koordinaten wurde so gewählt, daß eine maximale Genauigkeit von 160 Metern möglich ist. Für die Frage, ob GeoCast überhaupt technisch machbar ist, sind solche Adressen durchaus ausreichend. Doch wenn man nicht mehr die Frage nach der technischen Realisierbarkeit von GeoCast betrachtet, sondern Protokolle für den praktischen Einsatz designt, dann sind solche Adressen für die meisten Anwendungszwecke nicht geeignet. Zudem wird durch die Art der Adressierung GeoCast eigentlich nur für Outdoor-Szenarien nutzbar. Doch gerade die Indoor-Nutzung bietet gute Perspektiven für GeoCast.

In diesem Kapitel soll daher systematisch betrachtet werden, was für Möglichkeiten für die Adressierung einer lokationsabhängigen Nachricht es gibt, und welche dieser Alternativen sinnvoll wären. Man darf dabei allerdings nicht vergessen, daß ein Großteil der Anwender diese Adressen nie zu Gesicht bekommt, da die Anwendung meist eine graphische Eingabemöglichkeit (z.B. eine Karte) anbieten wird. Die Anwendung wird diese Eingabe dann aber auf eine entsprechende Adresse abbilden müssen.

5.2 Weltmodelle

Eine Adresse bezeichnet einen Ort in der Welt. Um eine Adresse angeben zu können, braucht man eine Vorstellung von der Welt. Insbesondere wenn mehrere Einheiten die Adresse auf gleiche Weise verstehen sollen ist es wichtig, daß alle Beteiligten das gleiche Modell verwenden. Die Fragen über die Einigkeit bestehen muß, sind u.a.:

- Wie sieht die Welt aus?
- Wie kann man einen Punkt in dieser Welt angeben?
- Wie wird ein Gebiet dort bezeichnet?

In dieser Sektion wird der Frage nachgegangen was für Weltmodelle vorstellbar sind. Für die verschiedenen Modelle wird außerdem die Eignung für lokationsabhängige Kommunikation betrachtet.

5.2.1 Koordinatensysteme

Eines der gebräuchlichsten Modelle ist es, die Welt mit einem Koordinatensystem zu versehen. Ein Punkt wird dann durch die entsprechenden Koordinaten gekennzeichnet. Dieses Prinzip begegnet uns in jedem Atlas oder auf jeder Straßenkarte.

Es gibt jedoch eine Vielzahl verschiedener Koordinatensysteme, die unterschiedliche Eigenschaften haben und auch unterschiedlich gut für GeoCast geeignet sind. Die Koordinatensysteme lassen sich in drei Kategorien einteilen:

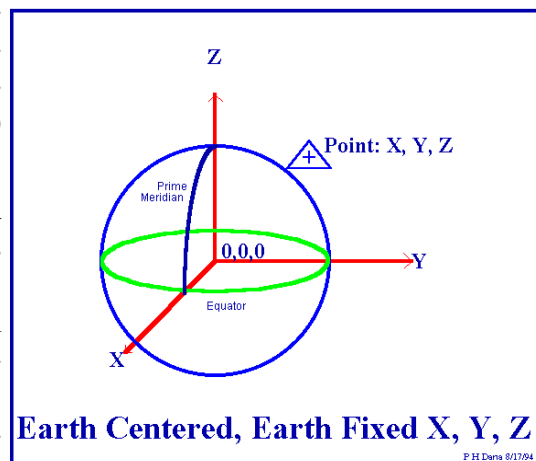
- Geozentrische Koordinatensysteme
- Ellipsenbasierte Koordinatensysteme
- Lokale Koordinatensysteme

Diese Kategorien werden im folgenden etwas detaillierter betrachtet.

5.2.1.1 Geozentrische Koordinaten

Ein geozentrisches Koordinatensystem (siehe Abbildung 5-1) hat seinen Ursprung im Erdmittelpunkt. Ein Punkt wird durch seine x,y und z Koordinaten (als Entfernung vom Mittelpunkt) bezeichnet.

Da die Koordinaten relativ zum Erdmittelpunkt und nicht relativ zur Oberfläche definiert sind, ist es schwierig Bezüge zur Erdoberfläche herzustellen. Ob z.B. zwei Punkte die gleiche Entfernung vom Boden haben, oder ob 3 Punkte auf einer „Oberflächenlinie“ liegen läßt sich nur schwer feststellen. Da für GeoCast natürlich immer der Bezug zur Erdoberfläche vorhanden ist, sind diese Koordinatensysteme weniger geeignet.



Earth Centered, Earth Fixed X, Y, Z

Abbildung 5-1. Geozentrische Koordinaten

5.2.1.2 Ellipsenbasierte Koordinaten

Um eine Adressierung in Abhängigkeit von der Erdoberfläche zu ermöglichen muß ein Modell der Erde zugrunde gelegt werden. Man hat festgestellt, daß die Erde keine echte Kugel ist. Als Näherung verwendet man Ellipsoide. Die Koordinatensysteme unterscheiden sich u.a. darin, welcher Ellipsoid verwendet wird. Ein weiteres Unterscheidungskriterium ist die Methode durch die Orte auf Koordinaten abgebildet werden. Die zwei gebräuchlichsten ellipsenbasierten Systeme werden im folgenden kurz betrachtet.

Gauß-Krüger. Das Gauß-Krüger Koordinatensystem basiert auf sogenannten „Transversalen Merkatorprojektionen“. Die Erdoberfläche wird hierbei nicht mit einem einheitlichen Koordinatensystem überzogen, sondern streifenweise auf eine Ebene abgebildet. Die Koordinaten sind jeweils pro Streifen definiert, da die Streifen sich überlappen kann ein Punkt mehrere Koordinaten haben kann. Zwei Punkte auf zwei verschiedenen Streifen haben Koordinaten, die keinen

direkten Bezug mehr zueinander haben. Zudem wird je nach Streifen ein geeigneter Ellipsoid ausgewählt.

Je nach verwendetem Streifen hat z.B. die Michaelis-Kirche in Lüneburg die folgenden Koordinaten:

- System des 9. Längengrades:
Rechtswert: 3 593 571,20 Hochwert: 5 902 863,21
- System des 12. Längengrades:
Rechtswert: 4 393 360,64 Hochwert: 5 903 137,40

Da mit Gauß-Krüger keine einheitliche fortlaufende Adressierung möglich ist, ist es für GeoCast weniger geeignet. Insbesondere, wenn Gebiete bezeichnet werden sollen, die über die Streifen-Grenzen hinweggehen, ist dies mit Gauß-Krüger nicht möglich.

WGS84. Auf der Basis des WGS84 Ellipsoiden gibt es ein Koordinatensystem, in welchem ein Punkt durch die Angabe der Länge und Breite bezeichnet wird. Durch Hinzufügen der Höhe über Normalnull kann jeder Punkt auf oder über der Erde adressiert werden. Dieses Koordinatensystem ist global einheitlich und hat fortlaufende Koordinaten. Es wird unter anderem von GPS genutzt.

Da mit WGS84 auch größere Gebiete einheitlich beschreiben werden, ist es gut geeignet für-GeoCast. Insbesondere da die Umwandlung von Koordinaten aus einem der anderen Koordinatensystemen in das ellipsenbasierte WGS84 Format kein Problem darstellt, ist es sinnvoll dieses Format als Standardkoordinatensystem für GeoCast zu verwenden.

5.2.1.3 Lokale Koordinatensysteme

Neben Koordinatensystemen die für jeden Punkt der Erdoberfläche eine eindeutige Koordinate anbieten, ist es auch möglich lokale Koordinatensystemen definieren. Hierzu wird ein beliebiger Punkt als Ursprung gewählt und andere Punkte durch die Entfernung in x,y und z Richtung zu diesem Punkt definiert. Hierzu ist es natürlich nötig die Lage der x,y und z Achse festzulegen und einen Maßstab vorzugeben. Nützlich sind lokale Koordinatensysteme, unter anderem, im Fall von mobilen Objekten. Wenn z.B. eine Stelle in einem Zug bezeichnet werden soll, so ist es sicherlich nicht sinnvoll diese in WGS84 Koordinaten anzugeben. Hierbei müßten die Koordinaten ständig geändert werden, da der Zug sich bewegt. Hier ist es praktischer, wenn man ein lokales Koordinatensystem definiert. So wird z.B. ein Punkt im Zug der Referenzpunkt, und alle anderen werden relativ zu diesem angegeben.

5.2.2 Symbolische Identifier

Symbolische Lokationsmodelle bezeichnen die Orte mit abstrakten Symbolen oder Namen wie z.B. "*Raum xyz*", "*Campus der Uni-Stuttgart*" oder "*St. Pauls Kirche*". In einem solchen Modell hat jeder (interessante) Ort einen Namen. Im Bereich des Modells muß dieses Symbol eindeutig sein. Die verschiedenen Ausprägungen werden im folgenden betrachtet. Der Nachteil eines solchen Modells soll nicht verschwiegen werden, es ist nicht möglich, Lokationen zu bezeichnen für die kein Symbol vorgesehen ist. Wenn also ausgedrückt werden soll, daß irgend-

wo auf einer Wiese etwas passiert, dann gibt es keine Möglichkeit diesen Ort anzugeben, wenn kein Symbol dafür definiert wurde.

Je nach Wahl des Schemas zur Bezeichnung der Lokationen können Beziehungen der Lokationen zueinander ausgedrückt werden.

5.2.2.1 Flaches Schema

Jeder Ort hat einen Namen, der ihn eindeutig bezeichnet. Es werden keinerlei Aussagen über die Beziehungen dieser Orte bzw. ihrer Symbole zueinander gemacht.

Ein Beispiel für ein flaches Adressierungsschema wäre im Fall einer Wohnung die Bezeichner:

- Küche
- Bad
- Wohnzimmer
- Esszimmer
- Flur

5.2.2.2 Hierarchisches Namensschema

Es ist möglich, die Symbole zu ordnen. Es bietet sich die „enthaltensein“ Relation an. Diese entspricht der Wahrnehmung im täglichen Leben. So ist Raum 2.018 im 2. Stock des Gebäudes Breitwiesenstr. 20-22.

Eine solche Ordnung kann auch zur Bildung von Namen verwendet werden. So läßt sich einfacher sicherstellen das alle Lokationen der Universität Stuttgart eindeutige Namen haben, wenn ihre Namen bezug auf die Objekte nehmen, in denen sie enthalten sind. Ein Beispiel wäre „Breitwiesenstrasse.2.018“.

Theoretisch wäre es denkbar ein einheitliches weltweites ID-Schema zu haben. Praktisch wird es das jedoch nicht geben. Stattdessen werden wohl lokal begrenzte Modelle genutzt werden. Dabei entscheidet sich eine Institution (z.B. die Universität), ein solches Modell einzuführen. Die Symbole sind dann nur innerhalb dieses lokal begrenzten Bereichs definiert.

5.3 Sensorik

Die Funktionalität von GeoCast und vor allem die Möglichkeiten zur Adressierung hängen auch immer von den verfügbaren Positionssensoren ab. Um die Anforderungen, die sich aus den Sensorsystemen ergeben, herauszuarbeiten, werden im folgenden die gebräuchlichen Systeme kurz betrachtet. Besonderer Wert wird darauf gelegt, in welcher Art und an welcher Stelle die Positionsangabe vorhanden ist. Dieses ist ein wichtiges Kriterium bei der Auswahl der Adressierungsart. Sicherlich wird es in Zukunft andere Positionierungsmethoden geben, aber die Grundprinzipien werden wahrscheinlich die gleichen bleiben. Von daher sollte die Betrachtung der jetzt erhältlichen bzw. in der Forschung verfügbaren Systeme einen Überblick über die grundsätzlichen Arten geben.

5.3.1 GPS oder DGPS

Funktionsweise

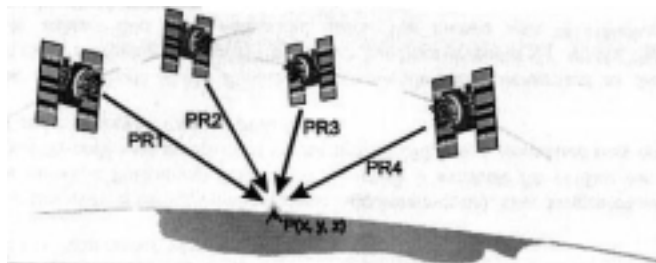


Abbildung 5-2. Funktionsweise von GPS

In Abbildung 5-2 wird das Prinzip der Positionsbestimmung mittels GPS dargestellt. Mit Hilfe von vier Satelliten kann der Empfänger, anhand der bekannten Positionen der Satelliten, seine eigene Position errechnen. Für einen korrekten Empfang der Signale dürfen sich keine Hindernisse zwischen Empfänger und Satelliten befinden, ansonsten fällt die Positionsbestimmung ungenau aus.

Aufgrund des kalten Krieges zwischen den USA und Rußland wurden die Satellitensignale so verschlüsselt, daß „nur“ eine Genauigkeit von +/- 50 Meter mit den gewöhnlichen GPS erreicht werden konnte. Diese künstliche Verschlechterung („selective availability“) des Signals wurde Mitte 2000 abgeschaltet. Inzwischen liefert GPS eine Genauigkeit zwischen 5-15 Metern.

Um die Genauigkeit der Positionsbestimmung mit GPS zu erhöhen, werden zusätzlich zu den Satellitensignalen noch Korrekturdaten benötigt. Diese gleichen die Fehler aus, die durch die „selective availability“ und durch atmosphärische Störungen, entstehen. Korrekturdaten werden von einer Referenzstation erzeugt, die an einem genau vermessenen Punkt einen Soll/Ist-Vergleich durchführt. Ein DGPS-fähiger GPS-Empfänger, dem diese Daten zugeführt werden, ist dann in der Lage, eine genauere Positionsbestimmung durchzuführen.

Output. Der Klient erhält die aktuelle Position in einem global eindeutigen Koordinatensystem (WGS84). Die Genauigkeit variiert je nach verwendetem System und nach Standort, liegt aber zwischen 2 bis 15 Meter. Die Auflösung des Koordinatensystems ist beliebig. GPS und DGPS sind nur „Outdoor“ verfügbar.

5.3.2 Active Badge

Das Active Badge System der AT&T Laboratories in Cambridge ist seit 1992 im Einsatz.

Funktionsweise: Das Active Badge System besteht im Prinzip aus 3 Komponenten. Es gibt das „Badge“, das der Nutzer bei sich trägt. Dieses sendet alle 15 Sekunden einen Impuls (Infrarot) aus. Dieses Signal wird von *Sensoren* aufgefangen, die im Gebiet verteilt sind. Diese melden einen Empfang des Signals an einen *Server*. Da Active Badge Infrarot benutzt, ist die kleinste Einheit ein Raum oder in großen Räumen die Sendereichweite.

Output: Die Information die auf dem Server vorhanden ist, ist zuerst einmal nur eine Verknüpfung von Badge-ID mit Sensor-ID. Durch entsprechende Programmierung können diese Rohdaten dann zu „sinnvollen“ Daten gewandelt werden. Z.B. wurde an den AT&T-Labs eine Anwendung für Telefonisten geschrieben die auf Basis dieser Daten eine Zuordnung macht, in der Nähe welches Telefons sich eine Person gerade befindet. Ein anderes Beispiel ist eine Zuordnung Personen → Raum, wobei häufig auch Raumteile genommen werden („Floor 1 near stairs“¹). Ich habe kein Beispiel gefunden wo eine explizite Objekt-Hierarchie definiert wurde, aber denkbar wäre es, die Namen bzw. Symbole in eine Hierarchie einzubetten.

Wichtig ist, daß der Client eigentlich keine Information über seinen Aufenthaltsort hat. Er kann diese, sofern das vorgesehen ist, vom Server abfragen. Der Server kann die Daten in beliebiger Form aufbereitet anbieten. Das eigentliche Wissen ist Badge-ID→Sensor-ID.

5.3.3 Active Bat

Das Active Bat System wurde ebenfalls von den AT&T Labs in Cambridge entwickelt. Es ist das Nachfolgesystem zum Active Badge System.

Funktionsweise: Auch im Active Bat System gibt es drei Komponenten. Es gibt wieder ein Tag, das der Anwender bei sich trägt. Die Sensoren sind beim Active Bat System im Abstand von 1,2 Metern an der Decke befestigt. Für jeden Sensor ist die genaue Position bekannt. Und es gibt auch wieder einen Steuerrechner.

Der Steuerrechner sendet einen Impuls an das Tag, das geortet werden soll. Dieses antwortet mit einem Ultraschall-Signal. Die Sensoren fangen dieses auf. Aufgrund der Laufzeit des Signals berechnen die Sensoren den Abstand des Tags (siehe Abbildung 5-3). Der Steuerrechner fragt dann bei den Sensoren nach dem Abstand. Mit mindestens 3 Abständen kann die Position des Tags bis auf 15 cm genau ermittelt werden.

Output: Auch hier liegen die Positionsdaten in der Infrastruktur bzw. dem Steuerrechner. Für jedes Objekt liegen die 3-D Daten seiner Position vor, oder können angefragt werden. Die Position kann dann in jedes beliebige Koordinatensystem umgewandelt werden. Welche Koordinaten nativ verwendet werden, wird nicht näher spezifiziert, hängt aber sicherlich davon ab welche Koordinaten die einzelnen Sensoren haben.

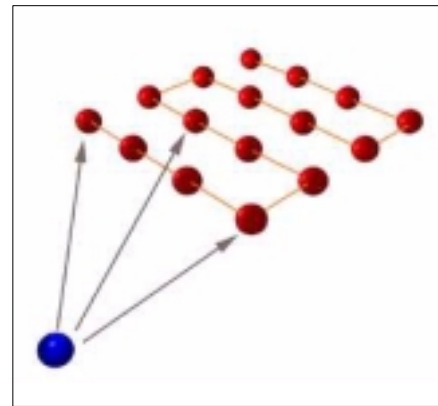


Abbildung 5-3. Funktionsweise des Active Bat Systems

5.3.4 PinPoint

PinPoint ist ein Beispiel für funkbasierte Positionierung. Es gibt mehrere Systeme die auf dieser Technik basieren. Die gemachten Aussagen über PinPoint treffen auch auf diese Systeme zu.

Funktion: Es gibt bei PinPoint mehrere Stationen die die Funksignale eines Tags empfangen. Die Information über Signallaufzeit werden an einen Server gesendet, der daraus die Position ermittelt.

Output: Die Position wird als Koordinate in einem Koordinatensystem angegeben. Genaue Informationen was für Koordinatensysteme verwendet werden, habe ich nicht in Erfahrung bringen können. Prinzipiell sollte es aber kein Problem sein, die Daten in beliebige Koordinatensystem umzurechnen. Auch hier sind die Positionsdaten nur am Server verfügbar. Nicht der Klient ermittelt seine Position, sondern die Position des Klienten wird von der Infrastruktur ermittelt.

1. Einer Anfrage auf der Website von AT&T Lab. entnommen.

5.3.5 Locust

Locust ist ein Positionierungssystem das am MIT entwickelt wurde.

Funktion: Locust basiert auf Infrarotbaken. Es werden Geräte verteilt, die über Infrarotsignale eine ID versenden. Der Klient empfängt das Signal und kann dann aufgrund der ID eine Abbildung vornehmen, um seine Position zu ermitteln. Die Abbildung kann selbstverständlich beliebig sein. Eine neuere Entwicklung verwendet statt Infrarot Funksignale und verbessert so die Verfügbarkeit des Systems.

Output: Der Klient erhält eine ID. Was er mit der ID macht, ist seine Sache. Wenn sichergestellt wurde, daß jede ID nur einmal vorhanden ist, dann kann der Klient diese ID für eine Abbildung ID→Lokation verwenden. Diese Abbildung bzw. die Tabelle muß „von Hand“ konfiguriert werden. Selbstverständlich kann jede beliebige Abbildung auf dem Client vorgenommen werden.

5.3.6 Fazit und Klassifizierung

Es gibt zwei Merkmale von Positionierungssystemen die für diese Arbeit wichtig sind.

1. Ermittelt der Klient oder die Infrastruktur die Position? Also, wo ist das Wissen über die Position der Klienten vorhanden?
2. In welcher Form ist Wissen über die Position vorhanden?

Wo ist das Wissen: Bei nahezu allen Systemen, die in Gebäuden nutzbar sind, liegt die Information über die Lokation des Klienten auf einem Server. Der Klient selber muß seine Position dort erfragen. Das heißt, hier bietet ein Server die Positionsdaten des Klienten an. Die Daten werden auch schon auf dem Server aufbereitet, d.h. eventuelle Abbildungen werden nicht auf dem Client sondern auf dem Server vorgenommen. Da für die Nutzung von GeoCast der Klient eine Verbindung zum Internet benötigt, sollte es kein Problem darstellen, die Positionsangabe vom Server an den Klienten zu übermitteln.

Was für Wissen ist vorhanden: Im Prinzip gibt es drei Formen in denen das Grundwissen über die Lokation vorhanden ist.

- Globale geographische Koordinaten (GPS, DGPS)
 - Lokale Koordinaten (Active Bat, PinPoint)
 - Nächster Sensor oder nächstes Badge (Active Badge, Locust)
- Da die Information, welcher Sensor in der Nähe ist, nicht viel aussagt, wird meist eine Abbildung auf „sinnvolle“ Information vorgenommen. So wird ein Badge mit einem Raum oder Raumteil verknüpft. Es kann aber genausogut mit einer Position, Telefonnummer, Terminal-ID, ... verknüpft werden. Selbstverständlich können die verknüpften Symbole die Blätter einer Hierarchie sein. Im Prinzip findet eine Abbildung auf ein Weltmodell statt.

5.4 Gebietsdefinition mit Koordinaten

Wenn ein koordinatenbasiertes Weltmodell verwendet wird, dann ist es erstmal nur möglich, einzelne Punkte darin zu bezeichnen. Dies kann natürlich verwendet werden um das Zielgebiet

als geometrische Figur zu beschreiben. Auch hier ist es wieder zwingend notwendig, daß alle Beteiligten das gleiche Verständnis über die möglichen Figuren haben. In diesem Abschnitt werden nun die verschiedenen Figuren auf ihre Eignung für GeoCast hin betrachtet. Zuvor möchte ich allerdings kurz auf die Anforderungen eingehen, die an diese Figuren gestellt werden.

5.4.1 Anforderungen

Für das Adressierungsmodell ergeben sich folgende Anforderungen:

- **Gute Annäherung**
Der Sender möchte eine Nachricht in ein bestimmtes Gebiet schicken. Die Figuren, die er zur Spezifikation des Zieles zur Verfügung hat, sollen eine möglichst genaue Angabe des gewünschten Zielgebietes erlauben. Dadurch wird verhindert, daß die Nachricht an Klienten geht, die die Nachricht nicht erhalten sollten, bzw. daß Klienten die Nachricht nicht bekommen obwohl sie im Gebiet sind, das der Sender als Empfangsgebiet definieren möchte.
- **Einfache Verarbeitung**
Einige der Klienten werden nur geringe Rechenleistung besitzen. Von daher sollten keine komplexen Berechnungen nötig sein, um festzustellen, ob sich eine Position innerhalb der Figur befindet. Auch sollte es für die Komponenten, die die Nachricht weiterleiten (z.B. GeoRouter), möglich sein, die Forwarding-Entscheidung möglichst schnell zu treffen. Auch hier sollten also komplexe Berechnungen vermieden werden.
- **Einfach Beschreibbar**
Die Adresse sollte möglichst einfach beschreibbar sein und vor allem sollte die Angabe der Adresse möglichst wenig Platz benötigen, d.h. es macht keinen Sinn eine Figur zu haben, die zwar das Zielgebiet exakt wiedergibt, aber bereits die ganze Paketlänge verbraucht.

5.4.2 2-D Figuren

Bisher erlaubt GeoCast nur die zweidimensionale Adressierung von Zielgebieten. Ein Polygon, dessen Eckpunkte durch Angabe der Länge und Breite auf dem WGS84 Ellipsoiden spezifiziert werden, ist die einzige Methode die bisher vorgesehen ist. Da für jede Koordinate nur 4 Byte verwendet werden, ist die maximale Genauigkeit bei 160 Metern. Dies ist für eine praktische Nutzung des Dienstes sicherlich nicht ausreichend. Durch eine Verwendung von 8 Byte pro Koordinate kann die Genauigkeit dann auf Teile eines Millimeters erhöht werden.

Doch neben der Genauigkeit stellt sich die Frage was für geometrische Figuren für die Adressierung vorgesehen werden sollten.

5.4.2.1 Kreis / Ellipse

Eine der am einfachsten spezifizierbaren Figuren ist der Kreis. Es wird nur der Mittelpunkt und der Radius benötigt, um das Ziel eindeutig zu beschreiben. Für Ellipsen ist noch ein weiterer Parameter notwendig.

Beliebige Flächen, insbesondere wenn sie eckig sind, lassen sich natürlich schlecht durch Kreise approximieren. So lassen sich z.B. Grundstücksgrenzen nicht einfach durch mehrere Kreise

annähern. Für diesen Fall ist der Kreis daher nur beschränkt geeignet. Häufig werden aber gerade Runde Zielgebiete verwendet werden, um z.B. eine Nachricht an alle zu senden, die in der Entfernung x um einen Punkt sind. Hierfür ist der Kreis logischerweise die ideale Figur. Der Kreis sollte daher als eine Methode der Adressierung vorgesehen werden.

5.4.2.2 Rechteck

Ein Rechteck kann entweder durch zwei Punkte und die Länge der Seiten oder durch drei Punkte definiert werden.

Ein Rechteck kann ein kreisförmiges Zielgebiet nur sehr ungenau annähern. Für die Beschreibung von eckigen Flächen ist das Rechteck eher geeignet. Durch die Kombination mehrerer Rechtecke lassen sich auch komplexe Flächen recht gut annähern. Da das Rechteck sich sehr platzsparend adressieren läßt, und vor allem da die notwendigen Enthaltenseinsberechnungen bei Rechtecken sehr effizient sind, sollte das Rechteck ins Adresskonzept aufgenommen werden.

5.4.2.3 Polygon

Das Polygon ist sicherlich die flexibelste zweidimensionale Figur. Mit ihr lassen sich beinahe alle anderen Figuren annähern. Daher kann mit einem Polygon jedes Zielgebiet angegeben werden. Es kommt nur auf die Anzahl der Eckpunkte an. Ein Polygon kann durch die Angabe aller Eckpunkte angegeben werden. Die Punkte werden im allgemeinen gegen den Uhrzeigersinn angegeben.

Das Polygon ist prinzipiell geeignet, jede beliebige Fläche (auch einen Kreis) mit hinreichender Genauigkeit anzunähern. Allerdings sollte die Anzahl der Seiten für GeoCast beschränkt sein, damit eine effiziente Bearbeitung möglich ist, und sich die Länge der Adresse in einem vernünftigen Rahmen hält.

5.4.3 3-D Figuren

Gerade im Indoor-Bereich sind zweidimensionale Adressen nicht mehr ausreichend. Zur Adressierung von Räumen oder Stockwerken ist es nötig, dreidimensionale Adressen zur Verfügung zu haben. Es gibt prinzipiell zwei Möglichkeiten, um solche Adressen darstellen zu können. Eine echte dreidimensionale Angabe des Zielgebietes, oder die eingeschränkte 2,5 dimensionale Adressierung.

5.4.3.1 Figuren mit gleichmäßiger Höhe (2,5 D)

Eine Möglichkeit, um dreidimensionale Zielräume zu definieren, ist eine Fläche mit einer Höhe zu versehen. Wenn die Punkte, die die Fläche definieren dreidimensional sind, die Fläche also auch eine Höhe über dem Boden hat, dann kann man damit Räume und Häuser adressieren. Diese Art Räume zu definieren wird oft auch 2,5 D genannt.

Für spezielle Figuren läßt sich das natürlich auch einfacher machen. So kann ein Quader auch durch die Angabe von 3 Punkten definiert werden.

Komplexere Räume (z.B. verwinkeltes Zimmer) können meist durch ein Polygon mit Höhe dargestellt werden (siehe Abbildung 5-4).

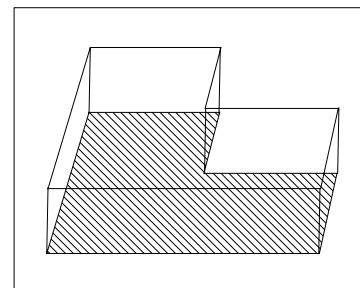


Abbildung 5-4. Polygon mit Höhe

Da die meisten Zielräume im dreidimensionalen Falle wohl Zimmer, Stockwerke oder Häuser sein werden, ist diese Art der Adressierung für GeoCast gut geeignet. Sie erlaubt eine kompakte und einfach zu verarbeitende Adresse.

Es sollten die Figuren unterstützt werden, die die Grundflächen haben, die in „2-D Figuren“ auf Seite 26 betrachtet wurden. D.h. der Zylinder, ein Quader und ein Polygon mit Höhe sind die Figuren die vorgesehen werden sollten.

5.4.3.2 Komplexe Räume

Um komplexere Räume, wie z.B. den Innenraum einer Kirche genau modellieren zu können, reicht die Angabe einer Höhe nicht aus. Das Dach der Kirche ist im Normalfall nicht gerade sondern formt eine Art Giebel. Hier müßte das Dach also genau modelliert werden.

Um beliebige 3D-Modelle definieren zu können, kann man zum einen alle Punkt und Kanten angeben, oder aber die Polygone, die die Wände modellieren. Beides ist sehr aufwendig und läßt sich auch nicht einfach als Adresse abbilden. Es ist hier wahrscheinlich empfehlenswerter, das Zielgebiet als Figur mit gleichmäßiger Höhe anzugeben, oder in mehrere solcher Figuren aufzuteilen.

Da diese Adressen zum einen sehr viel Platz benötigen und zum anderen die Berechnung von nötigen Operationen (z.B. Enthaltensein) erschweren sind sie für GeoCast ungeeignet. Daher sollte man auf solche Adressen verzichten und auf die 2,5 dimensional Adressen ausweichen.

5.5 Geschachtelte Adressen

Bisher haben wir betrachtet wie ein Zielgebiet in einem Weltmodell beschrieben werden kann. Es ist jedoch auch möglich, Adressen anzugeben deren Komponenten aus verschiedenen Teiladressen bestehen. So haben wir ja festgestellt, daß symbolische Identifier meist nur in einem begrenzten Gebiet genutzt und damit dort eindeutig sind. Um nun solch eine Adresse global eindeutig zu machen, kann man u.a. das Gebiet in dem sie gültig sind mitangeben. D.h. die Adresse hätte die Form „AREA.ID“. Das Gebiet („Area“) könnte dann z.B. in WGS84 Koordinaten angegeben werden. Dies ist nur ein Beispiel für geschachtelte Adressen. Prinzipiell sind beliebige

Kombinationen möglich. Es gibt allerdings einige Kombinationen die besonders viel Sinn machen.

1. Koordinatenbasiert.Symbolischer Identifier

Dies ist das Beispiel wie oben beschrieben. Eine Institution verwendet z.B. Active Batch und hat daher symbolische Identifier, die die Zimmer beschreiben. Der Zimmeradresse wird die Adresse des Firmengeländes, ausgedrückt in WGS84 Koordinaten, vorangestellt.

2. Symbolischer Identifier. Koordinatenbasiert

Wenn wir den Fall eines mobilen Objektes, z.B. eines Zuges betrachten, dann macht es keinen Sinn z.B. den Speisewagen durch WGS84 Koordinaten zu spezifizieren. Da der Zug sich bewegt, müßten die Koordinaten ständig geändert werden. Hier wird man ein lokales Koordinatensystem verwenden, das Positionen innerhalb des Zuges durch die Entfernung zu einem Bezugspunkt, z.B. dem Mittelpunkt des Zuges angibt. Der Zug selbst kann dann durch einen symbolischen Identifier spezifiziert werden.

3. Koordinatenbasiert.Symbolischer Identifier.Koordinatenbasiert

Die ersten zwei Teile der Adresse spezifizieren einen Raum. Wenn dieser Raum dann noch über ein Positionierungssystem verfügt, das die Position des Nutzers in lokalen Koordinaten angibt, dann kann man durch so eine Adresse Zielgebiete spezifizieren, die kleiner als ein Raum sind.

5.6 Scope und Architektur

Um Nachrichten effizient weiterleiten zu können, ist es unumgänglich einfache Adressen zu verwenden, bei denen effizient festgestellt werden kann, wohin die Nachricht weitergeleitet werden soll. Wenn man jedoch dreidimensionale Adressen verwendet, dann wird das einfach zu komplex. Die Verwendung dreidimensionaler Adressen erhöht den Aufwand, der für eine "Weiterleitungsentscheidung" getrieben werden muss, deutlich. Noch schwerwiegender ist jedoch das Problem, daß viele Gebäude nur mit Positonssensoren ausgestattet sind, die lokal gültige symbolische Identifier liefern. Entsprechende Klienten können also nichts mit Zielgebieten anfangen, die mit WGS84 Koordinaten definiert wurden. Da die Weiterleitung geographischer Nachrichten bisher nur auf WGS84 Koordinaten möglich ist, besteht für solche Klienten derzeit keine Möglichkeit an GeoCast zu partizipieren. Es ist zwar prinzipiell möglich ein Routing auch auf Identifiern durchzuführen, aber dieses würde bedeuten das zwei GeoCast-Routing Verfahren parallel nötig wären. Es wäre also gut, wenn man die Klienten auf andere Weise einbinden könnte.

Die Idee ist, daß jede Nachricht einen Scope und eine Zieladresse hat. Der Scope ist der Bereich in dem die Nachricht an alle Klienten verteilt wird. Die Klienten prüfen dann, ob sie im Zielgebiet sind, oder nicht. Als Scope wird das zweidimensionale Gebiet gewählt in dem das eigentliche Ziel komplett enthalten ist.

Im Falle von Koordinaten basierten Adressen kann z.B. das "bounding Rectangle" gewählt werden, also das kleinste Rechteck, daß die ursprüngliche Adresse komplett enthält. Im Fall einer Adresse, die auf symbolischen Identifiern beruht, muß der Scope zusammen mit der Adresse mitgeliefert werden. Es kann sich hier z.B. um die Grundfläche des Gebäudes handeln.

Wenn z.B. eine Nachricht in die oberen Stockwerke eines Gebäudes gesendet werden soll (siehe Abbildung 5-5), dann spezifiziert der Benutzer (oder die Anwendung) das dreidimensionale Zielgebiet. Das System bildet dieses dann auf den zweidimensionalen Scope ab, in diesem Fall das Rechteck, das die Grundfläche des Gebäudes enthält.

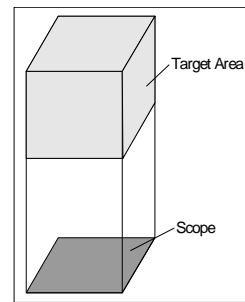


Abbildung 5-5.
Ziel und
Scope

Der Scope wird verwendet, um die Nachricht ins Zielgebiet zu routen. Die Nachricht wird dann an alle Klienten verteilt, die sich innerhalb des Scopes aufhalten. Auf diese Weise erhält jeder potentielle Empfänger die Nachricht. Bevor der Klient die Nachricht nun an die entsprechende Applikation weiterleitet, wird eine Filterung durchgeführt. Es wird überprüft, ob der Klient sich im tatsächlichen Zielgebiet aufhält. Die Nachricht wird also nur von Klienten in den oberen Stockwerken weitergeleitet. Die anderen verwerfen die Nachricht sofort. Wenn in dem Gebäude ein Positionierungssystem genutzt wird, das symbolische Identifier liefert, dann muß der Sender den Identifier für das Zielgebiet und den Scope liefern. Der Scope wäre auch in diesem Fall die Koordinaten des Gebäudes.

Die bereits existierenden Protokolle erwarten als Adresse ein Polygon, oder noch besser ein Rechteck. Durch die Verwendung des Rechteckes kann die Geschwindigkeit der Nachrichtenweiterleitung deutlich erhöht werden. Durch die Verwendung eines Rechteckes als Scope, ist eine einfache Integration mit den existierenden Protokollen möglich. Der Scope wird hierbei als Adresse für diese Protokolle verwendet. Die Klienten filtern die Nachrichten aber auf Basis der komplexen Adressen, die in einem inneren Header spezifiziert sind (see Abbildung 5-6).

Hierdurch werden z.B. Nachrichten die für einen Raum bestimmt sind, im ganzen Gebäude verteilt. Offensichtlich ist dies nicht sehr effizient. Durch eine kleine Änderung bei den GeoNodes kann man auch mit den existierenden Protokollen eine effizientere Verteilung realisieren. Die GeoNode muß nur überprüfen, ob ein Teil des Netzes, zu dem sie gehört, zum eigentlichen Zielgebiet gehört. Wenn ja, dann wird die Nachricht wie bisher gesendet, wenn nein, dann wird die Nachricht einfach verworfen. Auf diese Weise erhält die GeoNode immer noch Nachrichten die sie eigentlich nicht bräuchte, aber GeoNodes sind im Gegensatz zu Klienten normalerweise nicht über Funknetze angebunden. Und Bandbreite ist vor allem bei Funknetzen noch ein limitierender Faktor. Bei den kabelgebundenen Netzen fällt dies wesentlich weniger ins Gewicht.

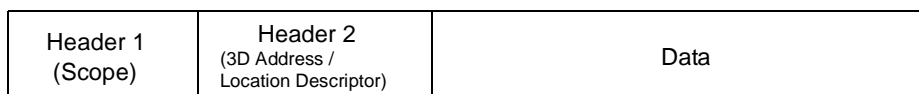


Abbildung 5-6. GeoCast Nachricht

5.7 Nachrichtenfilterung

Der Klient nimmt bei jeder GeoCast-Nachricht eine Filterung vor. Er prüft, ob seine Position innerhalb des angegebenen Zielgebietes liegt. Hierzu muß der Klient natürlich in der Lage sein, seine eigene Position mit dem Ziel zu vergleichen. Im Falle, daß beide das gleiche Modell verwenden ist dies kein Problem. Schwieriger wird es, wenn der Klient seine Position in einem anderen Modell zur Verfügung hat, als in der Adresse der Nachricht gegeben ist.

Betrachten wir zuerst den Fall, daß beide Formate gleich sind, die homogene Filterung und danach den anderen Fall, die heterogene Filterung.

5.7.1 Homogene Filterung

In den meisten Fällen wird der Klient die Adresse direkt vergleichen können. Im Outdoor-Bereich wird praktisch immer GPS bzw. DGPS und damit WGS84 eingesetzt. Im Indoorbereich funktioniert genau dieses System nicht. Hier funktionieren nur Systeme, die eine feste Infrastruktur im Gebäude haben. Es ist nicht anzunehmen, daß in einem Gebäudeteil mehrere unterschiedliche Positionierungssysteme installiert sind, und genutzt werden. Von daher kann der Sender durch Wahl des entsprechenden Modells alle Klienten erreichen. Die Frage, die sich stellt, ist woher der Sender das Wissen über die verwendeten Modelle hat.

Im Gegensatz zu "Outdoor"-GeoCast, wo der Sender keine Kenntnis über das Ziel haben muß, wird bei "Indoor"-GeoCast eine genau Kenntnis des Ziels vorausgesetzt. Der Sender muß entweder die verwendeten symbolischen Identifier angeben, oder etwa die geographischen Koordinaten der Räume kennen die er adressieren will. Diese Informationen müssen dem Sender entweder explizit bekannt sein, oder er nutzt einen entsprechenden Dienst, der entweder Namen oder gar kartenbasiert eine Auflösung auf die entsprechende Zieladresse erlaubt.

In jedem Fall kann dabei das Format gewählt werden, das die Klienten auch verstehen.

5.7.2 Heterogene Filterung

Wenn homogene Filterung der Normalfall ist, wann tritt dann überhaupt noch die heterogene Filterung auf? Nehmen wir einen Klienten, der sich innerhalb eines Gebäudes aufhält, und der mit symbolischen Identifiern arbeitet. Jetzt wird eine Nachricht an ein großes zweidimensionales Gebiet geschickt. Die Adresse wird in geographischen Koordinaten gegeben. Der Klient ist jedoch nicht in der Lage festzustellen, ob er sich in diesem Gebiet befindet oder nicht.

Ich schlage daher vor, daß für Klienten, die symbolische Identifier verwenden, auf jedem Klienten auch bekannt ist, in welchem Gebiet diese Identifier gültig sind. Auf diese Weise, kann der Klient zwar nicht feststellen ob er selber im Zielgebiet ist, aber kann feststellen, ob sich das Gebäude, in dem er ist im Zielgebiet befindet. Selbstverständlich kann der Fall auftreten, in dem ein Gebäude nur teilweise im Zielgebiet liegt. Dieses wird jedoch recht selten der Fall sein, weil es wohl kaum Nachrichten gibt, die für ein größeres Gebiet interessant sind, jedoch in einem Gebäude dann nur in einem Teil.

Daher kann in einem solchen Fall die Nachricht einfach angenommen werden, wenn ein Teil des Zieles sich mit den Koordinaten überdeckt.

Wenn mehr Anpassungsaufwand in das System gesteckt werden soll, dann kann man natürlich auch jedem Klienten bekannt machen, welches geographische Gebiet mit welcher Id übereinstimmt. Dann kann das oben genannte statt auf das ganze Gebäude, nun auf Raumbene gemacht werden. Dadurch erhöht sich einerseits die Genauigkeit, allerdings stellt die Bereitstellung dieser Information auch einen nicht zu unterschätzenden Aufwand dar.

5.8 Aliase

Dem Nutzer eines Systems kann es sicherlich nicht zugemutet werden, alle Koordinaten der möglichen Zielgebiete zu kennen, bzw. sich zu merken. Insbesondere, da solche Koordinaten keine erkennbare Bedeutung für den Nutzer haben. So wird sich der Anwender nur ungern mit

geographischen Koordinaten beschäftigen. Er wird die Koordinaten entweder über eine Karte o.ä. eingeben, oder aber einen Namen oder Alias angeben. Ich möchte hierbei darauf hinweisen, daß dies keine Funktionalität für GeoCast ist, sondern von der Anwendung darüber realisiert wird. Für das GeoCast System werden die Adressen aus den Komponenten bestehen, die bisher erläutert wurden. Der Vollständigkeit halber wird hier trotzdem noch ein Blick auf die Aliase geworfen.

5.8.1 Lokale Aliase

Lokale Aliase sind nur auf dem System des Nutzers bekannt. Er kann hierbei sowohl Punkten als auch gleich ganzen Zielgebieten einen Namen zuweisen. Ein Alias ist ein leichter lesbarer Ersatz für eine Koordinate oder ein Zielgebiet. Bevor die Nachricht dann an das eigentliche GeoCast weitergereicht wird, werden die Aliase durch die entsprechenden geographischen Koordinaten ersetzt.

5.8.2 Globale Aliase

Im Internet existiert ein ähnliches Problem. Kaum einer wird sich IP-Adressen merken. Die Zahlenkombination 208.202.218.15 läßt sich eben wesentlich schwerer merken als der dazugehörige Alias „www.amazon.com“. Ein Dienst im Netz, der Domain Name Server (DNS), sorgt für die Auflösung des Aliases auf die eigentliche Adresse.

Ein analoges Vorgehen ist auch für geographische Adressen denkbar. Anstatt die Koordinaten des Schloßplatzes in Stuttgart anzugeben, teilt der Nutzer dem System den Namen „Schloßplatz.Stuttgart.BW.ger“ mit. Ein, dem DNS analoger Dienst würde dann die entsprechenden Koordinaten zurückliefern. Da der Geo-DNS die Auflösung Name → Adresse nur auf Basis einer entsprechenden Tabelle vornehmen würde, und ihm keine Informationen über geographische Zusammenhänge bekannt wären, müßte der Name genau so angegeben werden, wie er dem Geo-DNS bekannt wäre. Analog dazu erhält man vom DNS auch keine (korrekte) IP-Adresse, wenn statt der oben genannten Adresse einfach www.amazon.com angegeben wird. Die Namen müssen so gewählt sein, daß sie global eindeutig sind, also kein zweites Mal vorkommen.

5.8.3 Namensschema für globale Aliase

Eine Frage, die zu klären wäre, ist wie ein allgemeines, überall gültiges Adreßschema für globale Aliase aussehen könnte. Eine Idee ist es, analog zur Postadresse vorzugehen. Ein solches Adreßschema könnte wie folgt aussehen:

zimmernr.stock.hausnr.strasse.stadt.land.staat.geo

18.2.20-22.breitwiesenstrasse.70565.b-w.germany.geo

5.8.4 Verwendung von Aliasen

Aliase können auf verschiedene Arten zur Spezifikation eines Zielgebietes herangezogen werden.

5.8.4.1 Alias bezeichnet Punkt

Ein Alias kann auf einen Punkt verweisen. Dies könnte z.B. die aktuelle Position eines mobilen Nutzers oder aber einfach eine häufig genutzte Koordinate (z.B. Stadtzentrum) sein. Ein solcher Alias kann wie eine geographische Koordinate zur Bildung von Adressen herangezogen werden.

5.8.4.2 Alias/Objekt bezeichnet Zielgebiet

Ein Alias kann auch das Zielgebiet bezeichnen. In diesem Fall reicht die Angabe des Namens um das Zielgebiet zu definieren. Ein Beispiel hierfür wären Häuser oder Zimmer. Durch die Angabe, daß eine Nachricht in ein bestimmtes Zimmer gesendet werden soll, ist die Adresse definiert. Weitere Angaben sind nicht nötig. Ein solches Zielgebiet kann beliebig komplex sein.

5.8.4.3 Zielgebiet wird als Punkt verwendet

Wenn ein Alias ein Zielgebiet definiert (z.B. Rathaus Esslingen) so kann der Anwender evtl. diesen Alias als Koordinaten für ein größeres Zielgebiet verwenden. Ein Beispiel ist in Abbildung 5-7 zu sehen. Hier wurden die Aliase für den Hauptbahnhof, das neue Schloß und das Rathaus verwendet um ein größeres Zielgebiet anzugeben. In diesem Fall werden die Objekte als Koordinaten für eine geometrische Figur verwendet. Da die Objekte aber keine Punkte sind, muß zur Berechnung der Figur jeweils ein Punkt (wenn möglich der Mittelpunkt) aus dem Zielgebiet herangezogen werden. Da auf diese Weise ein Teil der Leute im Objekt die Nachricht nicht erhalten würden, empfiehlt es sich auch noch das Objekt selber hinzuzunehmen.

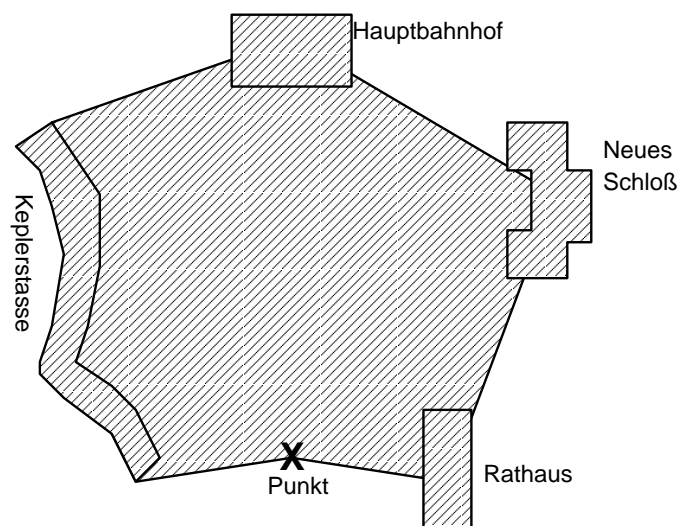


Abbildung 5-7. Polygon (Hauptbahnhof, Neues_Schloß, Rathaus, Punkt (L49.545,B9.841),Keplerstr)

5.9 Mobile Ziele

Mobile Ziele stellen besondere Anforderungen an die Adressierung. Ein mobiles Ziel kann sicher nicht mit Koordinaten eingegeben werden, sondern die Adresse muß zumindest auf einem mobilen Objekt (Auto, Person, Zug, etc.) beruhen. Hier stellt sich dann die Frage wie diese mobilen Objekte in ein allgemein gültiges Adreßschema passen könnten. Eine Möglichkeit wäre es einen eigenen „Zweig“ im Namensraum für mobile Objekte zu bilden.

Eine weitere Fragestellung ist, wie der Name auf die aktuelle Position aufgelöst werden kann. Wenn sich das Objekt dann auch noch mit großer Geschwindigkeit bewegt, muß nicht auf die derzeitige Position, sondern idealerweise auf die zukünftige Position aufgelöst werden. Eine Idee wäre es eine Art Home Agent zu haben, der die aktuelle Position kennt und die Nachricht entsprechend weiterleitet.

5.10 Verknüpfungen

Um auch komplexe Zielgebiete noch einfach spezifizieren zu können, sollten Verknüpfungen möglich sein. Einige Beispiele für solche Verknüpfungen sind in Abbildung 3-5 zu sehen.

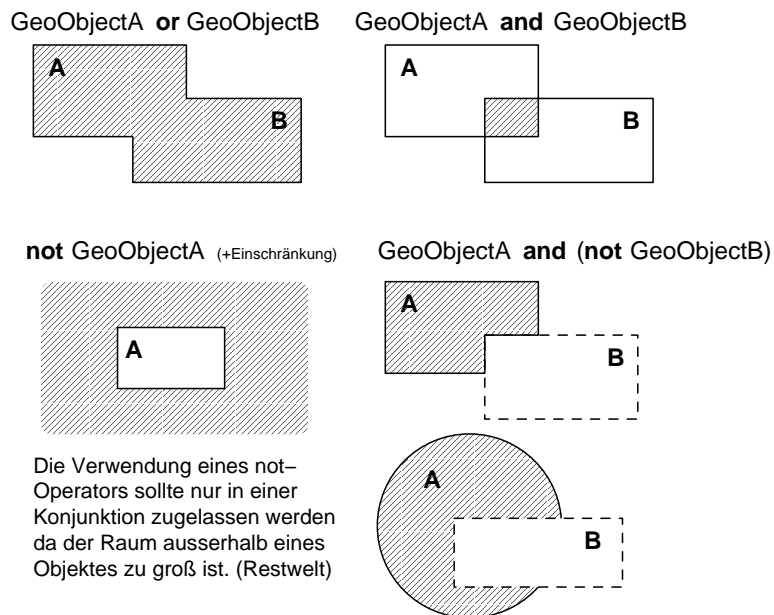


Abbildung 5-8. Logische Verknüpfungen von Zielen

6 Ausblick

Wir haben in dieser Arbeit die Bereiche Adressierung und Semantik eines GeoCastes intensiv betrachtet. Es wurde Wert darauf gelegt, daß Spektrum möglicher Alternativen auszuloten und die geeignetsten Lösungen hervorzuheben.

Für einen Erfolg von GeoCast ist es unabdingbar das es eine klare und sinnvolle Semantik gibt. Der Anwender muß verstehen können was passiert und wer seine Nachricht erhalten wird. Darüberhinaus ist natürlich eine klares und mächtiges Adressierungskonzept unabdingbar. Es muß mächtig genug sein, damit der Anwender die Zielgebiete formulieren kann, die für ihn sinnvoll erscheinen. Auf der anderen Seite steht eine effiziente Verarbeitbarkeit der Adressen. D.h. Die Adressen müssen eine effiziente Weiterleitung der Nachrichten erlauben.

Mit diesen beiden Problembereichen haben wir uns in dieser Arbeit intensiv beschäftigt, und versucht sinnvolle Lösungen aufzuzeigen. Wir denken das GeoCast an Bedeutung gewinnen wird, und in naher Zukunft so selbstverständlich genutzt wird, wie heute Uni- und Multicast.

Weitere Probleme die noch gelöst werden müssen sind im Bereich der Nachrichtenweiterleitung zu finden. Die vorgeschlagenen Protokolle sind zwar durchaus brauchbar solange die Gruppe der Nutzer klein bleibt. Wenn GeoCast aber wirklich erfolgreich wird, und dementsprechend viele Nutzer vorhanden sein werden, dann werden die bisherigen Protokolle schnell an Ihre Grenzen geraten. Insbesondere die Skalierbarkeit ist ein Problem das dringend eine Lösung sucht.

Ein weiteres Problem ist die Existenz von geeigneten Applikationen die es dem Nutzer erlauben die Vorteile von GeoCast zu nutzen. Wir haben eine Anwendung entwickelt, die das komfortable Versenden von geographischen Nachrichten erlaubt. Weitere Anwendungen werden sicherlich folgen.

7 Literatur

- [1] <http://www.cs.ukc.ac.uk/research/infosys/mobicomp/Fieldwork/Sticke/index.html>
- [2] <http://www.comp.lancs.ac.uk/computing/research/mpg/most/guide.html>
- [3] <http://nexus.informatik.uni-stuttgart.de>
- [4] F. Hohl, U. Kubach, A. Leonhardi, K. Rothermel, M. Schwehm: Next Century Challenges: Nexus - An Open Global Infrastructure for Spatial-Aware Applications, Proceedings of the Fifth Annual ACM/IEEE International Conference on Mobile Computing and Networking (MobiCom'99), Seattle, Washington, USA, August 15-20, 1999, T. Imielinski, M. Steenstrup, (Eds.), ACM Press, 1999, pp. 249-255
- [5] <http://athos.rutgers.edu/dataman/>
- [6] T. Imielinski and J. C. Navas, "GPS-Based Addressing and Routing" IETF RFC 2009, November 1996
- [7] T. Imielinski, J. C. Navas, "GPS-Based Geographic Addressing, Routing and Resource Discovery", *Communications of the ACM Vol. 42*, April 1999
- [8] J. C. Navas and T. Imielinski, "Multi-hop Dynamic Geographic Routing", Technical Report (LCSR-TR-364), Rutgers University Computer Science, May 29, 1998
- [9] D. Estrin, D. Farinacci, A. Helmy, D. Thaler, S. Deering, M. Handley, V. Jacobson, C. Liu, P. Sharma and L. Wei, "Protocol Independent Multicast-Sparse Mode (PIM-SM): Protocol Specification", IETF RFC 2362, Juni 1998
- [10] Young-Bae Ko and Nitin H. Vaidya, "Location-Aided Routing (LAR) in Mobile Ad Hoc Networks", Proceedings of Mobicom 98, Dallas, Texas, USA
- [11] Young-Bae Ko and Nitin H. Vaidya, "Geocasting in Mobile Ad Hoc Networks: Location-Based Multicast Algorithms", Proceedings of the 2nd Workshop on Mobile Computing Systems and Applications (WMCSA '99)

