

# **Vorabübertragung ortsbezogener Informationen zur Unterstützung mobiler Systeme**

Von der Fakultät Informatik der Universität Stuttgart  
zur Erlangung der Würde eines Doktors der  
Naturwissenschaften (Dr. rer. nat.) genehmigte Abhandlung

Vorgelegt von

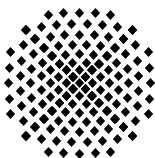
**Uwe Kubach**

aus Neckarsulm

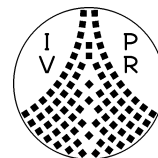
Hauptberichter: Prof. Dr. rer. nat. Dr. h. c. Kurt Rothermel

Mitberichter: Prof. Dr.-Ing. Dr. h. c. mult. Paul J. Kühn

Tag der mündlichen Prüfung: 7. Februar 2002



Institut für Parallele und Verteilte  
Höchstleistungsrechner (IPVR)  
der Universität Stuttgart  
2002





## Danksagung

An dieser Stelle möchte ich mich bei meinem Doktorvater, Prof. Dr. rer. nat. Dr. h. c. Kurt Rothermel, herzlich für seine tatkräftige Unterstützung bedanken. Er hat mit zahlreichen Anregungen und konstruktiver Kritik zum Gelingen dieser Arbeit beigetragen.

Ebenso gilt mein Dank Herrn Prof. Dr.-Ing. Dr. h. c. mult. Paul J. Kühn für seine Bereitschaft, den Mitbericht zu übernehmen sowie seine anregenden Fragen bei den Berichtskolloquien des Graduiertenkollegs Parallele und Verteilte Systeme.

Meinen Kollegen aus der Abteilung Verteilte Systeme bin ich für das angenehme Arbeitsklima und die gute Zusammenarbeit auf die gleiche Weise zu Dank verpflichtet.

Der Deutschen Forschungsgemeinschaft (DFG) danke ich für ihr großzügiges Stipendium, das diese Arbeit überhaupt erst ermöglicht hat.

Nicht zuletzt bedanke ich mich bei meinen Eltern und meiner Lebensgefährtin für ihre Unterstützung und ihr Verständnis.

Uwe Kubach

Stuttgart, den 18. 10. 2001



## Kurzfassung

Mit der zunehmenden Verbreitung mobiler Endgeräte wächst auch der Bedarf in mobilen Umgebungen auf entfernt gespeicherte Informationen zuzugreifen. Wie aktuelle Studien belegen, kommt dabei zukünftig ortsbezogenen Diensten (Location-Based Services, LBS) eine besondere Bedeutung zu. Diese Dienste ermöglichen es dem Benutzer, abhängig von seinem gegenwärtigen Aufenthaltsort, auf lokale Informationen, beispielsweise Informationen zu Restaurants, Kinoprogramme oder Staumeldungen, zuzugreifen.

Mobile Informationszugriffe erfolgen meist über drahtlose Netzwerke, in denen häufig Probleme wie geringe Bandbreite, große Verzögerungen oder Verbindungsausfälle auftreten. In dieser Arbeit wird ein Verfahren vorgeschlagen, das diese Probleme weitgehend löst, indem es vorhersagt, welche Informationen ein Benutzer voraussichtlich benötigen wird und sie im Voraus auf sein mobiles Endgerät überträgt. Greift der Benutzer dann tatsächlich auf die vorab übertragenen Informationsobjekte zu, sind sie bereits lokal auf dem mobilen Endgerät vorhanden und stehen somit unmittelbar zur Verfügung.

Für eine möglichst präzise Vorhersage der benötigten Informationen nutzt das beschriebene Verfahren Beziehungen zwischen Informationen und Orten, wie sie in mobilen Informationssystemen, insbesondere in ortsbezogenen Systemen, häufig zu finden sind, aus. Sie erlauben es, aus Kenntnissen über den aktuellen und zukünftigen Aufenthaltsort eines Benutzers Rückschlüsse auf die benötigten Informationen zu ziehen. Die Kenntnisse über den Aufenthaltsort erhält das Verfahren durch eigene Beobachtung oder wahlweise aus externen Informationsquellen. Ferner werden die Beziehungen zwischen Informationen und Orten selbständig erkannt. Das Verfahren reagiert sogar dynamisch auf Veränderungen in diesen Beziehungen.

Neben der ausführlichen Erläuterung des Verfahrens wird in dieser Arbeit auch seine Implementierung beschrieben. Für die Evaluierung des Verfahrens wurde ein Simulationsmodell für mobile, ortsabhängige Informationszugriffe erstellt und ein entsprechendes Simulationstool entwickelt. Das Modell sowie das Werkzeug werden ebenfalls vorgestellt.

Die Evaluierung des Verfahrens zeigt schließlich, dass es sich sehr gut für die Unterstützung mobiler, ortsabhängiger Systeme eignet. Des Weiteren wird nachgewiesen, dass die Möglichkeit, externe Informationsquellen zu nutzen, sehr vorteilhaft ist.



# Hoarding of Location-Dependent Information to Support Mobile Systems

## English Abstract of the Dissertation

Recent market projections indicate that the need for mobile information access, especially in location-based services, will be the driving force for the future development of the mobile communication market as well as the market for mobile devices, such as Notebooks, Personal Digital Assistants (PDAs), Smart Phones, or Web-Pads [Price-waterhouse-Coopers 2001, Heise-Newsticker 2001, Leite & Pereira 2001].

All major mobile network carriers in Germany have recognized the significance of location-based services and offer such services through their WAP<sup>1</sup> portals. In the research area of mobile computing, location-based systems also play an important role and thus many location-dependent information systems have been developed recently. The mobile information access always turned out to be one of the basic research challenges in such systems, since the deployed wireless communication technologies often led to problems like high delays, low bandwidths, or frequent disconnections.

In this thesis, a mechanism is proposed that overcomes these problems as far as possible. It allows a mobile information access independent of the availability of a mobile communication network. The users can even access information when no network is available at all (disconnected operation). The mechanism was especially designed for location-dependent information systems. In such systems, the probability a certain information object is accessed with depends on the user's location. This is, for example, the case in location-based services. In contrast to existing WAP-Services, the mechanism developed in this thesis supports access to much more extensive information objects like multimedia documents.

The basic idea of this mechanism is to predict the information objects that a user will probably need and to transfer these objects to his mobile device before they are actually accessed. This anticipatory information transfer is called information hoarding. If the user accesses one of the prefetched objects, it is already stored locally on his device and is available immediately.

---

<sup>1</sup>Wireless Application Protocol

The problem with hoarding is to predict what information a user will need (“hoarding problem”). However, a good prediction is important to avoid that information is transferred to the user’s device that he will never request. Otherwise, scarce resources, such as the time required for the transmission or the memory on the user’s device, are wasted. The mechanism proposed in this work uses relationships between information objects and locations to solve the hoarding problem. It uses these relationships to derive the information objects the user will probably need from knowledge about his current and future location [Kubach & Rothermel 2001b].

In contrast to existing approaches, these relationships neither have to be fixed nor do they have to be well known, but the proposed mechanism recognizes them on its own and can even adapt dynamically to changes in these relationships. The knowledge about a user’s future location is either gained autonomously by the mechanism through observation of the user or it is specified by external information sources like the application or the user himself. This makes the proposed mechanism much more flexible than other approaches, which usually use either an internal prediction mechanism or external knowledge, but not both.

This thesis consists of eight chapters. In Chapter 1, a motivation for this work is given and its structure is explained. Chapter 2 consists of a discussion of other mobile data management mechanisms and the related work. Afterwards, some basics, necessary to understand the subsequent chapters, are described in Chapter 3. In Chapter 4, the mechanism is described in detail. An overview of the implementation is given in Chapter 5. Before evaluating the mechanism in Chapter 7, a description of the simulation tool used for the evaluation is provided in Chapter 6. Finally, the thesis is concluded in Chapter 8. In the following sections of this abstract, a summary of the Chapters 2-7 is given.

## **Mobile Data Management**

In the first part of this chapter, an overview of mobile data management mechanisms is given. Basically, each of these mechanisms can be used to support mobile information accesses, however they differ in their suitability for location-dependent systems.

Replication systems aim to increase the availability of data and to reduce a system’s response time by maintaining several copies of each data item. The main questions are



where to store the copies and how to keep them consistent. In contrast to the replication of data, consistency is usually not considered in hoarding mechanisms. Sometimes hoarding systems are built on top of a replication system to ensure consistency. A further difference is that in a hoarding system the transfer of data is triggered by the client and the data is selected according to the individual needs of the user, whereas in replication systems the transfer of data is triggered by the system and the decision about replicating data is based on global criteria.

Caching is a very common method to speed up information accesses in both wired and wireless networks. The idea is to store information, once it is transferred from the server, locally on the user's device. Thus, the information is already available on the device, in case it is accessed again. If no more memory is available to locally store further information items, a caching strategy has to decide which information items should be removed from the cache. Most of these caching strategies are based on the assumption that there is a temporal locality in a user's request pattern. However, this might not be true in mobile environments, especially when location-dependent information is accessed. For example, when a user moves from one location to another, he will probably not be interested in information concerning the previous location anymore. Therefore, new caching strategies have been developed, which rely on a geographical locality. However, independent of the replacement strategy, caching never speeds up the first access to an information item. Thus the hit ratios achieved in location-dependent information systems are low, if the users do not frequently return to previously visited locations.

A further mobile data management mechanism is known as asynchronous information access. Asynchronous means that an information request that occurs while no or only a low bandwidth connection is available is delayed until a high bandwidth network connection becomes available. The main problem with this approach is that the users might then not be interested in the requested information anymore, since, in the meantime, they have moved on to another location.

Finally, broadcast data dissemination mechanisms are introduced. Their primary focus is to reduce the response time and to increase the scalability of the system. If they are location-aware, they are designed to support the users with the information items they need at their current location, e.g. within the coverage area of one cell of the dissemination system. Therefore, location-aware, broadcast-based dissemination mechanisms

do not make any predictions on the information items the users will need after leaving for another location.

The main result of the consideration of mobile data management mechanisms is that hoarding is better suited for location-dependent systems than the other mechanisms mentioned above.

In the second part of the chapter, the related work, i.e. mechanisms that transfer information in advance, is discussed. First of all, a taxonomy is proposed to categorize these mechanisms. It distinguishes between prefetching mechanisms, where a continuous network connection is assumed, and hoarding mechanisms, which are primarily designed to support disconnected operations. Further dimensions of the taxonomy are the type of the supported system, whether the decision about the objects to transfer is made by the system or interactively, and the criteria used for this decision.

Afterwards the most important prefetching and hoarding systems are described in detail, before they are eventually compared to the mechanism proposed in this thesis.

## **Basics**

In this chapter, the basics necessary to understand the subsequent elaborations are explained. In the first section, measures to describe concentrations and disparities are introduced, namely the concentration rate, the Herfindahl coefficient, and the Gini coefficient. They are used to define various metrics for the evaluation of the mechanism in Chapter 7.

Afterwards, an overview of current and future wireless communication technologies is given. It turns out that there are basically two types of such technologies: the technologies used in mobile telecommunication systems (GSM, HSCSD, EDGE, UMTS), which have a large coverage area, but support only low data rates, and the wireless local area technologies (WLANs, HiperLAN/2, DECT, IrDA), which usually support higher data rates, but have only a small range. The qualitative differences between these two classes will remain in the future, since wireless local area technology will always be faster than wireless wide area technology.

Then the system model on which this thesis is based is described. The description starts with the assumptions made for the supported information systems. The most

important assumed characteristic of the information systems is their location dependency.

Next the assumed communication infrastructure, a so-called infostation infrastructure, is introduced. This infrastructure combines the advantages of wireless local area technologies (high data rates) with those of wireless wide area networks (large coverage area). Infostations are islands of wireless LANs, which are embedded in an area that is otherwise not covered by a wireless network at all or only by a wireless wide area network.

The section concludes with a description of the requirements for the mobile devices and some required external services, namely a location service, an event service, and a directory service.

## **Hoarding of Location-Dependent Information**

This chapter constitutes the main part of the thesis. Here, the proposed hoarding mechanism is described in detail. Actually, two different variants of the mechanism, a basic variant and an enhanced one are described.

In the first section of the chapter, the hoarding problem is explained extensively. Furthermore, the hit ratio is introduced as a measure to rate the performance of a hoarding mechanism.

Thereafter, an overview of the basic functionality of the hoarding mechanism is given. A proxy server that coordinates the hoarding processes is associated with each infostation. In this thesis, the term infostation is used for the entirety of the wireless LAN and the associated proxy. The infostations serve geographically limited areas, their so-called hoarding areas. To each infostation belongs exactly one hoarding area and each hoarding area is assigned to exactly one infostation. In reasonable configurations, the hoarding area of an infostation should be bigger than the area covered by the wireless LAN.

The infostations maintain knowledge about what the information objects are that the users preferably access while moving through their hoarding area. When a user arrives at an infostation, this knowledge is used to predict the information objects that he will most probably request during his subsequent trip through the hoarding area of the

infostation. Based on this prediction the infostations determine the information objects to be hoarded on the user's device. The hoarding areas of adjacent infostations should overlap each other, so that, at each infostation, the users can be supplied with all the information that they will probably need until they reach the next infostation.

The hoarding process for each user is performed in a cyclic manner. The cycle consists of the following three phases:

1. *Download*: During the download period the user is located at an infostation, i.e. his mobile device is connected to a proxy server through a wireless LAN. Based on its knowledge about what the preferred information objects are in its hoarding area, the proxy determines the information objects that the user will most probably access before he reaches the next infostation and transfers them to the user's device. The amount of data that is transferred to the device is either limited by the time the user is staying within the area covered by the wireless LAN or by the memory available on the device.
2. *Disconnected Operation*: The second phase of the cycle begins after the user has left the infostation. It lasts until he reaches the next infostation. During this period of disconnected operation, the user's information requests are answered with the information that has been hoarded at the previous infostation. If the user requests an information object that has not been hoarded, a hoard miss occurs. Then the request has to be handled over the wireless WAN, if available and if the user accepts to pay for the WAN connection. Otherwise, it cannot be answered. Independent of whether or not a request succeeded, it is logged to a file together with the geographic position where the request occurred. The geographic position is determined by the external location service.
3. *Upload*: Finally, the third phase starts as the user reaches the next infostation. Then the information about the user's requests that has been logged on his device is uploaded to the proxy, which distributes this information to the other proxies. This distribution is done by sending the log file to each infostation, including the distributing infostation itself in whose hoarding area at least one of the logged information requests occurred. The information in the log file is used by the proxies to update their knowledge about the preferred information objects within their hoarding areas. Finally, the log file on the mobile device is deleted and the cycle starts again with the subsequent download process.

In the basic variant of the mechanism, the knowledge that each infostation uses for the hoarding decision simply consists of a table that contains the access probabilities of those information objects, which are the most popular ones within the infostation's hoarding area. This table is called the access probability table (APT) of an infostation. It is updated periodically with the information gained from the log files recorded on the users' devices.

With the APT the hoarding problem can easily be solved. If a maximum of  $m$  information objects can be transferred to the device, always those  $m$  objects will be transferred that have the highest access probabilities according to the APT.

The disadvantage of the basic variant is that the location dependency is only considered at the granularity of hoarding areas. Thus, knowledge about the user's future location, which often is easily available, can not be used for the hoarding decision. For example, a navigation system usually knows the exact path on which its user travels. If the navigation system could communicate this knowledge to the infostation and the infostation would know which information objects are preferred in the crossed subareas of the hoarding area, the hoarding decision could be much more precise.

The goal of the enhanced variant of the proposed mechanism, is to use such hints about the user's future location, in order to improve the prediction about the information objects he will probably need. Since this kind of information is not gained by the hoarding mechanism itself, it is called external knowledge. External knowledge might be offered by the applications or by the users themselves.

The basic idea of the enhanced variant is to divide each hoarding area into separate zones, which do not overlap, and to observe separately for each of these zones, which are the preferred information objects within the zone. Thus, for each of these zones, a separate APT is maintained. The geometry of the zones might be quite simple, e.g. equally sized squares, or it might reflect the geometry of buildings, streets, or other real world objects. In addition to the APTs, information about the visit probabilities of the zones is also maintained and stored in so-called visit probability maps.

External knowledge is specified in external visit probability maps. For example, the navigation application mentioned above can describe a user's future path by specifying a visit probability of one for the zones crossed by the path and a probability of zero for the remaining zones. In addition, it is also possible to specify relative external

visit probability maps, for example, to specify that certain zones are more likely to be visited than others.

The mechanism finally integrates all available external maps and the infostation's self-maintained map into one final visit probability map. Thereby, a configurable parameter determines how far the mechanism should trust the external sources. If it is set to one, the mechanism completely relies on the external maps, if it is set to zero the external knowledge is not considered at all.

From the final visit probability map and the APTs of all zones, each infostation can calculate the probability with which a certain user will request each information object within its hoarding area. In contrast to the basic variant, these probabilities now reflect all the available external knowledge about the user's future movement.

The external visit probability maps are only a low level interface to the proposed mechanism. Application dependent higher level interfaces can easily be implemented. Such an interface could, for example, allow a navigation application to specify a user's path through the coordinates of the visited locations. From these coordinates the external map could be calculated automatically, so that the user or application does not have to know anything about zones.

A discussion of common characteristics of both variants, like adaptivity, scalability, and privacy issues, follows the description of the mechanism. Furthermore, the role of important parameters, i.e. the size of hoarding areas and zones, as well as communication costs, storage costs, and computational costs for both variants are analyzed.

Finally, two further enhancements to the mechanism are described. The first one is a channel concept, that allows to consider a user's profile when selecting the information objects to hoard. The second enhancement is a filter mechanism, that allows a more user-specific hoarding decision.

## **Implementation**

The implementation of both variants of the mechanism is subject of Chapter 5 of this thesis. An important objective of the implementation was that it can be easily integrated into the neXus infrastructure [Hohl et al. 1999]. NeXus is an infrastructure for location-aware applications, which is currently developed at the Institute of Parallel and Distributed High-Performance Systems (IPVR).

The implemented hoarding system consists of infostation servers and hoarding clients. To operate an infostation, exactly one infostation server is necessary. The hoarding client must be installed on each mobile device that should be able to download information at an infostation.

In addition to these two components, basic versions of the required external services have also been implemented. By using the adapter design pattern, it was possible to ensure that these external services can be replaced without changing the hoarding system. Thus, integration into another infrastructure, like neXus, is very easy.

Finally, some application examples for the hoarding system are given in the last section of this chapter.

## **Simulation Environment**

Although a number of mobility models and information access models do exist, it was necessary to design a new simulation model for the evaluation of the hoarding mechanism. In Chapter 6 of the thesis, this model and the simulation tool that was implemented to realize the model, are described.

A review of existing models showed that the problem with them is, that they only model either the mobility or the information accesses of users, but not both. So they can not be used to model relationships between the mobility and the information accesses and thus questions like “where is a certain information object preferably accessed?” or “what information objects are accessed at a certain location?” can not be answered with these models.

The model designed in this thesis overcomes these disadvantages by integrating existing mobility models and information access models into one global model. In addition, it makes it possible to model the relationships between information objects and locations explicitly. The model is based on a graph representation of the real world, which is called the infrastructure graph, and consists of three submodels: the visit model, the information model, and the access model.

The visit model states the probability with which each location in the infrastructure graph is visited. The information model reflects the relationships between the information objects and the locations and the access model describes the users’ access behavior.

Finally, the simulation algorithm and the implementation of the simulation tool are also described in detail.

## **Evaluation**

In the first section of this chapter the characteristics of an information system that influence the efficiency of the hoarding mechanism and the advantage of the enhanced variant over the basic variant are identified. To quantify these characteristics new metrics are introduced using the concentration measures described in Chapter 3.

After a brief discussion of the expected influence of these characteristics on the hoarding mechanism, a mathematical analysis of the expected hit ratios is carried out.

At the end of the chapter, the extensive simulations which have been conducted with the simulation tool introduced in the previous chapter are described. The results show that the use of information about a user's current and future location makes the proposed mechanism well suited for location-dependent mobile information systems. Another important result is that the exploitation of external knowledge is highly beneficial in most cases.

The default values for the simulation parameters were chosen on the basis of an analysis of an existing mobile tourist guide. For this analysis log files recorded in field trials of the GUIDE project [Cheverst, Davies, Mitchell & Friday 2000] have been examined. A description of this analysis is also part of this chapter.



# Inhaltsverzeichnis

<b>1</b>	<b>Einleitung</b>	<b>1</b>
1.1	Motivation . . . . .	1
1.2	Aufbau der Arbeit . . . . .	3
<b>2</b>	<b>Mobile Datenverwaltung</b>	<b>5</b>
2.1	Einordnung der Arbeit . . . . .	5
2.1.1	Replikation . . . . .	5
2.1.2	Caching . . . . .	6
2.1.3	Asynchroner Informationszugriff . . . . .	8
2.1.4	Informationsaussendung durch Rundsenden . . . . .	9
2.1.5	Zusammenfassung . . . . .	12
2.2	Verwandte Arbeiten . . . . .	13
2.2.1	Taxonomie . . . . .	13
2.2.2	Prefetching-Verfahren . . . . .	14
2.2.3	Hoarding-Verfahren . . . . .	17
2.2.4	Diskussion . . . . .	20

<b>3 Grundlagen</b>	<b>22</b>
3.1 Konzentrationsmaße . . . . .	22
3.1.1 Konzentrationsrate . . . . .	23
3.1.2 Herfindahl-Koeffizient . . . . .	24
3.1.3 Gini-Koeffizient . . . . .	24
3.1.4 Vergleich Herfindahl- versus Gini-Koeffizient . . . . .	26
3.2 Drahtlose Kommunikationstechnologien . . . . .	27
3.2.1 Mobilfunknetze . . . . .	27
3.2.2 Drahtlose Regionalnetzwerke . . . . .	28
3.2.3 Lokale Kommunikationstechnologien . . . . .	29
3.2.4 Zusammenfassung . . . . .	30
3.3 Systemmodell . . . . .	30
3.3.1 Informationssystem . . . . .	31
3.3.2 Kommunikationsinfrastruktur . . . . .	32
3.3.3 Endgeräte . . . . .	34
3.3.4 Externe Dienste . . . . .	34
<b>4 Vorabübertragung ortsbezogener Informationen</b>	<b>37</b>
4.1 Problemstellung . . . . .	37
4.2 Überblick über die Verfahren . . . . .	39
4.2.1 Informationsübertragung . . . . .	42
4.2.2 Entkoppelter Betrieb . . . . .	43
4.2.3 Protokollübertragung . . . . .	43
4.3 Basisverfahren . . . . .	44
4.3.1 Auswertung der Protokolldateien . . . . .	45
4.3.2 Aktualisierung der Zugriffswahrscheinlichkeitstabelle . . . . .	47

4.3.3	Hoarding-Entscheidung . . . . .	49
4.3.4	Diskussion . . . . .	49
4.4	Erweitertes Verfahren . . . . .	50
4.4.1	Grundidee . . . . .	51
4.4.2	Bearbeitung einer Protokolldatei . . . . .	52
4.4.3	Aktualisierung der Zugriffswahrscheinlichkeitstabellen . . . . .	54
4.4.4	Besuchswahrscheinlichkeitskarten . . . . .	54
4.4.5	Externes Wissen . . . . .	56
4.4.6	Integration . . . . .	61
4.4.7	Hoarding-Entscheidung . . . . .	63
4.5	Diskussion . . . . .	64
4.5.1	Gemeinsame Eigenschaften beider Varianten . . . . .	64
4.5.2	Größe von Hoarding-Gebieten und Zonen . . . . .	66
4.5.3	Kosten . . . . .	72
4.6	Erweiterungen . . . . .	79
4.6.1	Informationskanäle . . . . .	80
4.6.2	Filter . . . . .	87
<b>5</b>	<b>Implementierung</b>	<b>89</b>
5.1	Architektur . . . . .	89
5.1.1	Infostation-Server . . . . .	90
5.1.2	Hoarding-Klient . . . . .	92
5.1.3	Externe Dienste . . . . .	97
5.2	Umsetzung . . . . .	99
5.2.1	Infostation-Server . . . . .	99
5.2.2	Hoarding-Klient . . . . .	101
5.2.3	Externe Dienste . . . . .	103
5.3	Einsatzbeispiele . . . . .	104
5.3.1	Web-Browser . . . . .	105
5.3.2	ViLis . . . . .	105
5.3.3	neXus . . . . .	106

<b>6</b>	<b>Simulationsumgebung</b>	<b>108</b>
6.1	Problemstellung . . . . .	108
6.2	Verwandte Arbeiten . . . . .	109
6.2.1	Mobilitätsmodelle . . . . .	109
6.2.2	Informationszugriffsmodelle . . . . .	111
6.3	Simulationsmodell . . . . .	113
6.3.1	Überblick . . . . .	113
6.3.2	Besuchsmodell . . . . .	115
6.3.3	Informationsmodell . . . . .	115
6.3.4	Zugriffsmodell . . . . .	116
6.4	Simulationsalgorithmus . . . . .	116
6.5	Implementierung . . . . .	120
6.5.1	Architektur . . . . .	120
6.5.2	Benutzungsoberfläche . . . . .	121
<b>7</b>	<b>Evaluierung</b>	<b>124</b>
7.1	Eigenschaften eines Informationssystems . . . . .	124
7.1.1	Definitionen . . . . .	125
7.1.2	Erwartetes Verhalten . . . . .	128
7.1.3	Mathematische Analyse . . . . .	129
7.2	Analyse einer existierenden Anwendung . . . . .	134
7.2.1	Beschreibung der Daten . . . . .	134
7.2.2	Ergebnisse . . . . .	136
7.2.3	Vorabübertragung von Informationen im GUIDE-Szenario . . . . .	139
7.2.4	Bewertung . . . . .	140
7.3	Simulationen . . . . .	141
7.3.1	Parametrisierung des Simulationsmodells . . . . .	141
7.3.2	Ergebnisse . . . . .	144

<b>8 Zusammenfassung und Ausblick</b>	<b>153</b>
<b>A Glossar und Abkürzungsverzeichnis</b>	<b>157</b>
<b>B Berechnung des Gini-Koeffizienten</b>	<b>162</b>
<b>Literaturverzeichnis</b>	<b>164</b>

# Abbildungsverzeichnis

2.1	Taxonomie für Verfahren zur Vorabübertragung von Informationen. . .	15
3.1	Lorenzkurve. . . . .	25
3.2	Funktechnologien im Vergleich. . . . .	30
3.3	Infostationen in der Stuttgarter Innenstadt. . . . .	33
4.1	Typische Infostationsinfrastruktur. . . . .	40
4.2	Infostationen und ihre Hoarding-Gebiete. . . . .	41
4.3	Zyklus des Vorabübertragungsverfahrens. . . . .	42
4.4	Wissen über die Bewegung eines Benutzers. . . . .	50
4.5	Informationsquellen für die Hoarding-Entscheidung. . . . .	51
4.6	Graphische Darstellung einer Besuchswahrscheinlichkeitskarte. . . . .	55
4.7	Externe Besuchswahrscheinlichkeitskarten. . . . .	58
4.8	Integration aller verfügbaren BWK in eine endgültige BWK. . . . .	62
4.9	Berechnung der Zugriffswahrscheinlichkeiten. . . . .	63
4.10	Abdeckung eines Gebiets durch einzelne Hoarding-Gebiete. . . . .	68
4.11	Trefferrate in Abhängigkeit von der Größe des Hoarding-Gebiets. . . . .	70
4.12	Trefferrate in Abhängigkeit von der Anzahl zusammengefasster Zonen. . . . .	72
4.13	Kommunikationsaufwand zur Verteilung von Protokolldateien. . . . .	75

4.14	Alternative 1 zur Berücksichtigung von Informationskanälen. . . . .	84
4.15	Alternative 2 zur Berücksichtigung von Informationskanälen. . . . .	85
5.1	Komponenten der Infostation-Server. . . . .	90
5.2	Komponenten der Hoarding-Klienten. . . . .	93
5.3	Transparente Koexistenz. . . . .	96
5.4	Beschreibung eines Hoarding-Gebiets. . . . .	101
5.5	Benutzungsoberflächen der Hoarding-System-Komponenten. . . . .	102
5.6	Benutzungsoberfläche des ViLis-Systems. . . . .	106
5.7	Integration des Hoarding-Systems in neXus. . . . .	107
6.1	Zipf-Verteilung und Zipf-ähnliche Verteilung. . . . .	112
6.2	Infrastrukturgraph zur Modellierung der Stuttgarter Innenstadt. . . . .	114
6.3	Berechnung der kürzesten Route. . . . .	119
6.4	Die Architektur der Simulationsumgebung. . . . .	121
6.5	Hauptfenster der Simulationsumgebung. . . . .	122
7.1	Ortsbezogene und lokal fokussierte Zugriffsmatrix. . . . .	129
7.2	Trefferraten für unterschiedlich fokussierte Informationsräume. . . . .	130
7.3	Trefferrate in Abhängigkeit von der Besuchsrate. . . . .	134
7.4	Die Benutzungsoberfläche der GUIDE-Anwendung. . . . .	135
7.5	Verteilung der Zugriffe auf einzelne Informationsobjekte. . . . .	136
7.6	Anzahl der Informationsobjekte und Informationszugriffe. . . . .	137
7.7	Lorenzkurven. . . . .	139
7.8	Durchschnittlich erzielte Trefferraten. . . . .	140
7.9	Trefferraten in Abhängigkeit von Anzahl der Informationsobjekte und Zugriffswahrscheinlichkeitsverteilung. . . . .	145

7.10 Trefferraten in Abhängigkeit von Zugriffswahrscheinlichkeit und Auswahlwahrscheinlichkeit. . . . .	146
7.11 Trefferraten in Abhängigkeit von Anzahl der Objekte und Anzahl der Protokolldateien. . . . .	148
7.12 Trefferraten in Abhängigkeit von Größe des Hoarding-Gebiets und Anzahl der Zonen. . . . .	149
7.13 Besuchsreihenfolge und Nutzen von unscharfem Wissen. . . . .	150



# Tabellenverzeichnis

2.1	Vergleich von Verfahren zur mobilen Datenverwaltung. . . . .	13
3.1	Verteilung der Zugriffshäufigkeiten. . . . .	25
3.2	Häufigkeitsverteilungen. . . . .	26
5.1	XML-Tags zur Beschreibung eines Hoarding-Gebiets. . . . .	100
7.1	Werte der Metriken für das GUIDE-System. . . . .	138
7.2	Werte der Metriken für das simulierte Informationssystem. . . . .	143



## Kapitel 1

# Einleitung

### 1.1 Motivation

Aktuelle Studien belegen, dass sich der Wunsch vieler Menschen nach einem mobilen Informationszugriff, insbesondere in Form so genannter ortsbezogener Dienste (Location-Based Services, LBS), als Motor für die weitere Entwicklung des Mobilfunkmarktes sowie des Marktes für mobile Endgeräte, wie zum Beispiel Notebooks, internetfähige Mobiltelefone (Smart-Phones), Web-Pads und Persönliche Digitale Assistenten (Personal Digital Assistant, PDA), erweisen wird [Pricewaterhouse-Coopers 2001, Heise-Newsticker 2001, Leite & Pereira 2001].

Ortsbezogene Dienste ermöglichen es dem Benutzer, abhängig von seinem gegenwärtigen Aufenthaltsort, auf lokale Informationen, beispielsweise Informationen zu Restaurants, Kinoprogramme oder Staumeldungen, zuzugreifen. Solche ortsbezogenen Dienste werden als Schlüsseltechnologie für die weitere Verbreitung mobiler Endgeräte und mobiler Datenkommunikationsdienste angesehen, weil sie die Alleinstellungsmerkmale mobiler Endgeräte in idealer Weise ausnutzen [Pricewaterhouse-Coopers 2001].

Einer EMNID-Studie [Heise-Newsticker 2001] zufolge, die im Auftrag von America Online (AOL) durchgeführt wurde, geben 88% der Besitzer eines internetfähigen, mobilen Endgeräts den mobilen Informationszugriff als Vorteil ihres Geräts an. Nach dieser Studie besitzen bereits zwölf Millionen Deutsche ein solches Gerät. Neun Millionen planen eine entsprechende Anschaffung innerhalb der nächsten zwei Jahre.

Auch die Funknetzbetreiber haben die Bedeutung ortsbezogener Dienste erkannt, so dass alle deutschen Netzanbieter inzwischen ortsbezogene Dienste über ihre WAP<sup>1</sup>-Portale anbieten. Die bislang mangelnde Akzeptanz dieser WAP-Dienste ist nach [Pricewaterhouse-Coopers 2001] auf die Leistungsbeschränkung der Funknetze beim mobilen Informationszugriff zurückzuführen. Im Zuge der Erforschung mobiler Rechnerumgebungen (Mobile Computing) wurden ebenfalls schon zahlreiche ortsbezogene Systeme entwickelt [Abowd et al. 1997, Pascoe 1997, Cheverst, Davies, Mitchell & Friday 2000]. Als eine der zentralen Herausforderungen hat sich auch dabei immer wieder der mobile Informationszugriff erwiesen, da die verwendeten Funknetztechnologien häufig zu Problemen, wie hohen Verzögerungen, geringen Bandbreiten oder vollständigen Verbindungsausfällen, führen.

Aus dieser Motivation heraus ist die vorliegende Arbeit entstanden. Ihr Ziel ist es, diese Probleme soweit wie möglich zu überwinden und überall einen mobilen Informationszugriff zu ermöglichen, unabhängig von der Verfügbarkeit eines Mobilfunknetzes am aktuellen Aufenthaltsort des Benutzers. Selbst wenn gar keine Kommunikationsmöglichkeit vorhanden ist, soll in diesem so genannten entkoppelten Betrieb, ein möglichst uneingeschränkter Zugriff möglich sein. Der Schwerpunkt dieser Arbeit liegt dabei auf ortsbezogenen Informationssystemen. Als Standardbeispiel für ein solches System taucht in dieser Arbeit immer wieder ein elektronischer, mobiler Touristenführer auf. Im Gegensatz zu den existierenden WAP-Diensten ermöglicht das vorgestellte Verfahren einen Zugriff auf sehr viel umfangreichere, auch multimediale Informationsobjekte.

Die Grundidee dieses Verfahrens ist es, vorherzusagen auf welche Informationsobjekte ein Benutzer voraussichtlich zugreifen wird, und diese Objekte im Voraus auf sein mobiles Endgerät zu übertragen. Diese Vorabübertragung von Informationen wird in der Literatur als Hoarding bezeichnet. Greift der Benutzer nun tatsächlich auf die vorab übertragenen Informationsobjekte zu, sind sie bereits lokal auf dem mobilen Endgerät vorhanden und stehen somit unmittelbar zur Verfügung. Um zu vermeiden, dass knappe Ressourcen, wie zum Beispiel der auf dem mobilen Endgerät verfügbare Speicherplatz oder die Zeit für die Übertragung der Informationen, vergeudet werden, sollten möglichst nur solche Informationsobjekte vorab übertragen werden, auf die der Benutzer auch tatsächlich zugreifen wird. Daher ist eine möglichst präzise Vorhersage der vom Benutzer zukünftig benötigten Informationen notwendig. Das entwickelte

---

<sup>1</sup>Wireless Application Protocol

Verfahren nutzt für diese Vorhersage Beziehungen zwischen Informationen und Orten, die in ortsbezogenen Systemen vorhanden sind, um aus Kenntnissen über den aktuellen und die zukünftigen Aufenthaltsorte eines Benutzers Rückschlüsse auf die benötigten Informationen zu ziehen.

Im Gegensatz zu existierenden Ansätzen werden jedoch keine festen, bekannten Beziehungen zwischen Orten und Informationen vorausgesetzt, sondern das Verfahren erkennt diese selbständig und kann sogar dynamisch auf Veränderungen in diesen Beziehungen reagieren. Die Kenntnisse über die zukünftigen Aufenthaltsorte eines Benutzers kann das Verfahren aus selbständig durchgeführten Beobachtungen oder durch verschiedene externe Informationsquellen gewinnen. Damit ist es sehr viel flexibler als vergleichbare Ansätze [Kubach & Rothermel 2001c, Kubach & Rothermel 2001b, Kubach & Rothermel 2001a].

Eine wichtige Randbedingung bei der Entwicklung des Verfahrens war, dass es in die neXus-Plattform integrierbar sein sollte. Diese Plattform, die im Rahmen einer Forschergruppe der Deutschen Forschungsgemeinschaft (DFG) unter Federführung der Abteilung Verteilte Systeme des Instituts für Parallele und Verteilte Höchstleistungsrechner (IPVR) entwickelt wird, wird eine neuartige Infrastruktur für ortsbezogene Systeme bereitstellen [Hohl et al. 1999].

## 1.2 Aufbau der Arbeit

Im anschließenden Kapitel 2 wird die Idee der Vorabübertragung von Informationen anderen mobilen Datenverwaltungsmechanismen gegenübergestellt und dabei verdeutlicht, dass die Vorabübertragung speziell in mobilen, ortsbezogenen Informationssystemen besonders vorteilhaft ist. Ferner wird die Arbeit dadurch in den Forschungskontext eingeordnet. Die Diskussion existierender Ansätze zur Vorabübertragung von Informationen ist ebenfalls Gegenstand dieses Kapitels.

In Kapitel 3 schließt sich die Besprechung einiger Grundlagen an. Zunächst werden verschiedene stochastische Konzentrationsmaße vorgestellt, die später für die Evaluierung der Arbeit benötigt werden. Aus einer Diskussion der verfügbaren Mobilfunktechnologien wird danach die Motivation für die verwendete Infostationsinfrastruktur abgeleitet. Diese Infrastruktur ist neben dem Informationssystem und den verwendeten externen Diensten, wie Ortsinformationsdienst und Ereignisdienst, ein wichtiger

Bestandteil des ebenfalls in Kapitel 3 vollständig beschriebenen Systemmodells, das dieser Arbeit zugrunde liegt.

Das zentrale Kapitel dieser Arbeit ist Kapitel 4. Hier wird das Vorabübertragungsverfahren im Detail beschrieben. Eigentlich werden zwei Verfahren beschrieben, ein Basisverfahren und ein erweitertes Verfahren. Während das Basisverfahren keine externen Informationen über die zukünftigen Aufenthaltsorte nutzen kann, ist dies im erweiterten Verfahren mit einer beliebig feinen Granularität möglich. Dadurch kann das erweiterte Verfahren zum Beispiel Pfadinformationen aus einer Navigationsanwendung beziehen oder auch Richtungsangaben von einer Komponente, die die Bewegungshistorie eines Benutzers beobachtet. Als Mittel zur Spezifikation von solchem externen Wissen werden so genannte Besuchswahrscheinlichkeitskarten eingeführt. Schließlich werden noch zwei weitere Konzepte vorgestellt, mit denen die Präzision der Vorhersage der benötigten Informationsobjekte weiter gesteigert werden kann. Das erste Konzept sind die Informationskanäle, die es ermöglichen das Profil eines Benutzers bei der Entscheidung, welche Objekte übertragen werden sollen, zu berücksichtigen. Diese Entscheidung wird nachfolgend als Hoarding-Entscheidung bezeichnet. Das zweite Konzept sind Filter. Sie erlauben es bei der Hoarding-Entscheidung stärker auf die individuellen Bedürfnisse eines einzelnen Benutzers einzugehen.

Die vollständige Implementierung beider Verfahren wird in Kapitel 5 vorgestellt. Neben einer Beschreibung der Architektur sowie ihrer Umsetzung werden Beispiele für den Einsatz des entstandenen Systems gegeben.

Für die Evaluierung der Vorabübertragungsverfahren musste ein spezielles Simulationsmodell entworfen werden, das die besonderen Charakteristika eines mobilen, ortsabhängigen Informationszugriffs widerspiegelt. Dieses Simulationsmodell wird in Kapitel 6 vorgestellt. Ferner wird die Simulationsumgebung, die auf Basis dieses Modells entwickelt wurde und für die Evaluierung eingesetzt wurde, beschrieben.

In Kapitel 7 werden die damit durchgeführten Simulationen erläutert und dadurch die Effizienz der Verfahren nachgewiesen. Zu Beginn des Kapitels werden einige theoretische Betrachtungen angeführt, sowie eine existierende ortsbezogene Anwendung analysiert, um daraus die Simulationsparameter abzuleiten.

Kapitel 8 beschließt diese Abhandlung schließlich mit einer Zusammenfassung der wichtigsten Ergebnisse und einem Ausblick auf mögliche zukünftige Forschungsgebiete.

## **Kapitel 2**

# **Mobile Datenverwaltung**

Zunächst erfolgt in diesem Kapitel eine Einordnung der vorliegenden Arbeit in das Forschungsumfeld der mobilen Datenverwaltung. Ferner werden existierende Verfahren zur Vorabübertragung von Informationen vorgestellt und klassifiziert, um sie gegenüber dem in dieser Arbeit vorgestellten Verfahren abzugrenzen.

## **2.1 Einordnung der Arbeit**

Um den Nachteilen der drahtlosen Netzwerke beim mobilen Informationszugriff entgegen zu wirken, wurde bereits eine Reihe von Methoden entwickelt. Diese Methoden werden in der Literatur häufig unter dem Oberbegriff mobile Datenverwaltung (mobile data management) zusammengefasst. In diesem Abschnitt wird verdeutlicht, worin sich die Vorabübertragung von Informationen von den übrigen Verfahren zur mobilen Datenverwaltung unterscheidet. Ferner wird die Eignung dieser Verfahren für den ortsabhängigen Informationszugriff diskutiert.

### **2.1.1 Replikation**

Ziel der Replikation ist es, durch das Vorhalten von Datenkopien die Verfügbarkeit von Daten zu erhöhen und die Antwortzeiten von Software-Systemen zu reduzieren.

Wesentliche Fragen dabei sind, wo die Kopien vorgehalten werden sollen und wie ihre Konsistenz gewährleistet werden kann.

In mobilen Umgebungen werden die Datenkopien entweder an den Übergangspunkten vom Festnetz zum drahtlosen Netzwerk (Access Points) oder auf den mobilen Endgeräten selbst vorgehalten. Bei den Verfahren zur Gewährleistung der Konsistenz unterscheidet man die pessimistischen und die optimistischen Verfahren [Davidson et al. 1985]. Die pessimistischen Verfahren verbieten durch entsprechende Sperrprotokolle, dass gleichzeitig auf verschiedenen Kopien geschrieben wird. Nach einem Schreibzugriff werden Änderungen automatisch an alle Kopien propagiert. Pessimistische Verfahren können nicht eingesetzt werden, wenn in einem entkoppelten Betrieb ein Schreibzugriff auf den lokalen Kopien eines mobilen Endgeräts möglich sein soll, weil die für die Sperrprotokolle notwendige Kommunikation nicht durchgeführt werden kann. Daher können dann nur optimistische Verfahren eingesetzt werden. Sie erlauben grundsätzlich das Schreiben auf den Datenkopien. Sollten dabei Inkonsistenzen durch unterschiedliche Änderungen an den gleichen Daten auftreten, müssen diese zu einem späteren Zeitpunkt, eventuell unter Einbeziehung der Benutzer, aufgelöst werden [Satyanarayanan et al. 1990, Heidemann et al. 1992, Demers et al. 1994].

Im Gegensatz zur Replikation werden bei Vorabübertragungsverfahren, wie dem in dieser Arbeit beschriebenen Verfahren, in der Regel Konsistenzbedingungen nicht betrachtet. Teilweise setzen die Vorabübertragungsverfahren auf Replikationssysteme auf, um dennoch eine konsistente Datenhaltung zu garantieren. Ein weiterer Unterschied ist, dass die Vorabübertragung von Informationen vom Klienten initiiert wird, während sie in den meisten Replikationssystemen vom System beziehungsweise den Replikations-Servern getrieben wird. Bei der Vorabübertragung von Informationen spielt auch die Vorhersage der Bedürfnisse eines einzelnen Klienten eine wichtige Rolle für die Entscheidung, welche Daten auf ein Endgerät kopiert werden. Bei den Replikationsverfahren hingegen basiert diese Entscheidung in der Regel auf globaleren Kriterien.

### 2.1.2 Caching

Eine weit verbreitete Methode zur Beschleunigung von Informationszugriffen, die sowohl in Festnetzen als auch in drahtlosen Netzwerken Verwendung findet, ist das Caching von Informationen. Die Grundidee beim Caching ist es, Informationen, die vom



Server zum Klienten übertragen wurden, für eine spätere Verwendung aufzubewahren. Wird dann auf die gleiche Information erneut zugegriffen, ist diese bereits auf dem Klienten vorhanden und muss nicht erneut vom Server geholt werden.

Im Gegensatz zur Replikation wird das Caching vom Klienten getrieben, das heißt Kopien von Daten werden ausschließlich dann angelegt, wenn eine Applikation auf dem Klienten eine Information anfordert. Ferner wird beim Caching grundsätzlich nicht von schreibenden Zugriffen auf den lokalen Kopien ausgegangen. Einige Arbeiten, insbesondere im Bereich mobiler Umgebungen, beschäftigen sich damit, wie die Klienten davon in Kenntnis gesetzt werden können, dass ihre lokalen Kopien veraltet sind, damit sie gegebenenfalls das Original neu anfordern können [Barbará & Imieliński 1994, Chan et al. 1998].

Greift ein Benutzer auf viele verschiedene Informationen zu, wird irgendwann der Speicherplatz, der für die Aufbewahrung der Informationen zur Verfügung steht, aufgebraucht sein. Dann muss eine so genannte Ersetzungsstrategie entscheiden, welche Informationen gelöscht werden sollen, um wieder freien Speicherplatz zu schaffen. Die meisten Ersetzungsstrategien beruhen auf dem Prinzip der temporalen Lokalität. Dieses Prinzip besagt, dass auf eine Information, auf die zu einem bestimmten Zeitpunkt zugegriffen wird, mit hoher Wahrscheinlichkeit in naher Zukunft erneut zugegriffen werden wird. Entsprechend löscht beispielsweise die häufig eingesetzte Strategie Least Recently Used (LRU) immer die Information, auf die am längsten nicht mehr zugegriffen wurde. Neben der LRU-Strategie gibt es noch eine Reihe ähnlicher Strategien, die jedoch seltener eingesetzt werden [Effelsberg & Härder 1984].

Beim ortsabhängigen Informationszugriff verändert sich durch einen Ortswechsel der Kontext eines Benutzers jedoch meist so stark, dass dadurch die temporale Lokalität in den Informationszugriffen verloren geht. Sobald ein Benutzer seinen Aufenthaltsort wechselt, ist es unwahrscheinlich, dass er weiterhin auf die kurz zuvor referenzierten Informationen zugreifen wird, da diese primär an seinem vorherigen Aufenthaltsort von Interesse waren oder sogar nur von diesem aus abrufbar waren.

Daher wurden speziell für mobile Umgebungen semantische Ersetzungsstrategien entwickelt, die Beziehungen zwischen dem Kontext des Benutzers und den referenzierten Informationen ausnutzen. Sehr häufig wird dabei anstatt von einer temporalen Lokalität von einer geographischen Lokalität ausgegangen, das heißt es wird angenommen, dass die Informationen mit dem Ort, an dem auf sie zugegriffen wird, in Beziehung

stehen. Ren & Dunham (2000) beschreiben den Kontext eines Benutzers ausschließlich durch seinen Aufenthaltsort. Wie in dem in dieser Arbeit beschriebenen Verfahren wird versucht, Beziehungen zwischen den Aufenthaltsorten des Benutzers und den dort abgerufenen Informationsobjekten zu erkennen, das heißt es wird versucht, die Informationsobjekte geographischen Orten zuzuordnen. Auf Basis dieser Zuordnung können dann gegebenenfalls die Informationsobjekte aus dem Speicher des mobilen Endgeräts gelöscht werden, die den am weitesten vom aktuellen Aufenthaltsort des Benutzers entfernten Orten zugeordnet sind (Furthest Away Replacement Strategie).

Chim et al. (1998) beschreiben eine Ersetzungsstrategie, die speziell für die Verwaltung eines Caches zur Unterstützung eines Spaziergangs durch eine virtuelle Welt entwickelt wurde. Die Autoren nennen ihre Strategie Most Required Movement und berücksichtigen neben den Orten der virtuellen Objekte zusätzlich die Blickrichtung des Benutzers. Gwertzman & Seltzer (1995) gehen in ihrem Web-Caching Ansatz für Festnetze ebenfalls von einer geographischen Lokalität in den Informationsanfragen aus.

In Mobilfunksystemen werden ebenfalls semantische Ersetzungsstrategien eingesetzt, zum Beispiel zur Verwaltung von Caches, die Informationen über den Aufenthaltsort der Benutzer enthalten. Dabei stellt die geringe Gültigkeitsdauer der Informationen eine besondere Herausforderung dar. Häufig entscheidet die Mobilität eines Benutzers darüber, ob es sich überhaupt lohnt, Informationen über seinen Aufenthaltsort in einem Cache zu speichern [Jain et al. 1994].

Unabhängig von der verwendeten Ersetzungsstrategie bleibt beim Caching jedoch der Nachteil, dass es den jeweils ersten Zugriff auf eine bestimmte Information nicht beschleunigt. Entsprechend muss der Benutzer bei einem solchen ersten Zugriff, im Gegensatz dazu wie es durch die Vorabübertragung von Informationen möglich ist, grundsätzlich eine Netzwerkverbindung zum Server haben und darauf warten bis die Information vom Server zum Klienten übertragen wurde. Beim ortsabhängigen Informationszugriff wirkt sich dieser Nachteil besonders stark aus, da sehr häufig überhaupt nur einmal auf die jeweiligen Informationsobjekte zugegriffen wird. Selbst wenn ein Benutzer zu einem bereits besuchten Ort zurückkehrt, ist es dennoch unwahrscheinlich, dass er erneut Informationen über diesen Ort anfordert.

### 2.1.3 Asynchroner Informationszugriff

Der ARTour Web Express ist ein System von IBM, das es ermöglichen soll, World Wide Web (WWW) Anwendungen, insbesondere Browser, über drahtlose Weitverkehrsnet-

ze zu betreiben [Chang et al. 1997, Housel & Lindquist 1996]. Dazu wurden zahlreiche Optimierungen am Hypertext Transferprotokoll (HTTP) [Fielding et al. 1999], das üblicherweise von solchen Anwendungen benutzt wird, vorgenommen. Diese Optimierungen, die sich im Wesentlichen in einer Reduktion der zu übertragenden Daten auswirken, sind jedoch bei vollständigen Verbindungsausfällen nutzlos.

Um dennoch einen entkoppelten Betrieb zu ermöglichen, wird zusätzlich ein asynchroner Informationszugriff unterstützt. Sobald Informationsanfragen auf Grund von Netzwerkausfällen nicht beantwortet werden können, werden sie in eine Warteschlange eingereiht. Der WWW-Anwendung wird eine Seite mit der Meldung „Anfrage wurde für eine spätere Bearbeitung gepuffert“ als Antwort auf die Anfrage zurückgeliefert. Für die Anwendung ist der asynchrone Zugriff also transparent. Sobald eine Netzwerkverbindung verfügbar wird, beginnt ARTour die Warteschlange abzuarbeiten. Wenn das vollständige Ergebnis einer Anfrage vorliegt, wird es in den lokalen Cache eingetragen. Danach wird der Benutzer wahlweise über ein Informationsfenster, einen Statusbalken, der allen angezeigten Seiten hinzugefügt wird oder eine Statusseite darüber informiert, dass eine seiner Anfragen vollständig bearbeitet ist, und er kann sie über den ihm angezeigten Link aus dem lokalen Cache abrufen.

Im Rover HTTP-Proxy, der mit Hilfe des Rover Toolkits entwickelt wurde, wird der gleiche Ansatz verfolgt [Joseph et al. 1997]. Auch hier werden Anfragen nach nicht lokal vorgehaltenen Informationsobjekten zurückgestellt, bis eine Netzwerkverbindung verfügbar wird.

In vielen Fällen, wie beispielsweise dem in [Chang et al. 1997] geschilderten Szenario, ermöglicht der asynchrone Informationszugriff einen sinnvollen entkoppelten Betrieb. Beim ortsabhängigen Informationszugriff hingegen ergibt sich jedoch das Problem, dass ein Benutzer unter Umständen erst dann angeforderte Informationen erhält, wenn er seinen Standort gewechselt hat und daher an den Informationen, die seinen vorherigen Standort betrafen, nicht mehr interessiert ist.

### 2.1.4 Informationsaussendung durch Rundsenden

In [Tan & Ooi 2000] wird eine umfassende Taxonomie für Strategien zur Datenübertragung in mobilen Umgebungen vorgeschlagen. Diese Taxonomie charakterisiert die Strategien auf Grund von vier verschiedenen Dimensionen. Für die Betrachtungen in diesem Abschnitt sind jedoch lediglich die ersten drei Dimensionen relevant.

In der ersten Dimension werden die Ansätze dadurch unterschieden, ob sie die Daten ausschließlich durch Rundsenden übertragen oder ausschließlich Klientenanfragen beantworten oder ob sie eine Mischform der beiden ersten Möglichkeiten verwenden. Im letzten Fall spricht man von sogenannten hybriden Ansätzen.

Durch die Betrachtung der Mechanismen zur Datenaussendung ergibt sich die zweite Dimension. Hier werden die Ansätze unter anderem danach unterschieden, ob die Datenaussendung durch den Server oder den Klienten initiiert wird. Auch die Frage, ob die Daten auf Grund eines festen Zeitplans (Zeitplan-getriebene Verfahren) oder auf Grund von bestimmten Ereignissen (Ereignis-getriebene Verfahren) verschickt werden, spielt hier eine Rolle.

Die dritte Dimension befasst sich mit der Organisation der Datenaussendung. Wenn die Datenaussendung über Rundsende-Kommunikation erfolgt, kann man je nachdem nach welcher Vorschrift die Daten gesendet werden, verschiedene Klassen identifizieren. Die erste Klasse bilden die Strategien, bei denen es eine feste Reihenfolge gibt, in der die Daten verschickt werden. Hier kann man noch unterscheiden, ob innerhalb eines Rundsendezyklus jede Information genau einmal versendet wird, Tan & Ooi (2000) spricht dann von flachen Rundsende-Programmen, oder ob häufig benötigte Informationen mehrmals innerhalb eines Zyklus verschickt werden (nicht-flache Rundsende-Programme). Bei der zweiten Klasse von Strategien werden die Daten ad hoc verschickt. Dabei kann zwischen Strategien, die mehrere Anfragen puffern, um die Zahl der Übertragungen der gleichen Information zu minimieren, und solchen, die das nicht tun, unterschieden werden. Zusätzlich können zu Beginn und am Ende eines Rundsendezyklus Meta-Daten verschickt werden, die beschreiben, welche Daten innerhalb des Zyklus gesendet werden bzw. wurden. Je nachdem, ob solche Indizes verschickt werden oder nicht, teilt Tan & Ooi (2000) die Verfahren in indizierte und nicht indizierte Verfahren ein.

Bezüglich der ersten Dimension dieser Taxonomie gehören die hier vorgestellten Verfahren entweder zur Klasse der ausschließlich rundsende-basierten Ansätze oder zur Klasse der hybriden Ansätze. Die Grundidee bei diesen Verfahren ist es, häufig benötigte Informationen periodisch zu allen Endgeräten innerhalb einer Funkzelle des drahtlosen Netzwerks zu senden. Die Endgeräte puffern diese Informationen, so dass sie bei einem Zugriff des Benutzers bereits auf dem Endgerät vorhanden sind. Man spricht daher in diesem Zusammenhang auch von Push-Caching. Neben einer Reduzierung der Antwortzeiten ergibt sich durch das Rundsenden auch eine Ersparnis in

der benötigten Bandbreite, da die häufig benötigten Informationen jeweils nur einmal übertragen werden müssen, anstatt sie für jede Anfrage eines Klienten erneut übertragen zu müssen. Die hybriden Ansätze unterscheiden sich von den ausschließlich rundsende-basierten Ansätzen darin, dass sie zusätzlich zum Rundsenden der Informationen auch einzelne Anfragen von Klienten zulassen und diese direkt beantworten. Dies ist zum Beispiel von Vorteil, wenn ein Klient eine Funkzelle, unmittelbar nachdem eine von ihm benötigte Information gesendet wurde, betreten hat. Dann kann er diese direkt anfordern und muss nicht darauf warten bis sie ein weiteres Mal durch das Rundsenden verbreitet wird.

Im mobilen Touristenführer GUIDE [Cheverst, Davies, Mitchell & Friday 2000] wird ein reines Rundsendeverfahren eingesetzt [Davies et al. 1999]. Die Kommunikationsinfrastruktur besteht aus WaveLAN<sup>1</sup>-Funkzellen. In jeder dieser Funkzellen gibt es einen Proxy-Server, der einen Cache für die innerhalb der Funkzelle abgerufenen Informationen verwaltet. Eine Teilmenge dieses Caches wird periodisch durch Rundsenden verschickt. Leider machen die Autoren keine genaueren Angaben darüber, wie diese Teilmenge bestimmt wird. Zu Beginn und am Ende eines jeden Rundsendezyklus wird eine Liste der Bezeichner der Objekte, die für das Rundsenden vorgesehen sind, verschickt. Damit kann ein Klient entscheiden, ob er ein von ihm benötigtes Informationsobjekt im aktuellen bzw. nächsten Rundsendezyklus empfangen kann. Falls nicht, kann er es beim Proxy-Server anfordern, der es daraufhin in den nächsten Zyklus aufnimmt. Obwohl es sich bei dem Verfahren um ein reines Rundsendeverfahren handelt, kommt die Initiierung einer Informationsübertragung also von den Klienten. Entsprechend der Taxonomie von Tan & Ooi (2000) lässt sich das Verfahren weiter als indiziertes, Zeitplan-getriebenes Verfahren mit flachem Rundsendeprogramm charakterisieren.

Eine weitere wichtige Klasse von Rundsendeverfahren sind die so genannten Rundsende-Platten (Broadcast Discs), wie zum Beispiel das in [Acharya et al. 1995] beschriebene Verfahren. In ihrer Grundform handelt es sich bei diesen Verfahren um reine, Server-initiierte, Zeitplan-getriebene Rundsendeverfahren. In der Regel benutzen sie ein nicht-flaches Rundsendeprogramm. Der Name dieser Verfahren kommt von der Idee, dass der knappe lokale Speicher der mobilen Endgeräte dadurch erweitert wird, dass sie auf die Daten im Rundsendezyklus wie auf eine, wenn auch sehr langsame, Festplatte zugreifen.

---

<sup>1</sup><http://www.orinocowireless.com/>

Weitere Arbeiten zum Rundsenden von Informationen beschäftigen sich damit, wie Indizes genutzt werden können, um den Energiebedarf der mobilen Endgeräte zu senken [Imieliński et al. 1994]. Die Idee dabei ist, durch die Indizes die Endgeräte darüber zu informieren, wann für sie interessante Daten gesendet werden. Dann müssen sie während der übrigen Zeit nicht im Empfangsmodus arbeiten, wodurch Energie gespart werden kann.

Der Unterschied aller Verfahren zur Informationsaussendung durch Rundsenden zu dem in dieser Arbeit vorgestellten Verfahren zur Vorabübertragung von Informationen ist, dass sie keinen entkoppelten Betrieb unterstützen. Zwar können auch per Rundsendeverfahren in einen lokalen Cache übertragene Informationen in einem entkoppelten Betrieb abgerufen werden, aber die Rundsendeverfahren unterstützen dies nicht explizit.

Bei einem ortsabhängigen Informationszugriff würde sich das darin äußern, dass ein Benutzer ausschließlich Informationen erhält, die innerhalb der aktuellen Funkzelle von Interesse sind. Wollte man einen entkoppelten Betrieb unterstützen, müsste man alle Benutzer, unabhängig davon in welche Richtung sie die Funkzelle verlassen, durch den gleichen Rundsendezyklus mit den zukünftig benötigten Informationen versorgen. Bei einem ortsabhängigen Informationszugriff würden sich die Informationsbedürfnisse dieser Benutzer je nach Bewegungsrichtung jedoch stark unterscheiden. Aber gerade bei solchen stark unterschiedlichen Informationsbedürfnissen sind die Rundsendeverfahren nicht sehr effizient. Wie später noch erläutert wird, können die Rundsendeverfahren dennoch zur Optimierung der Vorabübertragung der Informationen, wie sie in dieser Arbeit beschrieben wird, eingesetzt werden (siehe Abschnitt 4.2).

### 2.1.5 Zusammenfassung

Tabelle 2.1 fasst die Unterschiede der vorgestellten Verfahren zur mobilen Datenverwaltung zusammen. Die in der Tabelle enthaltenen Aussagen betreffen jeweils typische Vertreter der einzelnen Verfahren. Selbstverständlich gibt es auch untypische Vertreter, für die ein Teil der jeweiligen Aussagen nicht zutreffen.

	Initiator	Konsistenz	Vorhersage	Entkoppelter Betrieb
Replikation	Server	ja, Schreib- u. Lesezugriffe	teilweise, globale Kriterien	möglich bei optimistischen Verfahren
Caching	Klient	teilweise, nur Lesezugriffe	Ersetzungsstrategie	möglich
Asynchroner Informationszugriff	Klient	entfällt	nein	ja, Daten erst später verfügbar
Rundsendeverfahren	Server	teilweise, nur Lesezugriffe	teilweise, globale Kriterien	möglich
Vorabübertragung	Klient	nein	ja, individuell	ja, explizit unterstützt

**Tabelle 2.1:** Vergleich von Verfahren zur mobilen Datenverwaltung.

## 2.2 Verwandte Arbeiten

Zunächst wird in diesem Abschnitt eine Taxonomie für Verfahren zur Vorabübertragung von Informationen vorgestellt. Mit dieser Taxonomie wird das gesamte Spektrum der existierenden Verfahren aufgezeigt, bevor schließlich konkrete Verfahren beschrieben werden.

### 2.2.1 Taxonomie

In der englischsprachigen Literatur werden die Verfahren zur Vorabübertragung von Informationen in so genannte Prefetching-Verfahren und Hoarding-Verfahren unterteilt. Bei den Prefetching-Verfahren wird davon ausgegangen, dass stets eine Netzwerkverbindung verfügbar ist. Folglich findet in diesen Verfahren die Vorabübertragung von Informationen kontinuierlich und ausschließlich zur Reduzierung der Antwortzeiten statt. Hoarding-Verfahren hingegen zielen hauptsächlich darauf ab, einen entkoppelten Betrieb zu ermöglichen. Die Vorabübertragung der Informationen findet bei diesen Verfahren zu einem wohldefinierten Zeitpunkt unmittelbar vor der Entkopplung vom Netzwerk statt. Die Unterscheidung zwischen Prefetching- und Hoarding-Verfahren wird ausschließlich in diesem Abschnitt aufrechterhalten, um die Unterschiede zwischen den beschriebenen Verfahren zu verdeutlichen.

Bei der Art des unterstützten Systems lassen sich bei den existierenden Verfahren bislang zwei Klassen unterscheiden. Die meisten Verfahren sind für die Unterstützung von Dateisystemen entworfen worden. Ein kleinerer Teil unterstützt Informationssysteme, wie beispielsweise Navigationsanwendungen oder den mobilen Zugriff auf das World Wide Web (WWW).

Weiter können die Verfahren danach unterschieden werden, ob der Benutzer zwingend in die Entscheidung, welche Informationen im Voraus übertragen werden sollen, involviert ist oder nicht. Im ersten Fall spricht man von interaktiven Verfahren, im zweiten Fall von automatischen Verfahren.

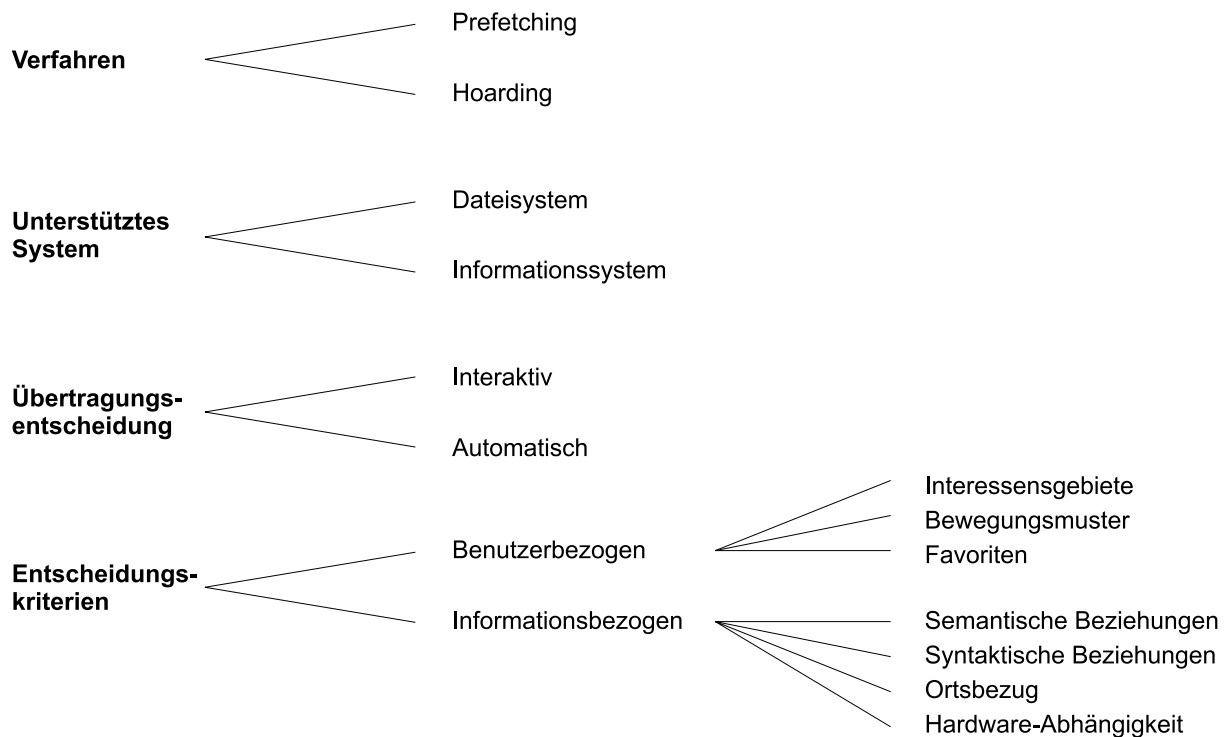
Auch die Kriterien, die für die Entscheidung, welche Informationen übertragen werden sollen, herangezogen werden, sind sehr unterschiedlich. Man kann sie in die Klasse der benutzerbezogenen Kriterien und der informationsbezogenen Kriterien einteilen. Benutzerbezogene Kriterien sind zum Beispiel Interessensgebiete, Favoriten (Lesezeichen in Browsern) oder typische Bewegungsmuster eines Benutzers. Beispiele für informationsbezogene Kriterien sind semantische Beziehungen zwischen den Informationen, wie sie häufig beim Hoarding von Dateien ausgenutzt werden [Tait et al. 1995, Kuenning & Popek 1997], aber auch syntaktische Beziehungen, wie Hypertext-Links. Der Ortsbezug einer Information und die Darstellbarkeit einer Information auf dem Endgerät des Benutzers sind weitere informationsbezogene Kriterien.

Abbildung 2.1 fasst die vier Dimensionen der vorgestellten Taxonomie und ihre möglichen Ausprägungen nochmals zusammen. Bei dem in dieser Arbeit vorgestellten Verfahren handelt es sich um ein automatisches Hoarding-Verfahren, das ortsabhängige Informationssysteme unterstützt. Als Entscheidungskriterien werden sowohl benutzerbezogene Kriterien, wie Bewegungsmuster und Interessensgebiete, als auch informationsbezogene Kriterien, insbesondere der Ortsbezug von Informationen, verwendet.

### 2.2.2 Prefetching-Verfahren

Die Idee, Daten basierend auf Vorhersagen schon zu übertragen, bevor sie tatsächlich benötigt werden, ist nicht neu. Bereits sehr früh wurde diese Technik beim Prozessor-Pipelining verwendet. Hier wird versucht, Verzweigungen eines Maschinenprogramms vorherzusagen, um die entsprechenden Befehle und die zugehörigen Operanden bereits vor der eigentlichen Ausführung des Befehls in eine Prozessor-Pipeline





**Abbildung 2.1:** Taxonomie für Verfahren zur Vorabübertragung von Informationen.

zu laden [Hennessy & Patterson 1990]. Auch bei der seitenbasierten Realisierung von verteiltem gemeinsamen Speicher [Li & Hudak 1989] wurden vergleichbare Techniken bereits eingesetzt.

Frühe Arbeiten im Bereich des Datei-Prefetchings gehen von einem sequentiellen Dateizugriff aus und beschreiben Dateisysteme, die Blöcke einer Datei im Voraus lesen [Smith 1985]. Patterson et al. (1993) schlagen vor, Hinweise von Anwendungen für ein Datei-Prefetching auszunutzen. Als Beispiel führen sie an, dass das Werkzeug make mit Hilfe des Makefiles sehr genau vorhersagen kann, welche Dateien für die Übersetzung eines Programms benötigt werden [Patterson et al. 1993].

In zellbasierten, mobilen Umgebungen werden zur Effizienzsteigerung häufig die von einem Benutzer benötigten Daten und Dienste in der Nähe des mit dem Benutzer assoziierten Zugangspunkts repliziert oder in einem Cache bereitgestellt. Wechselt ein Benutzer die Funkzelle, ändert sich auch der assoziierte Zugangspunkt. Durch die-

sen als Handover bezeichneten Vorgang, kann der Benutzer einen vorübergehenden Leistungseinbruch wahrnehmen, weil an seinem neuen Zugangspunkt die benötigten Daten und Dienste noch nicht zur Verfügung stehen. Um dies zu vermeiden, werden zum Teil ebenfalls Prefetching-Verfahren eingesetzt. In [Liu et al. 1996] wird ein solches Verfahren beschrieben, das die Bewegung eines Benutzers auf Grundlage von Bewegungsmustern vorhersagt. Entsprechend dieser Vorhersage werden die benötigten Daten und Dienste im Voraus am zukünftigen Zugangspunkt des Benutzers bereitgestellt. Es wird also ausschließlich der Aufenthaltsort vorhergesagt und lediglich die aktuelle Umgebung des Benutzers am zukünftigen Aufenthaltsort bereitgestellt. Eine Vorhersage, ob sich die benötigten Daten und Dienste ändern, wird nicht gemacht.

Neben der bereits erwähnten Caching-Strategie beschreiben die Autoren in [Chim et al. 1998] auch eine Prefetching-Strategie, die ebenfalls auf der Vorhersage der Bewegung eines Benutzers beruht. Für die Vorhersage wird versucht, von in der Vergangenheit beobachteten Bewegungsvektoren eines Benutzers auf seinen zukünftigen Bewegungsvektor zu schließen. Es werden drei Varianten des Vorhersagealgorithmus beschrieben, die sich darin unterscheiden wie die jeweiligen Bewegungsvektoren bei der Berechnung des neuen Vektors gewichtet werden. In der ersten Variante werden alle Vektoren gleich gewichtet, in der zweiten werden grundsätzlich nur die letzten  $n$  Vektoren betrachtet, während bei der dritten wiederum alle Vektoren berücksichtigt werden, allerdings mit exponentiell abnehmender Gewichtung der Vektoren entsprechend ihrem Alter. Auf Grundlage der vorhergesagten Position des Benutzers ist die Vorhersage der dort für die Darstellung der virtuellen Welt benötigten Objekte trivial. Da alle Objekte eine feste Position besitzen, kann einfach berechnet werden, welche Objekte von einer bestimmten Position aus sichtbar sind.

In einer sehr theoretisch ausgerichteten Arbeit stellen de Nitto Personè et al. (1998) ein Kostenmodell zur Bewertung von Prefetching-Strategien für kontextbezogene Informationssysteme vor und evaluieren das Potenzial unterschiedlicher Strategien bezüglich der Reduzierung der Antwortzeiten dieser Systeme. Sie gehen davon aus, dass es eine Menge der möglichen Kontexte eines Benutzers gibt und dass jede Information nur in einem Teil dieser Kontexte von Interesse ist. Die Untermenge der Kontexte, in denen eine Information von Interesse ist, wird als ihr Gültigkeitsbereich bezeichnet. Die Autoren machen keine Aussage darüber, wie die Gültigkeitsbereiche der Informationen bestimmt werden. Die Evaluierung verschiedener Prefetching-Strategien beschränkt sich auf Strategien, bei denen der Kontext eines Benutzers ausschließ-

lich durch seine geographische Position beschrieben wird. Um eine vollständige mathematische Evaluierung durchführen zu können, wird ein vergleichsweise einfaches Mobilitätsmodell verwendet. Das Gebiet, durch das sich der Benutzer bewegt, ist in einzelne Zellen unterteilt und zu jedem Zeitpunkt bewegt sich der Benutzer entweder mit der jeweils gleichen Wahrscheinlichkeit in eine der Nachbarzellen oder verweilt in der aktuellen Zelle. Nur für den eingeschränkten Fall einer eindimensionalen Bewegung entlang einer Straße werden unterschiedliche Wahrscheinlichkeiten für die beiden möglichen Richtungen betrachtet. Schließlich wird gezeigt, dass sich in dem beschriebenen Szenario die Antwortzeiten des Systems durch Prefetching stark reduzieren lassen, wobei allerdings eine weitere Vergrößerung einer bereits großen Menge vorabübertragener Informationen nur noch sehr geringe Vorteile bringt.

### 2.2.3 Hoarding-Verfahren

Wie bei den Prefetching-Verfahren gibt es auch beim Hoarding eine Reihe von Verfahren, die auf die Vorabübertragung von Dateien spezialisiert sind. Im Andrew File System (AFS) [Howard 1988], das als eines der ersten Dateisysteme einen entkoppelten Betrieb explizit unterstützt hat [Huston & Honeyman 1993], wurde allerdings noch gar keine Vorabübertragung von Dateien durchgeführt. Stattdessen standen im entkoppelten Betrieb ausschließlich die Dateien zur Verfügung, die sich ohnehin im lokalen Cache des Endgeräts des Benutzers befanden.

In [Alonso et al. 1990] wird bei unbedarften Benutzern, die keine expliziten Angaben zu den von ihnen benötigten Dateien machen können, ebenso vorgegangen. Benutzer, die sich sehr gut mit dem System auskennen, haben die Möglichkeit eine Liste der Dateien zu spezifizieren, mit denen sie während des entkoppelten Betriebs arbeiten möchten. Zusätzlich besteht die Möglichkeit, diese Liste halbautomatisch zu erzeugen oder zu ergänzen, in dem alle Dateizugriffe innerhalb eines vom Benutzer angegebenen Zeitraums in die Liste aufgenommen werden.

Das Coda-Dateisystem unterstützt ebenfalls einen entkoppelten Betrieb [Satyanarayanan et al. 1990, Kistler & Satyanarayanan 1992, Satyanarayanan et al. 1993]. Dort werden implizite und explizite Informationsquellen benutzt, um auf der Grundlage von Prioritäten zu entscheiden, welche Dateien für einen entkoppelten Betrieb bereit gestellt werden sollen. Die implizite Information besteht aus einer Zugriffshistorie, wie

sie auch in vielen Caching-Strategien benutzt wird. Die expliziten Informationen werden aus einer so genannten Hoard-Datenbank entnommen, die für jeden Arbeitsplatz die dort bevorzugten Dateien identifiziert. Diese Hoard-Datenbank kann der Benutzer unter zu Hilfenahme von Hoard-Profilen manipulieren. Diese Hoard-Profile sind im Wesentlichen Kommando-Skripte, die der Hoard-Datenbank Dateibezeichner hinzufügen oder auch daraus löschen. Damit ist die Möglichkeit gegeben, die Hoard-Datenbank auf die Verwendung bestimmter Anwendungen oder der Bearbeitung bestimmter Projekte zu optimieren, indem Hoard-Profile aufgerufen werden, die die von der jeweiligen Anwendung beziehungsweise dem Projekt benötigten Dateien beschreiben. Wie bei [Alonso et al. 1990] kann die Erstellung von Hoard-Profilen unterstützt werden, in dem alle Dateizugriffe innerhalb eines vom Benutzer angegebenen Zeitraums aufgezeichnet werden.

Im Hoarding-System Seer [Kuenning & Popek 1997] wird eine vollständig automatisierte Vorabübertragung von Dateien unterstützt. Die Auswahl der Dateien basiert auf so genannten semantischen Entfernungen zwischen Dateien. Dabei ist die semantische Entfernung zwischen dem Öffnen einer Datei *A* und dem Öffnen einer Datei *B* als null definiert, wenn *A* nicht geschlossen wird, bevor *B* geöffnet wird, ansonsten entspricht die semantische Entfernung der Anzahl der zwischenzeitlich geöffneten Dateien. Mit Hilfe dieser semantischen Entfernungen identifiziert ein Gruppierungsalgorithmus Gruppen zusammengehörender Dateien, die als Projekte bezeichnet werden. Schließlich werden für die Vorabübertragung der Dateien die aktiven Projekte bestimmt und die zu den am höchsten priorisierten Projekten gehörenden Dateien werden so weit wie möglich übertragen.

Auch das von Tait et al. (1995) beschriebene Verfahren ist vollständig automatisiert. Allerdings hat der Benutzer hier auf Wunsch die Möglichkeit, die Entscheidung, welche Dateien übertragen werden, zu beeinflussen. Dieses Verfahren basiert auf der Erkennung von so genannten Arbeitsmengen. Um diese Arbeitsmengen zu bestimmen, werden die Dateizugriffe eines Benutzers in einer Protokolldatei aufgezeichnet. Wenn ein entkoppelter Betrieb eingeleitet werden soll, wird diese Protokolldatei analysiert und Baum-Strukturen aufgebaut, die Programmausführungen repräsentieren. Jeder in der Protokolldatei aufgezeichnete Dateizugriff ist ein Knoten in einem Baum. Kinder eines Knotens sind alle Dateien, die vom Vaterknoten ausgeführt oder geöffnet wurden. Unmittelbar von der Benutzeroberfläche aus gestartete Programme werden zu Wurzelknoten eines Baums. Nachdem ein Baum aufgebaut wurde, wird er in einer einfachen

Datenbank gespeichert. In der beschriebenen Implementierung sind in dieser Datenbank für jedes Programm maximal die letzten 15 konstruierten Bäume enthalten. Zu der Arbeitsmenge eines Programms gehören alle Dateien, die in mindestens einem zu diesem Programm gehörenden Bäume enthalten sind. Schließlich kann der Benutzer auswählen, für welche Programme die Arbeitsmengen auf sein Endgerät übertragen werden sollen. Trifft der Benutzer keine Auswahl, werden die seit einem bestimmten, konfigurierbaren Zeitpunkt ausgeführten Programme ausgewählt.

Ein Hoarding-Verfahren zur Vorabübertragung von Karteninformationen wird in [Ye et al. 1998] beschrieben. Das dort betrachtete Fahrzeugnavigationssystem stellt eine Straßenkarte bereit, die in einzelne Segmente unterteilt ist. Jedes dieser Segmente steht in drei verschiedenen Detaillierungsstufen zur Verfügung. Zu Beginn einer Reise teilt der Benutzer den Start- und End-Punkt seiner Reise dem System mit. Daraufhin wird eine Karte im niedrigsten Detaillierungsgrad zum Endgerät des Benutzers übertragen. Auf dieser Karte ist die geplante Reiseroute eingezeichnet. Während der Reise lädt das System selbständig einzelne Segmente der Karte in höheren Detaillierungsstufen nach. Dieses Nachladen erfolgt im Voraus, das heißt bevor der Benutzer in dem durch das Kartensegment beschriebenen Gebiet eintrifft. Es wird ausschließlich an speziellen Kommunikationspunkten, so genannten Infostationen, an denen ein Netzwerkzugang mit hoher Bandbreite möglich ist, durchgeführt. In die Entscheidung, welche Segmente übertragen werden sollen, wird die geplante Reiseroute, die Position des Benutzers sowie seine Bewegungsrichtung miteinbezogen. Die Wahl des Detaillierungsgrads erfolgt entsprechend der Bewegungsgeschwindigkeit des Benutzers. Bei sich schnell bewegenden Benutzern werden nur niedrigere Detaillierungsstufen übertragen, weil davon ausgegangen wird, dass sich der Benutzer nur auf großen Straßen fortbewegt und daher relativ grobe Karten ausreichend sind. Dies hat den Vorteil, dass pro Zeiteinheit mehr Segmente übertragen werden können und damit ein größerer Kartenausschnitt als bei der Verwendung eines höheren Detaillierungsgrads. Bei sich langsam bewegenden Benutzern hingegen wird ein höherer Detaillierungsgrad benutzt, da sie sich durchaus auch auf kleineren Straßen bewegen.

Durch die Spezialisierung des eingesetzten Hoarding-Verfahrens auf die Vorabübertragung von Kartensegmenten ergeben sich zwei Faktoren, die das Verfahren entscheidend vereinfachen. Zum einen sind die zukünftigen Aufenthaltsorte des Benutzers bekannt, weil er zu Beginn Start- und End-Punkt seiner Reise angibt und zum anderen ist genau bekannt, welche Teile der Karte an diesen zukünftigen Aufenthaltsorten

benötigt werden. Entsprechend einfach kann entschieden werden, welche Segmente jeweils zu übertragen sind.

AvantGo<sup>2</sup> ist ein Hoarding-System, das bereits in der Praxis eingesetzt wird. Dabei handelt es sich um einen Dienst, der es seinen Benutzern ermöglicht, in einem entkoppelten Betrieb auf Inhalte des World Wide Web (WWW) zuzugreifen. Der Dienst sowie die benötigte Software, die es in Versionen für PocketPCs, Palm OS und Mobiltelefone gibt, sind für den Endbenutzer kostenfrei. Die jeweiligen Inhalte werden für die Darstellung auf den mobilen Endgeräten optimiert. Der Benutzer wählt aus einem breiten Angebot von Kanälen, wie beispielsweise „Wall Street Journal“, „ExpediaToGo“ oder „Tagesschau mobil“, diejenigen aus, auf die er im entkoppelten Betrieb zugreifen möchte. Durch die in den mobilen Endgeräten verwendete Desktop-Synchronisation kommunizieren die Endgeräte mit den AvantGo-Servern und übertragen die Web-Seiten der ausgewählten Kanäle in ihren lokalen Speicher, so dass diese im entkoppelten Betrieb verfügbar sind.

In der Diplomarbeit [Reichart 2000] wurde untersucht, welche Informationen aus der Beschreibung eines Arbeitsablaufs in Workflow-Systemen gewonnen werden können, um den Benutzer im Voraus mit den Daten zu versorgen, die er zur Durchführung eines oder mehrerer Arbeitsschritte benötigt. Die im Rahmen dieser Arbeit entwickelte Hoarding-Komponente für Workflow-Systeme ermöglicht es dem Benutzer, aus einer Liste von ihm zugewiesenen Aufgaben einen Teil für die Bearbeitung im entkoppelten Betrieb auszuwählen. Die Hoarding-Komponente überträgt daraufhin automatisch alle für die Bearbeitung der ausgewählten Schritte notwendigen Daten. Die Vorhersage der benötigten Daten beruht jeweils auf vorangegangenen Beobachtungen der in jedem Arbeitsschritt referenzierten Daten.

### 2.2.4 Diskussion

Die existierenden Verfahren zur Vorabübertragung von Informationen können in ortsbezogenen Informationssystemen bestenfalls bedingt eingesetzt werden. Die Verfahren, die für Dateisysteme konzipiert sind, können ganz offensichtlich nicht eingesetzt werden, weil sie Beziehungen zwischen Dateien ausnutzen, die es zwischen den Informationsobjekten in einem Informationssystem in dieser Form nicht gibt.

---

<sup>2</sup><http://www.avantgo.com>

Doch auch die für mobile Informationssysteme entwickelten Verfahren, die bereits den Ortsbezug von Informationen ausnutzen, stellen für bestimmte Informationssysteme entwickelte Speziallösungen dar. Insbesondere das Verfahren von Ye et al. (1998) ist ausschließlich zur Verwendung mit der beschriebenen Navigationsanwendung ausgelegt. de Nitto Personè et al. (1998) legen sich zwar nicht auf ein konkretes Informationssystem fest, setzen aber voraus, dass es feste Beziehungen zwischen Orten und Informationen gibt, die a priori bekannt sind. Das in dieser Arbeit beschriebene Verfahren kann hingegen solche Beziehungen selbständig feststellen und sogar Veränderungen dieser Beziehungen über die Zeit hinweg erkennen.

Des Weiteren sind in den existierenden Arbeiten entweder gar keine Vorhersagen der zukünftigen Aufenthaltsorte eines Benutzers notwendig, da sie vorab bekannt sind [Ye et al. 1998] oder sie werden ausschließlich auf der Basis der Beobachtung eines einzelnen Benutzers vorgenommen [Chim et al. 1998, Liu et al. 1996]. Dadurch ist es unmöglich, plötzliche Richtungswechsel vorherzusagen. Im Gegensatz zu diesen Verfahren beobachtet das hier vorgestellte Verfahren alle Benutzer eines Informationssystems und kann so sehr genau feststellen, welche Orte die Benutzer bevorzugen und mit welchen Wahrscheinlichkeiten einzelne Orte besucht werden. Zusätzlich bietet dieses Verfahren Schnittstellen, die es externen Informationsquellen erlauben, Wissen über die zukünftigen Aufenthaltsorte eines Benutzers in die Hoarding-Entscheidung einzubringen.

## Kapitel 3

# Grundlagen

In diesem Kapitel werden grundlegende Sachverhalte erläutert, deren Kenntnis zum weiteren Verständnis dieser Abhandlung notwendig ist. Zur Beschreibung diverser Eigenschaften von ortsbezogenen Informationssystemen werden stochastische Maße benötigt, mit denen die Konzentration einer Häufigkeitsverteilung ausgedrückt werden kann. Diese so genannten Konzentrationsmaße werden in Abschnitt 3.1 eingeführt.

Ferner wird ein Überblick über existierende drahtlose Kommunikationstechnologien gegeben, der notwendig ist, um die Motivation für die in dieser Arbeit verwendete Infostationsinfrastruktur zu verstehen. Diese Infrastruktur wird gemeinsam mit den Annahmen über das zugrunde liegende Informationssystem sowie einigen benötigten externen Diensten im Abschnitt über das Systemmodell beschrieben.

### 3.1 Konzentrationsmaße

Für die Effektivität der Vorabübertragung von ortsbezogenen Informationen ist es entscheidend, wie die Informationszugriffe auf die einzelnen Informationsobjekte und Orte verteilt sind. Von besonderem Interesse ist dabei, wie stark sich diese Informationszugriffe auf die einzelnen Objekte und Orte konzentrieren. Im Folgenden werden die gebräuchlichsten Maße vorgestellt, mit denen solche Konzentrationen beschrieben werden können [Bohley 1996]. Sie werden in Abschnitt 7.1.1 benötigt, um Begriffe wie Ortsbezogenheit exakt definieren zu können. Im Allgemeinen werden Konzentrationsmaße sehr häufig benutzt, um Disparitäten in Verteilungen zu quantifizieren, zum



Beispiel die Konzentration der Bevölkerung in Ballungsgebieten oder die Konzentration eines großen Teils des Gesamteinkommens einer Bevölkerung auf einen kleinen Teil dieser Bevölkerung.

In dieser Arbeit werden die Konzentrationsmaße als Kennzahlen zur Beschreibung von Häufigkeitsverteilungen benutzt. Als Grundlage für eine Häufigkeitsverteilung betrachtet man in der Regel eine so genannte Urliste mit Primärdaten. Dies kann beispielsweise eine Liste mit Bezeichnern von Informationsobjekten eines Informationsraums sein, die nacheinander von den Benutzern referenziert wurden. Eine Häufigkeitsverteilung gibt nun an, wie häufig jede mögliche Ausprägung in der Urliste enthalten ist. Im Beispiel bedeutet das, wie häufig ein Informationsobjekt mit einem bestimmten Bezeichner referenziert wurde.

Im Folgenden bezeichnet  $n$  die Länge der jeweils betrachteten Urliste,  $N$  die Anzahl der möglichen Ausprägungen und  $x_i$  die absolute Häufigkeit, mit der eine bestimmte Ausprägung in der Urliste auftaucht.

### 3.1.1 Konzentrationsrate

Zur Berechnung der Konzentrationsrate müssen die einzelnen Häufigkeiten  $x_i$  absteigend sortiert werden, d.h. es muss gelten  $x_1 \geq x_2 \geq \dots \geq x_{N-1} \geq x_N$ . Danach kann die Konzentrationsrate  $CR_j$  berechnet werden:

$$CR_j = \frac{\sum_{i=1}^j x_i}{\sum_{i=1}^N x_i} = \frac{\sum_{i=1}^j x_i}{n}$$

Die Konzentrationsrate  $CR_j$  gibt also das Verhältnis der Teilsumme der  $j$  größten Häufigkeiten zu der Länge der betrachteten Urliste an.

Möchte man eine Verteilung mit nur einer Kennzahl beschreiben, stellt sich bei der Konzentrationsrate die Frage, welche konkrete Rate  $CR_j$  verwendet werden soll. Ein weiterer Nachteil ist, dass in einer einzelnen Konzentrationsrate  $CR_j$  keine Aussage über die Größenverteilung der einzelnen Werte in der betrachteten Teilsumme und der restlichen Werte enthalten ist. Auf Grund dieser Nachteile wird die Konzentrationsrate in dieser Arbeit nicht weiter verwendet.

### 3.1.2 Herfindahl-Koeffizient

Der Herfindahl-Koeffizient  $HK$  ist definiert als die Summe der quadrierten relativen Häufigkeiten, mit der die einzelnen Ausprägungen auftreten:

$$HK = \sum_{i=1}^N \left( \frac{x_i}{\sum_{j=1}^N x_j} \right)^2 = \sum_{i=1}^N \left( \frac{x_i}{n} \right)^2$$

Tritt überhaupt nur eine der möglichen Ausprägungen in der Urliste auf, so nimmt der Herfindahl-Koeffizient seinen Maximalwert 1 an. Sind die Häufigkeiten der einzelnen Ausprägungen gleichverteilt, dann wird der Herfindahl-Koeffizient minimal. Der Minimalwert hängt von der Anzahl  $N$  der möglichen Ausprägungen ab und beträgt  $\frac{1}{N}$ .

### 3.1.3 Gini-Koeffizient

Die Grundlage für den Gini-Koeffizienten bildet die Lorenzkurve. Mit ihr kann die Konzentration einer Verteilung graphisch dargestellt werden. Um eine Lorenzkurve zu konstruieren, müssen zunächst die Häufigkeiten  $x_i$  aufsteigend sortiert werden, d.h. es muss gelten  $x_1 \leq x_2 \leq \dots \leq x_{N-1} \leq x_N$ . Die Lorenzkurve kann dann als Polygonzug durch die Punkte  $(0, 0), (u_1, v_1), (u_2, v_2), \dots, (u_{N-1}, v_{N-1}), (u_N, v_N)$  konstruiert werden. Die Koordinaten der Punkte werden folgendermaßen berechnet:

$$u_j = \frac{j}{N} \quad \text{und} \quad v_j = \sum_{i=1}^j \frac{x_i}{\sum_{k=1}^N x_k} = \sum_{i=1}^j \frac{x_i}{n}$$

Man trägt also auf der Abszissen-Achse den Anteil der Ausprägungen ab und auf der Ordinaten-Achse die zugehörigen kumulierten relativen Häufigkeiten. Wenn die betrachteten Häufigkeiten gleichverteilt sind, dann stimmt die Lorenzkurve mit der Hauptdiagonalen überein. Je größer die Fläche zwischen der Lorenzkurve und der Hauptdiagonalen ist, desto höher ist der Grad der Konzentration.

Als Beispiel ist in Abbildung 3.1 die Lorenzkurve zur Häufigkeitsverteilung aus Tabelle 3.1 angegeben. Zum Vergleich ist die Hauptdiagonale mit eingezeichnet. Die Häufigkeitsverteilung gibt die Häufigkeiten wider, mit denen die einzelnen Informationsobjekte eines zehn Informationsobjekte umfassenden Informationsraums referenziert werden, wenn für die Verteilung der Zugriffswahrscheinlichkeiten eine Zipf-Verteilung angenommen wird (siehe Abschnitt 6.2.2).

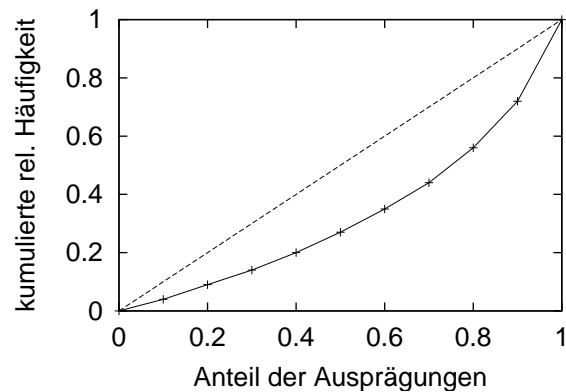


Abbildung 3.1: Lorenzkurve.

Informations- objekt	Anzahl der Zugriffe	Anteil der Ausprägungen	kumulierte rel. Häufigkeit
1	4	0.1	0.04
2	5	0.2	0.09
3	5	0.3	0.14
4	6	0.4	0.20
5	7	0.5	0.27
6	8	0.6	0.35
7	9	0.7	0.44
8	12	0.8	0.56
9	16	0.9	0.72
10	28	1.0	1.0

Tabelle 3.1: Verteilung der Zugriffshäufigkeiten auf einzelne Informationsobjekte.

Der Gini-Koeffizient  $GK$  ist definiert als das Verhältnis der Fläche zwischen der Hauptdiagonalen und der Lorenzkurve zur Fläche zwischen der Hauptdiagonalen und der Abszissen-Achse. Wie in Anhang B gezeigt wird, gilt für den Gini-Koeffizienten:

$$GK = \frac{2 \cdot \sum_{i=1}^N i \cdot x_i}{N \cdot \sum_{i=1}^N x_i} - \frac{N+1}{N} = \frac{2 \cdot \sum_{i=1}^N i \cdot x_i}{N \cdot n} - \frac{N+1}{N}$$

Bei einer egalitären Verteilung nimmt der Gini-Koeffizient seinen Minimalwert null an. Der Maximalwert  $\frac{N-1}{N}$  des Gini-Koeffizienten wird bei maximaler Konzentration erreicht, d.h. wenn ausschließlich eine der möglichen Ausprägungen in der Urliste auftritt.

### 3.1.4 Vergleich Herfindahl- versus Gini-Koeffizient

Der Unterschied zwischen den beiden Koeffizienten ist, dass in den Gini-Koeffizient ausschließlich die relative Konzentration der Häufigkeiten eingeht, während der Herfindahl-Koeffizient sowohl eine absolute als auch eine relative Konzentration reflektiert.

Dabei spricht man von absoluter Konzentration, wenn sich alle in der Urliste enthaltenen Ausprägungen auf eine kleine Zahl möglicher Ausprägungen konzentrieren. Die Verteilung der Häufigkeiten über die wenigen möglichen Ausprägungen muss dann nicht zwangsläufig ungleich sein.

Von relativer Konzentration spricht man hingegen, wenn sich ein großer Anteil der in der Urliste vorkommenden Ausprägungen auf einen kleinen Anteil der möglichen Ausprägungen verteilt. Dies hat zwangsläufig eine Ungleichverteilung der Häufigkeiten zur Folge.

Verteilung 1		Verteilung 2	
Informations- objekt	Anzahl der Zugriffe	Informations- objekt	Anzahl der Zugriffe
1	50	1	50
2	50	2	50
3	50	3	50
		4	50
		5	5
		6	1
		7	1
		8	1
		9	1
		10	1
$GK = 0$		$GK \approx 0.5619$	
$HK = \frac{1}{3}$		$HK \approx 0.2274$	

**Tabelle 3.2:** Häufigkeitsverteilungen mit unterschiedlichen absoluten und relativen Konzentrationen.

Um den Unterschied zwischen den beiden Koeffizienten zu verdeutlichen, sind in Tabelle 3.2 zwei Verteilungen dargestellt, bei denen sich die zwei Koeffizienten entgegengesetzt verhalten. Während der Herfindahl-Koeffizient bei der ersten Verteilung höher ist als bei der zweiten, ist es beim Gini-Koeffizient genau umgekehrt. Die Erklärung ist,

dass die absolute Konzentration der ersten Verteilung, bei der sich alle Zugriffe auf nur drei Informationsobjekte beziehen, höher ist als in der zweiten Verteilung. Hingegen liegt in der ersten Verteilung keine relative Konzentration vor, da auf alle Informationsobjekte gleich oft zugegriffen wird. In der zweiten Verteilung konzentrieren sich dann die meisten Zugriffe auf die ersten fünf Informationsobjekte.

## 3.2 Drahtlose Kommunikationstechnologien

Um einen mobilen Informationszugriff zu ermöglichen, muss ein drahtloses Kommunikationsmedium eingesetzt werden. Inzwischen existiert ein breites Spektrum drahtloser Kommunikationstechnologien, das von speziellen Technologien für den Nahbereich über regionale Netze bis hin zu Mobilfunknetzen reicht.

Im Folgenden wird untersucht inwiefern sich diese Technologien für den Einsatz in mobilen ortsbezogenen Informationssystemen, wie sie in Abschnitt 1.1 beschrieben wurden, eignen. Wichtig dabei ist, dass die verfügbare Bandbreite auch für die Übertragung multimedialer Daten ausreichend ist. Um eine möglichst hohe Akzeptanz bei den potenziellen Benutzern zu erreichen, sollten die Kommunikationskosten für den Benutzer möglichst gering sein. Eine weitere Anforderung ist, dass auch die Installationskosten der Infrastruktur für den Informationsanbieter nicht zu hoch sind.

### 3.2.1 Mobilfunknetze

Mobilfunknetze bieten eine große Reichweite bei einer vergleichsweise niedrigen Datenrate. Das Abdeckungsgebiet eines Mobilfunknetzes umfasst typischerweise das Gebiet eines Staates. Durch bilaterale Abkommen, so genannte Roaming-Abkommen, der Netzbetreiber unterschiedlicher Staaten ist inzwischen auch eine Sprach- und Datenkommunikation über Staatsgrenzen hinweg möglich. Die maximale Größe einer einzelnen Funkzelle liegt im Bereich weniger Kilometer.

In Europa hat sich für Mobilfunknetze der GSM-Standard (Global System for Mobile Communications) [Eberspächer & Vögel 1998] durchgesetzt. Inzwischen gewinnt er auch weltweit an Verbreitung. GSM bietet typischerweise eine Datenrate von 9.6 kbit/s, was für die Übertragung multimedialer Daten nicht ausreichend ist.

Bei HSCSD (High Speed Circuit Switched Data) [Walke 2000a] handelt es sich um eine Erweiterung des GSM-Standards, mit der durch die Bündelung von bis zu 4 Kanälen à 14.4 kbit/s Datenraten bis zu 57.6 kbit/s erreicht werden können. Wie in GSM werden auch bei HSCSD die Daten leitungsvermittelt übertragen, daher werden auch Ressourcen verbraucht, wenn gar keine Daten übertragen werden und der Benutzer muss für die kontinuierlich bestehende Verbindung bezahlen.

Dieser Nachteil wird in der paketorientierten Erweiterung GPRS (General Packet Radio System) [Witt 2000] des GSM-Standards überwunden. Mit GPRS können theoretisch Daten mit bis zu 171.2 kbit/s übertragen werden, allerdings bieten die Netzbetreiber momentan nur Datenraten zwischen 40 kbit/s und 115.2 kbit/s.

Mit EDGE (Enhanced Data Rates for Global Evolution) ist noch eine weitere Verbesserung der Datenübertragung in GSM-Netzen geplant. Hier werden durch die Änderung des Modulationstyps und die Bündelung von bis zu acht Kanälen Datenraten von bis zu 384 kbit/s möglich.

Das kommende Mobilfunksystem UMTS (Universal Mobile Telecommunications System) [Walke 2000b] wird mehrere Modi für die Datenübertragung unterstützen. Im Allgemeinen wird die Datenrate bei 384 kbit/s liegen. In späteren Ausbaustufen sind Datenraten von bis zu 1.92 Mbit/s über kurze Strecken hinweg (ohne Zellenwechsel) geplant. Die Broschüre [Ericsson GmbH 2000] bietet einen leichtverständlichen Überblick über die hier vorgestellten Mobilfunktechnologien.

Wird ein Mobilfunknetz für den Zugriff auf ein Informationssystem eingesetzt, entstehen für den Benutzer vergleichsweise hohe Kommunikationskosten. Bei den leitungsvermittelten Verfahren werden diese Kosten nach der Verbindungszeit abgerechnet, während bei den paketorientierten Verfahren nach übertragenem Datenvolumen abgerechnet wird.

#### 3.2.2 Drahtlose Regionalnetzwerke

Das einzige drahtlose Regionalnetzwerk, das einen höheren Bekanntheitsgrad erreicht hat, ist das Ricochet Netzwerk der Firma Metricom<sup>1</sup>. Momentan gibt es in 13 amerikanischen Großstädten sowie 15 Flughäfen Ricochet Installationen, die eine Datenübertragungsrate von 128 kbit/s ermöglichen. Zusätzlich gibt es in zwei weiteren Großstädten Installationen, deren Datenübertragungsrate auf 28.8 kbit/s beschränkt ist.

---

<sup>1</sup><http://www.metricom.com>

Metricom-Kunden müssen eine monatliche Grundgebühr bezahlen, die zwischen US\$ 60 und US\$ 80 liegt. Darin ist ein beliebiges Datentransfervolumen enthalten. Die Preise für die benötigten Modems liegen bei ca. US\$ 100. Außer den unter Umständen geringeren Kosten für den Benutzer bietet das Ricochet Netzwerk keine Vorteile gegenüber den modernen Weitverkehrstechnologien und ist zudem in Europa nicht verfügbar.

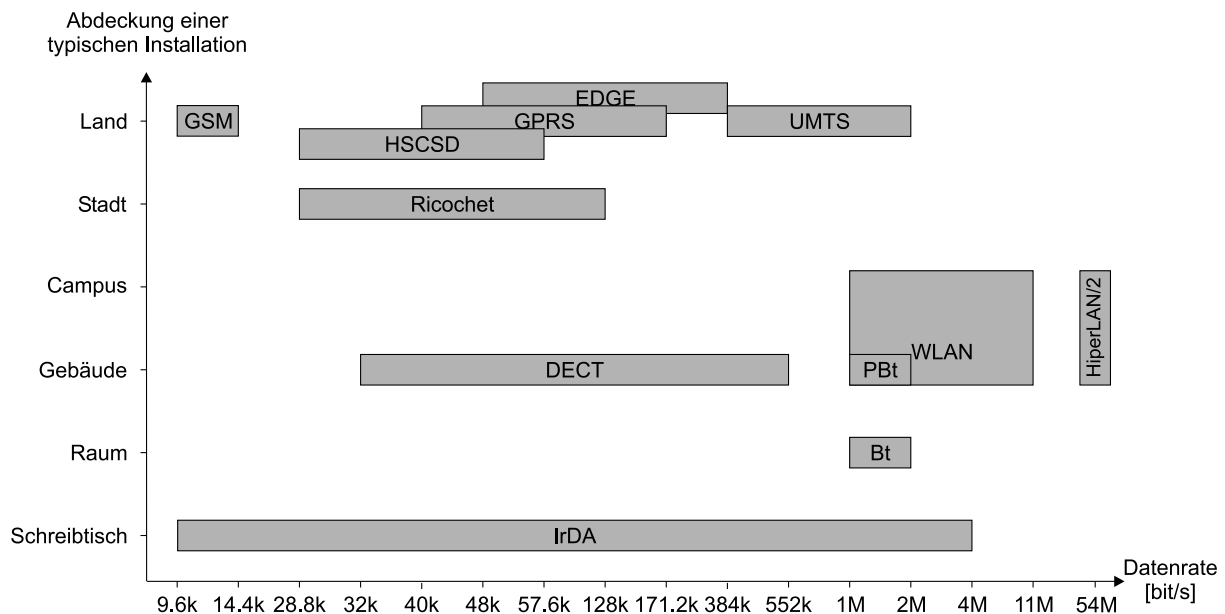
### 3.2.3 Lokale Kommunikationstechnologien

Die gegenwärtig verbreiteste Technologie zur drahtlosen Kommunikation im Nahbereich sind die drahtlosen Nahverkehrsnetzwerke nach dem IEEE 802.11b Standard [Gergeleit 2001]. Mit dieser Technologie ist es möglich, einen drahtlosen Zugang zu einem vorhandenen verdrahtetem Nahverkehrsnetzwerk herzustellen. Dazu werden die mobilen Endgeräte üblicherweise mit einer Funkkarte ausgerüstet. Auf der Seite des verdrahteten Netzwerks muss ein Zugangspunkt installiert werden.

Bei einer guten Funkverbindung zwischen Endgerät und Zugangspunkt können Daten mit bis zu 11 Mbit/s übertragen werden. Die maximale Reichweite der Funkverbindung liegt in der Regel bei 200 m. Falls Abschattungen vorliegen, beispielsweise in Gebäuden, kann sie auch deutlich geringer sein.

Bluetooth [Miller & Bisdikian 2000] ist eine weitere Technologie zur drahtlosen Kommunikation im Nahbereich mit großem Potenzial. Sie wurde für die Kommunikation zwischen Endgeräten entworfen. Ein besonderes Augenmerk lag bei der Entwicklung von Bluetooth auf seiner Eignung für den Massenmarkt und damit auf der Möglichkeit, es in günstigen Chipsätzen zu realisieren. Je nach der verwendeten Variante liegt die Reichweite bei Bluetooth-Verbindungen zwischen 10 m und 100 m, wobei Datenübertragungsraten von 1-2 Mbit/s erzielt werden.

Neben diesen beiden wichtigsten Standards für die drahtlose Nahbereichskommunikation existieren noch eine Reihe weiterer Standards, wie zum Beispiel die Standards der IrDA (Infrared Data Association), DECT (Digital Enhanced Cordless Telecommunications) und HomeRF. Prinzipiell sind sie mit den hier vorgestellten Technologien, WLAN und Bluetooth, bezüglich Reichweite und Datenrate vergleichbar. Der vom ETSI (European Telecommunications Standards Institute) vorgeschlagene Funknetzstandard HiperLAN/2 bietet höhere Datenraten von 54 Mbit/sec, ist verbindungsorientiert und kann verschiedene Dienstgüten garantieren.



**Abbildung 3.2:** Verschiedene Funktechnologien im Vergleich (Bt – Bluetooth, PBt – PowerBluetooth).

#### 3.2.4 Zusammenfassung

In Abbildung 3.2 sind für die einzelnen Technologien die Datenraten sowie die Abdeckungsgebiete typischer Installationen zusammengefasst. Die verfügbaren Technologien stellen stets einen Kompromiss zwischen der Größe des Abdeckungsgebiets und der maximalen Datenrate dar. Mit den zukünftigen Technologien EDGE und UMTS werden erstmals höhere Datenraten bei einer globalen Abdeckung möglich sein. Allerdings muss aufgrund der hohen Investitionskosten der Mobilfunkbetreiber für diese neuen Technologien davon ausgegangen werden, dass die Nutzungskosten dieser neuen Dienste ähnlich hoch oder höher als die für HSCSD oder GPRS sein werden.

### 3.3 Systemmodell

Nachfolgend wird das Systemmodell beschrieben, von dem bei der Entwicklung des in dieser Arbeit vorgestellten Verfahrens ausgegangen wurde.



### 3.3.1 Informationssystem

In den betrachteten Informationssystemen greifen die Benutzer über ihre mobilen Endgeräte auf eine Menge von Informationen, den so genannten Informationsraum, zu. Es wird davon ausgegangen, dass dieser Informationsraum aus diskreten, unveränderlichen Informationsobjekten besteht. Jedes Informationsobjekt hat einen global eindeutigen Bezeichner über den darauf zugegriffen werden kann. Wie diese Bezeichner genau aussehen, ist für das Vorabübertragungsverfahren irrelevant, möglich sind zum Beispiel URLs (Uniform Resource Locators) [Berners-Lee et al. 1994].

Sollten die Informationsobjekte veränderlich sein, dann würde das hier vorgestellte Verfahren trotzdem prinzipiell funktionieren. Allerdings müsste dann damit gerechnet werden, dass der Benutzer veraltete Informationen erhält, wenn er auf vorabübertragene Informationen zugreift, die seit dem Zeitpunkt der Vorabübertragung verändert wurden.

Des Weiteren ist das vorgeschlagene Verfahren auf einen ortsabhängigen Informationszugriff optimiert, der wie folgt definiert ist:

**DEFINITION 3.1 (ORTSABHÄNGIG, ORTSBEZOGEN)**

Der Zugriff auf ein Informationssystem heißt *ortsabhängig*, wenn die Wahrscheinlichkeit, mit der ein Benutzer auf ein bestimmtes Informationsobjekt zugreift, vom Aufenthaltsort des Benutzers abhängt. Die Informationsobjekte werden dann als *ortsbezogene Informationsobjekte* bezeichnet, das Informationssystem selbst als *ortsbezogenes Informationssystem*.

Die Ortsabhängigkeit bzw. die Ortsbezogenheit ist dabei keine binäre Eigenschaft, vielmehr kann ein Informationszugriff mehr oder weniger ortsabhängig sein und entsprechend ein Informationssystem mehr oder weniger ortsbezogen. In Abschnitt 7.1 wird daher ein Maß definiert, mit dem der Grad der Ortsbezogenheit erfasst werden kann. Je höher die Ortsbezogenheit des Informationssystems ist, desto effektiver wird im Allgemeinen das hier vorgeschlagene Verfahren arbeiten.

Entsprechend dem Grad ihrer Ortsbezogenheit kann man zwei Klassen von ortsbezogenen Informationssystemen unterscheiden:

1. **Inhärent ortsbezogene Informationssysteme:** In dieser Klasse von Systemen ist jedes Informationsobjekt einem bestimmten geographischen Ort fest zugeordnet

und die Benutzer greifen ausschließlich dann auf ein bestimmtes Informationsobjekt zu, wenn sie sich an dem Ort befinden, dem es zugeordnet ist. Ein Beispiel ist die in [Ye et al. 1998] beschriebene Kartenanwendung, in der die Benutzer an jedem Ort ausschließlich auf den Kartenausschnitt zugreifen, der zu ihrer Umgebung gehört. Ein weiteres Beispiel ist eine Augmented Reality Anwendung, in der der Benutzer auf ein Objekt der realen Welt zeigen muss, um darüber Informationen zu erhalten. Hier muss sich der Benutzer in der Nähe des Objekts befinden, um die Information abrufen zu können.

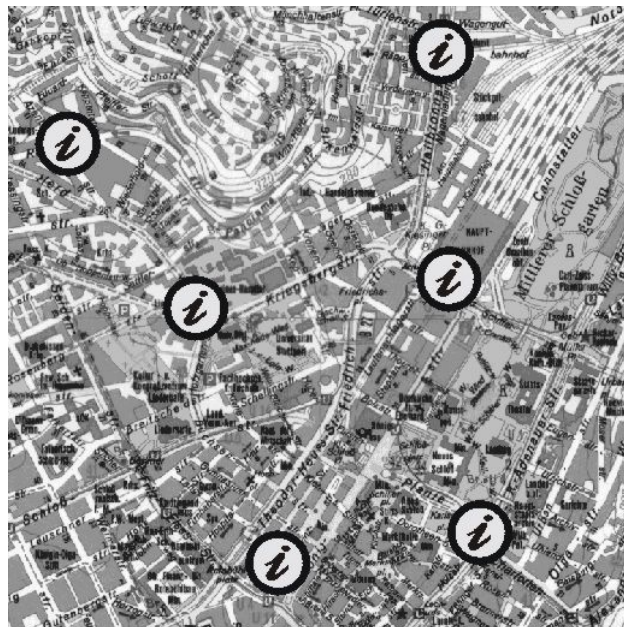
2. Nicht inhärent ortsbezogene Informationssysteme: In dieser Art von Systemen greifen die Benutzer an jedem Ort potenziell auf alle Objekte im Informationsraum zu. Ein Beispiel für ein solches System ist ein Browser-basierter Touristenführer. Auch wenn der Informationszugriff hier eine hohe Ortsabhängigkeit aufweisen wird, so kann der Benutzer doch zumindest theoretisch auch Informationen über eine Attraktion *A* abrufen, wenn er selbst vor der Attraktion *B* steht.

#### 3.3.2 Kommunikationsinfrastruktur

Wie in Abschnitt 3.2 erläutert wurde, bieten die heute verfügbaren drahtlosen Kommunikationstechnologien entweder nur eine geringe Reichweite oder eine vergleichsweise niedrige maximale Datenrate. Mit den neueren und zukünftigen Standards zur mobilen Datenübertragung (GPRS, UMTS) werden zwar erstmals hohe Datenraten bei einer globalen Abdeckung möglich, aber dennoch wird sich an der Diskrepanz zwischen lokalen Kommunikationstechnologien und Mobilfunknetzen qualitativ nichts ändern, da drahtlose lokale Netze auch zukünftig stets eine höhere Datenrate bieten werden und günstiger sein werden als Mobilfunknetze. Ferner wird mit steigenden Datenraten auch der Bedarf an Bandbreite weiter wachsen, da neue Anwendungen und die Unterstützung umfangreicherer Informationstypen, wie beispielsweise Video, ermöglicht werden.

Um die Vorteile der drahtlosen lokalen Kommunikationstechnologien (hohe Datenraten) mit denen der Mobilfunknetze (großes Abdeckungsgebiet) kombinieren zu können, wird in dieser Arbeit eine sogenannte Infostationsinfrastruktur verwendet. Bei den Infostationen handelt es sich um drahtlose LANs, die über ein Gebiet verteilt sind,

das ansonsten gar nicht von einer drahtlosen Kommunikationsinfrastruktur oder nur durch ein Mobilfunknetz abgedeckt wird. Durch die Infostationen entstehen also Teilgebiete, in denen eine hohe Bandbreite verfügbar ist, in einem Gebiet, in dem ansonsten gar keine Kommunikation möglich wäre oder nur mit der maximalen Datenrate des Mobilfunknetzes kommuniziert werden könnte. Abbildung 3.3 zeigt wie solche Infostationen über die Innenstadt von Stuttgart verteilt werden könnten. Die Infostationen sollten möglichst an Orten platziert werden, an denen die Benutzer des unterstützten Informationssystems häufig vorbeikommen, damit sie keine Umwege in Kauf nehmen müssen, um eine Infostation zu erreichen.



**Abbildung 3.3:** Infostationen in der Stuttgarter Innenstadt. Die Kreise markieren die von den drahtlosen LANs abgedeckten Gebiete.

Am naheliegendsten ist es sicherlich, die Infostationen als drahtlose LANs nach dem IEEE 802.11b Standard zu realisieren. Von dieser Umsetzung des Infostation-Konzepts wird in dieser Arbeit, wie auch in [Ye et al. 1998] und [Badrinath et al. 1996], ausgegangen. In [Ye et al. 1998] ist ein Szenario beschrieben, in dem die Infostationen entlang von Straßen platziert sind, um Autofahrer mit aktuellen Umgebungskarten zu versorgen. In [Badrinath et al. 1996] werden drei Kategorien von Infostationen unterschieden, je nach dem ob sich die Benutzer länger im Abdeckungsgebiet einer Infostation aufhalten, nur durch das Abdeckungsgebiet hindurch gehen oder ob sie hindurch fah-

ren.

Für das in dieser Arbeit vorgestellte Verfahren ist die Abdeckung des gesamten Gebiets durch ein Mobilfunknetz nicht notwendig. Wenn allerdings ein Mobilfunknetz verfügbar ist, kann das in bestimmten Situationen diverse Vorteile haben, beispielsweise wenn auf ein Informationsobjekt zugegriffen wird, das nicht vorab auf das Endgerät übertragen wurde. Kommuniziert ein Endgerät in einer Infostationsinfrastruktur über eine Mobilfunkverbindung wird dies in dieser Arbeit als schwach gekoppelter Betrieb bezeichnet, um den Unterschied zum entkoppelten Betrieb zu verdeutlichen, in dem gar keine Kommunikation möglich ist.

Prinzipiell kann das Infostation-Konzept zu beliebigen Szenarien verallgemeinert werden, in denen es Gebiete mit unterschiedlich guter Netzwerkanbindung gibt. So wird zum Beispiel in [Goodman et al. 1997] vorgeschlagen die Gebiete im Kern von Mobilfunkzellen, in denen oftmals höhere Datenraten als an den Rändern erreicht werden können, als Infostationen zu betrachten und primär dort Daten zu übertragen.

#### **3.3.3 Endgeräte**

Bezüglich der verwendeten mobilen Endgeräte gibt es keine prinzipiellen Einschränkungen von Seiten des Verfahrens. Sie müssen lediglich genügend Speicherplatz haben, um vorab übertragene Informationsobjekte lokal speichern zu können. Es hat sich gezeigt (siehe Kapitel 7), dass bereits sehr gute Ergebnisse erzielt werden, wenn vier MBytes Speicherplatz für das Verfahren zur Verfügung stehen. Zusätzlich müssen die Endgeräte über Einrichtungen zur drahtlosen Kommunikation verfügen.

#### **3.3.4 Externe Dienste**

Die Endgeräte müssen in der Lage sein, ihren Ort zu bestimmen und zu erkennen, wenn sie sich im Abdeckungsbereich einer Infostation befinden. Aus Sicht des in dieser Arbeit vorgestellten Verfahrens ist es unerheblich, ob dies lokal auf dem mobilen Endgerät oder unter Verwendung entfernter Dienste geschieht. In beiden Fällen nutzt das Verfahren aus seiner Sicht externe Dienste, die ihm eine geforderte Funktionalität bieten. Neben dem Ortsinformationsdienst und dem Ereignisdienst benötigt das Vorabübertragungsverfahren zusätzlich einen Verzeichnisdienst.

#### **Ortsinformationsdienst**

Der Ortsinformationsdienst muss dem Vorabübertragungsverfahren jederzeit die aktuelle geographische Position des mobilen Endgeräts mitteilen können. Ein sehr einfacher Dienst könnte beispielsweise einen am Endgerät befestigten GPS-Sensor (Global Position System) abfragen [Hofmann-Wellenhof et al. 2001, Hässler 2000]. Bei GPS handelt es sich um ein Satelliten-gestütztes System, das es ermöglicht die geographische Position bis auf wenige Meter genau zu bestimmen. Der Nachteil dabei ist allerdings, dass die Funktion von GPS-Sensoren von der Sichtbarkeit der Satelliten abhängt. Daher kann die Funktion durch hohe Gebäude, Berge oder Ähnlichem eingeschränkt werden. Innerhalb von Gebäuden kann GPS prinzipiell nicht eingesetzt werden.

Einen Ausweg können hier funkzellenbasierte Dienste, wie sie in Mobilfunksystemen eingesetzt werden, bieten. Diese sind zwar innerhalb von Gebäuden verfügbar, besitzen dafür aber eine vergleichsweise geringe Genauigkeit. Eine sehr gute Lösung stellen generische Ortsinformationsdienste dar [Leonhardi & Rothermel 2001], die mehrere Positionierungssysteme integrieren und dadurch sowohl innerhalb von Gebäuden als auch im Außenbereich verlässliche Ortsinformationen mit hinreichender Genauigkeit liefern können.

#### **Ereignisdienst**

Der Ereignisdienst muss dem Vorabübertragungsverfahren mitteilen können, wenn sich die Netzwerkanbindung des mobilen Endgeräts ändert, insbesondere wenn das Abdeckungsgebiet einer Infostation betreten wird. Auch hier kann ein einfacher lokaler Dienst genutzt werden, der beispielsweise durch periodisches Nachfragen feststellt, wann ein drahtloses LAN verfügbar ist oder es können komplexere, entfernte Ereignisdienste eingesetzt werden, wie sie zum Beispiel gegenwärtig im neXus-Projekt entwickelt werden.

#### **Verzeichnisdienst**

Wie in Abschnitt 4.2 noch erläutert wird, benötigt das Vorabübertragungsverfahren zusätzlich einen Verzeichnisdienst. Das Verfahren nutzt mehrere Server, die jeweils einen

bestimmten geographischen Zuständigkeitsbereich besitzen. Dieser Zuständigkeitsbereich wird dem Verzeichnisdienst bei der Initialisierung eines Servers mitgeteilt.

Der Verzeichnisdienst muss dann in der Lage sein, dem Vorabübertragungsverfahren zu einer gegebenen geographischen Position alle Server mitzuteilen, in deren Zuständigkeitsbereich sich die gegebene Position befindet.

## Kapitel 4

# Vorabübertragung ortsbezogener Informationen

Zunächst wird in diesem Kapitel die Problemstellung bei der Vorabübertragung von Informationen im Detail erläutert. Anschließend wird ein Überblick über die grundsätzliche Funktionsweise des im Rahmen von dieser Arbeit entwickelten Verfahrens gegeben, bevor es schließlich in zwei verschiedenen Varianten ausführlich beschrieben wird. Die erste Variante, das so genannte Basisverfahren, berücksichtigt den Ortsbezug von Informationen nur auf einem groben Niveau, während die zweite Variante, das erweiterte Verfahren, diesen Ortsbezug auf einem beliebig feinen Niveau ausnutzen kann.

### 4.1 Problemstellung

Wie bereits erwähnt, ist die Grundidee bei der Vorabübertragung von Informationen, die Informationen auf das mobile Endgerät zu übertragen, solange eine gute Netzwerkanbindung verfügbar ist, um sie später, wenn sie in einem schwach gekoppelten oder einem entkoppelten Betrieb benötigt werden, bereits auf dem Endgerät vorliegen zu haben.

Auf Grund von Platz- und/oder Zeit-Restriktionen können in der Regel nicht alle in einem Informationssystem verfügbaren Informationen im Voraus auf das Endgerät eines

Benutzers übertragen werden. Bei sehr kleinen Endgeräten, wie persönlichen digitalen Assistenten, spielen die Platzrestriktionen die Hauptrolle, da diese Geräte im Allgemeinen über wenig Speicherplatz verfügen. Typisch sind heute vier bis acht MBytes. Wenn man davon ausgeht, dass vier MBytes für die Speicherung vorab übertragener Informationen zur Verfügung stehen, entspricht das ca. 80 WWW-Seiten.

Bei größeren Endgeräten, wie Laptops, steht in der Regel sehr viel mehr Speicherplatz zur Verfügung, daher treten dort die Beschränkungen durch den Speicherplatz in den Hintergrund und die Zahl der vorab übertragbaren Informationen wird durch die maximal zur Verfügung stehende Übertragungszeit bestimmt. Insbesondere wenn sich die Benutzer nicht zu Fuß sondern in Fahrzeugen fortbewegen, ist die Übertragungszeit stark eingeschränkt. Von einem Fahrzeug, das sich mit  $50 \frac{km}{h}$  bewegt, wird eine Infostation, deren Abdeckungsgebiet einen Durchmesser von 150 Metern hat, in nur 10.8 Sekunden durchquert. Diese Zeit reicht selbst bei wenig ausgelasteten WLANs (Durchsatz pro Benutzer 1.5 - 3 Mbps) lediglich zur Übertragung von 2.025 - 4.05 Mbyte. Dies entspricht ca. 40 - 80 WWW-Seiten.

Da nicht alle Informationen übertragen werden können, muss bei der Vorabübertragung entschieden werden, welche Informationen im Voraus übertragen werden sollen. In der englischsprachigen Literatur wird das Problem, diese Hoarding-Entscheidung zu treffen, auch als „hoarding problem“ bezeichnet [Tait et al. 1995]. Für die Hoarding-Entscheidung ist es wichtig, möglichst genau vorherzusagen, welche Informationen ein Benutzer zukünftig benötigen wird, da durch die Übertragung nicht benötigter Informationen knappe Ressourcen, wie der auf dem Endgerät verfügbare Speicherplatz, die Übertragungszeit oder Bandbreite des WLANs, verschwendet werden. Ziel der Hoarding-Entscheidung ist es, mit den verfügbaren Ressourcen eine möglichst hohe Trefferrate im entkoppelten bzw. schwach gekoppelten Betrieb zu erzielen.

##### DEFINITION 4.1 (TREFFERRATE)

Die innerhalb eines Zeitintervalls  $T$  erzielte Trefferrate ist der Quotient aus der Anzahl der Anfragen, die während des Intervalls  $T$  gestellt wurden und mit vorab übertragenen Informationsobjekten beantwortet werden konnten, und der Gesamtzahl der während des Intervalls  $T$  gestellten Anfragen.



## 4.2 Überblick über die Verfahren

Die im Systemmodell (siehe Abschnitt 3.3) vorausgesetzte Ortsabhängigkeit des Informationszugriffs wird im vorgestellten Verfahren dahingehend ausgenutzt, dass das System Informationen darüber sammelt, welche Informationsobjekte in welchen Gebieten von einem durchschnittlichen Benutzer oder bestimmten Benutzergruppen bevorzugt abgerufen werden. Diese Information wird bei der Hoarding-Entscheidung benutzt, um vom aktuellen und dem vorhergesagten Aufenthaltsort eines Benutzers auf die von ihm zukünftig benötigten Informationsobjekte zu schließen. Im erweiterten Verfahren wird zusätzlich Wissen über die bevorzugten Aufenthaltsorte der Benutzer eingesetzt, um die zukünftigen Aufenthaltsorte eines Benutzers präziser vorhersagen zu können.

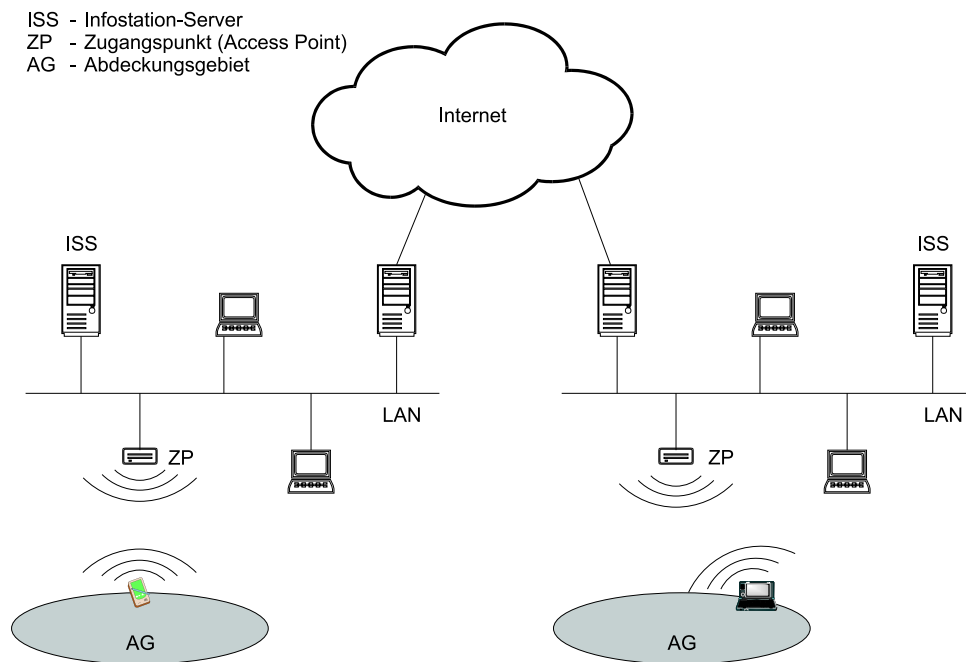
Wie in Abschnitt 3.3.2 bereits erwähnt wurde, beruht das hier vorgestellte Verfahren auf der Verwendung von Infostationen. Die Vorabübertragung von Informationen wird ausschließlich in den von den Infostationen abgedeckten Gebieten durchgeführt. Wenn zusätzlich zu den Infostationen ein Mobilfunknetz verfügbar ist, kann dieses eingesetzt werden, um Informationen, die zwar nicht vorab übertragen wurden, aber dennoch vom Benutzer referenziert werden, bei Bedarf zu übertragen.

Jeder Infostation ist genau ein Software-Prozess zugeordnet, der die Vorabübertragungsvorgänge an der jeweiligen Infostation koordiniert. Dieser so genannte Infostation-Server sollte auf einem Knoten ausgeführt werden, der über eine möglichst schnelle Kommunikationsmöglichkeit mit dem Zugangspunkt des drahtlosen LANs der Infostation verfügt, damit die Kommunikation mit den mobilen Endgeräten im Sendebereich der Infostation nicht unnötig verlangsamt wird. Idealerweise sollte sich der Knoten des Infostation-Servers im gleichen Festnetz-LAN befinden wie der Zugangspunkt des Funknetzes (siehe Abbildung 4.1). Im Folgenden wird der Ausdruck Infostation für die Gesamtheit aus drahtlosem LAN und dem zugehörigen Server-Prozess verwendet.

Jeder Infostation ist ein geographisch begrenztes Gebiet, ihr so genanntes Hoarding-Gebiet, zugeordnet.

### DEFINITION 4.2 (HOARDING-GEBIET EINER INFOSTATION)

Das Hoarding-Gebiet einer Infostation ist ein geographisch begrenztes Gebiet, das ihr fest zugeordnet ist.



**Abbildung 4.1:** Typische Infostationsinfrastruktur.

Beim Hoarding-Gebiet handelt es sich um eine logische Verwaltungseinheit, die nichts mit dem Abdeckungsgebiet des Funknetzes einer Infostation zu tun hat.

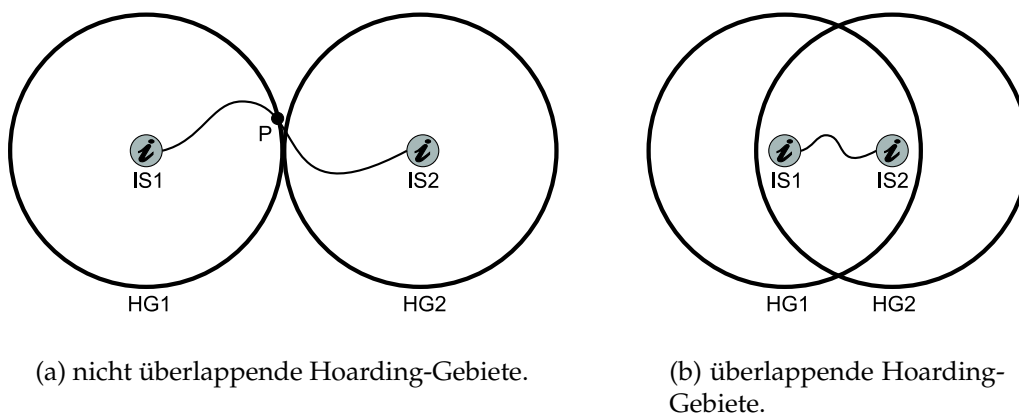
### DEFINITION 4.3 (ABDECKUNGSGEBIET EINER INFOSTATION)

Das Gebiet, in dem das Funknetz einer Infostation verfügbar ist, wird als das Abdeckungsgebiet der Infostation bezeichnet.

Zu jeder Infostation gehört genau ein Hoarding-Gebiet und umgekehrt ist jedes Hoarding-Gebiet genau einer Infostation zugeordnet. Jede Infostation verwaltet Wissen darüber, welche Informationsobjekte die Benutzer des unterstützten Informationssystems abrufen, während sie sich durch ihr Hoarding-Gebiet bewegen. Kommt nun ein Benutzer an einer Infostation an, wird dieses Wissen über die Popularität der Informationsobjekte benutzt, um vorherzusagen auf welche Informationsobjekte der Benutzer während seiner weiteren Bewegung durch das Hoarding-Gebiet am wahrscheinlichsten zugreifen wird. Schließlich werden nur die am wahrscheinlichsten benötigten Informationsobjekte auf sein Endgerät übertragen. Jede Infostation versorgt den Benutzer also ausschließlich mit den Informationsobjekten, die er während seiner Bewegung durch das Hoarding-Gebiet der Infostation wahrscheinlich benötigen wird.

Das Hoarding-Gebiet einer Infostation sollte stets größer sein als ihr Abdeckungsgebiet, da ansonsten der Benutzer ausschließlich mit Informationsobjekten versorgt wird, die er benötigt solange er sich innerhalb des Abdeckungsgebiets aufhält. Auf diese Objekte könnte er jedoch auch direkt über das drahtlose LAN der Infostation zugreifen.

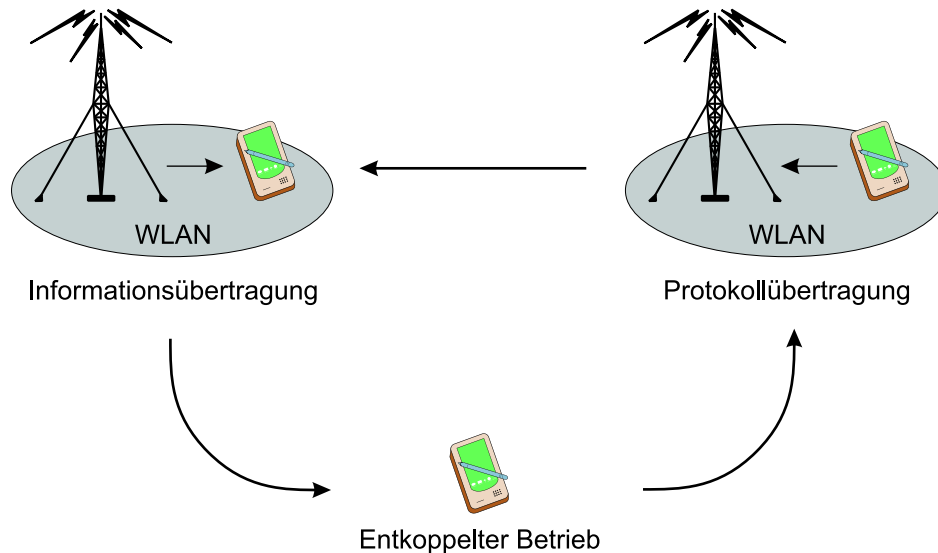
Um den Benutzer möglichst lückenlos mit Informationen versorgen zu können, sollten sich die Hoarding-Gebiete benachbarter Infostationen überlappen. Dies wird am Beispiel aus Abbildung 4.2, die den disjunkten und den überlappenden Fall darstellt, deutlich. Nehmen wir an, ein Benutzer bewegt sich auf dem eingezeichneten Pfad von der Infostation *IS1* zur Infostation *IS2*. Falls sich die Hoarding-Gebiete nicht überlappen (Abbildung 4.2(a)), wird er am Punkt *P* das Hoarding-Gebiet der Infostation *IS1* verlassen. Da er jedoch zu diesem Zeitpunkt die Infostation *IS2* noch nicht erreicht hat, hatte er noch keine Gelegenheit die Informationsobjekte, die im Hoarding-Gebiet *HG2* benötigt werden, auf sein Endgerät zu übertragen. Folglich muss er den Rest des Weges bis zur Infostation *IS2* zurücklegen, ohne mit den entsprechenden Informationen versorgt worden zu sein. Im Fall von überlappenden Hoarding-Gebieten (Abbildung 4.2(b)) hingegen können bereits an der Infostation *IS1* alle Informationsobjekte, die der Benutzer auf seinem Weg bis zur Infostation *IS2* benötigen wird, auf sein Endgerät übertragen werden. Bei der anschließenden Vorabübertragung an der Infostation *IS2* wird darauf geachtet, dass Informationsobjekte, die bereits auf dem Endgerät des Benutzer vorhanden sind, nicht nochmals übertragen werden.



**Abbildung 4.2:** Infostationen (IS) und ihre Hoarding-Gebiete (HG).

Das Vorabübertragungsverfahren wird für jeden einzelnen Benutzer in einer zyklischen Art und Weise durchgeführt (siehe Abbildung 4.3). Jeder Zyklus des Verfah-

rens besteht aus den drei Phasen Informationsübertragung, Entkoppelter Betrieb und Protokollübertragung.



**Abbildung 4.3:** Zyklus des Vorübertragungsverfahrens.

### 4.2.1 Informationsübertragung

Während der Informationsübertragungsphase befindet sich der Benutzer innerhalb des Abdeckungsgebiets einer Infostation, das heißt sein mobiles Endgerät ist mit dem Infostation-Server über ein drahtloses LAN verbunden. Basierend auf dem Wissen über die innerhalb des Hoarding-Gebiets bevorzugten Informationsobjekte, bestimmt der Infostation-Server die Objekte, auf die der Benutzer am wahrscheinlichsten zugreifen wird, und überträgt sie auf das Endgerät des Benutzers. Die Menge der übertragenen Informationsobjekte wird dabei entweder durch die Aufenthaltszeit des Benutzers im Abdeckungsgebiet der Infostation oder den auf dem Endgerät verfügbaren Speicherplatz begrenzt. Da sich die Verwaltung des Wissens über die bevorzugten Informationsobjekte sowie die Bestimmung der zu übertragenden Objekte in den beiden Varianten des Verfahrens unterscheiden, werden sie in den Abschnitten 4.3 und 4.4 getrennt beschrieben.

Wenn sich häufig zur gleichen Zeit mehrere Benutzer im Abdeckungsgebiet einer Infostation befinden, kann die Vorübertragung eventuell durch den Einsatz von Rundsendeverfahren, wie sie in Abschnitt 2.1.4 beschrieben wurden, optimiert werden.

### 4.2.2 Entkoppelter Betrieb

Die zweite Phase des Zyklus beginnt, sobald der Benutzer das Abdeckungsgebiet der Infostation verlassen hat und dauert bis er das Abdeckungsgebiet der nächsten Infostation erreicht. Während der Phase des entkoppelten Betriebs werden die Informationsanfragen des Benutzers mit den Informationen beantwortet, die lokal auf dem Endgerät des Benutzers vorhanden sind, weil sie in einer der vorangehenden Informationsübertragungsphasen dorthin übertragen wurden. Fordert der Benutzer eine Information an, die nicht vorab übertragen wurde, tritt ein Vorabübertragungsfehler (Hoard Miss) auf. Ist in einem solchen Fall ein Mobilfunknetz verfügbar (schwach gekoppelter Betrieb) und ist die Information dem Benutzer so wichtig, dass er bereit ist, die für die Beschaffung notwendigen Verbindungskosten zu tragen, kann die Information über eine Mobilfunkverbindung vom Informationssystem angefordert werden. Andernfalls muss die Informationsanfrage unbeantwortet bleiben.

Unabhängig davon, ob eine Informationsanfrage beantwortet werden kann oder nicht, wird sie immer in einer Protokolldatei zusammen mit dem Ort, an dem sie gestellt wurde, aufgezeichnet. Die Bestimmung des Ortes erfolgt über einen Ortsinformationssdienst (siehe Abschnitt 3.3.4), der natürlich auch im entkoppelten Betrieb verlässliche Informationen liefern muss. Die Protokolldatei stellt eine Liste dar, deren Einträge die Form  $\langle i, (x, y) \rangle$  haben, wobei  $i$  der global eindeutige Bezeichner des referenzierten Informationsobjekts ist und  $(x, y)$  die Koordinaten sind, an denen die Informationsanfrage gestellt wurde.

### 4.2.3 Protokollübertragung

Die dritte Phase des Zyklus beginnt schließlich, wenn der Benutzer das Abdeckungsgebiet der nächsten Infostation erreicht. In dieser Phase wird die Protokolldatei, die während des entkoppelten beziehungsweise des schwach gekoppelten Betriebs erstellt wurde, zum Infostation-Server übertragen.

Der Infostation-Server verteilt die in der Protokolldatei enthaltenen Informationen an alle Infostationen für die sie relevant sind. Um festzustellen welche Infostationen dies sind, benutzt der Infostation-Server den in Abschnitt 3.3.4 erwähnten Verzeichnisdienst. Selbstverständlich muss jede Infostation bei ihrer Initialisierung ihr Hoarding-Gebiet zuvor bei diesem Verzeichnisdienst angemeldet haben. Die Informationen aus

den Protokolldateien werden von den Infostationen verwendet, um ihr Wissen über die innerhalb ihres Hoarding-Gebiets bevorzugten Informationsobjekte ständig zu aktualisieren.

Schließlich wird die Protokolldatei auf dem mobilen Endgerät gelöscht und der nächste Zyklus beginnt mit der nachfolgenden Informationsübertragungsphase. Da sich die Bearbeitung der Protokolldateien in den beiden Varianten des Verfahrens ebenfalls unterscheidet, wird sie auch getrennt in den Abschnitten 4.3 und 4.4 behandelt.

Obwohl die Protokolldateien anonym zu den Infostationen übertragen werden, müssen die Infostationen die Geräteadressen der mobilen Endgeräte kennen, um mit ihnen kommunizieren zu können. Dadurch ist es theoretisch möglich, Informationen über das Verhalten eines Benutzers mit der Adresse seines Endgeräts in Verbindung zu bringen. Das kann zu Datenschutzproblemen führen, falls der Benutzer dem Betreiber der Infostationsinfrastruktur nicht traut (siehe Abschnitt 4.5.1).

### 4.3 Basisverfahren

Zu Beginn jeder Informationsübertragungsphase muss die Infostation, an der sich der Benutzer gegenwärtig befindet, entscheiden, welche Informationsobjekte zum mobilen Endgerät des Benutzers übertragen werden sollen. Wie bereits erwähnt wurde, wird für diese Entscheidung Wissen über die Popularität der Informationsobjekte ausgenutzt. Im Basisverfahren besteht dieses Wissen ausschließlich aus einer Tabelle, die die Zugriffswahrscheinlichkeiten der innerhalb des Hoarding-Gebiets am häufigsten referenzierten Informationsobjekte beinhaltet. Diese Tabelle wird als Zugriffswahrscheinlichkeitstabelle (ZWT) bezeichnet. Die Einträge in dieser Tabelle haben die Form  $\langle i, a(i) \rangle$ , wobei  $i$  wiederum der global eindeutige Bezeichner eines Informationsobjekts ist und  $a(i)$  die Zugriffswahrscheinlichkeit auf das Objekt  $i$  angibt.

Im Basisverfahren muss jede Infostation im Wesentlichen drei Aufgaben erledigen: die Auswertung der empfangenen Protokolldateien, die periodische Aktualisierung der Zugriffswahrscheinlichkeitstabelle und die Auswahl der zu übertragenden Informationsobjekte. Die periodische Aktualisierung erfolgt nach Ablauf eines Aktualisierungsintervalls  $\Delta$  dessen Länge frei konfigurierbar ist und durch die Anzahl der bearbeiteten Protokolldateien bestimmt wird.

### 4.3.1 Auswertung der Protokolldateien

Jedes Mal wenn eine Infostation eine Protokolldatei empfängt, wird diese unmittelbar ausgewertet. Wurde die Protokolldatei von einer anderen Infostation empfangen, wird sie durch die empfangende Infostation nicht mehr weiter verteilt. Wurde sie hingegen von einem mobilen Endgerät in der Protokollübertragungsphase empfangen, müssen die in der Protokolldatei enthaltenen Informationen von der empfangenden Infostation an alle betroffenen Infostationen weiter verteilt werden.

Die Verteilung dieser Informationen erfolgt gemäß Algorithmus 4.1. Für jeden Eintrag  $\langle i, (x, y) \rangle$  in der empfangenen Protokolldatei  $\mathcal{P}$  wird beim Verzeichnisdienst eine Liste  $\mathcal{L}$  der Infostationen angefordert innerhalb deren Hoarding-Gebiete sich die Koordinaten  $(x, y)$  befinden. Für jede in dieser Liste enthaltene Infostation  $s$  wird, sofern nicht schon in einer vorherigen Iteration geschehen, eine neue Protokolldatei  $\mathcal{P}_s$  angelegt. Schließlich wird der Eintrag  $\langle i, (x, y) \rangle$  in die Protokolldatei  $\mathcal{P}_s$  jeder in der Liste  $\mathcal{L}$  enthaltenen Infostation  $s$  eingetragen.

```

method verteileProtokolldatei(in:Protokolldatei  $\mathcal{P}$  out:Protokolldatei)
  for all  $\langle i, (x, y) \rangle \in \mathcal{P}$  do
     $\mathcal{L} := \text{verzeichnisdienst.getInfostations}(x, y);$ 
    for all  $s \in \mathcal{L}$  do
5:      if  $\mathcal{P}_s$  existiert nicht then
        Neue Protokolldatei  $\mathcal{P}_s$  anlegen;
      end if
       $\mathcal{P}_s := \mathcal{P}_s + \langle i, (x, y) \rangle;$ 
    end for
10: end for
    for all  $\mathcal{P}_s$  do
      if  $s \neq$  verteilende Infostation then
        Sende  $\mathcal{P}_s$  zur Infostation  $s$ ;
      else
15:       $\text{result} := \mathcal{P}_s;$ 
      end if
    end for
  return result;

```

**Algorithmus 4.1:** Methode einer Infostation zur Verteilung der in einer Protokolldatei  $\mathcal{P}$  enthaltenen Informationen.

```

method bearbeiteProtokolldatei(in:Protokolldatei  $\mathcal{P}$ )
    bearbeiteteProtokolldateien := bearbeiteteProtokolldateien + 1;
    if  $\mathcal{P}$  direkt von mobilem Endgerät empfangen then
         $\mathcal{P}$  := this.verteileProtokolldatei( $\mathcal{P}$ );
5: end if
    for all  $\langle i, (x, y) \rangle \in \mathcal{P}$  do
        anzahlAnfragen := anzahlAnfragen + 1;
        if Zähler  $r_{\Delta}(i)$  existiert in Vektor  $c$  then
             $r_{\Delta}(i)$  :=  $r_{\Delta}(i) + 1$ ;
10: else
            Neuen Zähler  $r_{\Delta}(i)$  an Vektor  $c$  anhängen;
             $r_{\Delta}(i)$  := 1;
        end if
    end for
15: if bearbeiteteProtokolldateien =  $|\Delta|$  then
    bearbeiteteProtokolldateien := 0;
    tabelle.aktualisiere();
    anzahlAnfragen := 0;
    Lösche Inhalt von Vektor  $c$ ;
20: end if

```

**Algorithmus 4.2:** Methode einer Infostation zur Bearbeitung der Einträge einer Protokolldatei  $\mathcal{P}$ .

Der Verteilungsalgorithmus liefert als Ergebnis eine verkürzte Protokolldatei zurück, die ausschließlich Anfragen enthält, die im Hoarding-Gebiet der verteilenden Infostation aufgetreten sind. Für die weitere Verarbeitung wird nur noch diese verkürzte Protokolldatei verwendet. Unabhängig davon ob die Protokolldatei ursprünglich direkt von einem mobilen Endgerät oder von einer anderen Infostation empfangen wurde, wird also im Folgenden nur noch eine Protokolldatei betrachtet, die ausschließlich für die Infostation relevante Einträge enthält. Dabei ist die Relevanz eines Eintrags für eine Infostation wie folgt definiert:

**DEFINITION 4.4 (RELEVANZ EINES PROTOKOLLDATTEI-EINTRAGS)**

Ein Eintrag  $\langle i, (x, y) \rangle$  in einer Protokolldatei  $\mathcal{P}$  ist für eine Infostation  $\mathcal{I}$  relevant, genau dann wenn die Koordinaten  $(x, y)$  im Hoarding-Gebiet von  $\mathcal{I}$  liegen.

Nach der Verteilung der Protokolldatei beginnt die eigentliche Bearbeitung der inzwischen verkürzten Protokolldatei. Für jeden Eintrag  $\langle i, (x, y) \rangle$  in der verkürzten Pro-



tokolldatei wird ein Zähler *anzahlAnfragen* für die Gesamtzahl der beobachteten Informationszugriffe um eins inkrementiert, falls vorhanden, wird auch der Zähler  $r_{\Delta}(i)$  für die Anzahl der Zugriffe auf das Informationsobjekt  $i$  um eins erhöht. Ist noch kein Zähler  $r_{\Delta}(i)$  vorhanden, wird ein neuer angelegt, mit dem Wert eins initialisiert und an den Vektor  $c$ , der alle Zähler  $r_{\Delta}(i)$  enthält, angehängt. Zusätzlich wird die Anzahl der bearbeiteten Protokolldateien im Zähler *bearbeiteteProtokolldateien* festgehalten. Sobald eine konfigurierbare Anzahl  $\Delta$  an Protokolldateien bearbeitet wurde, wird die Aktualisierung der Zugriffswahrscheinlichkeitstabelle angestoßen. In Algorithmus 4.2 ist zusammenfassend der gesamte Algorithmus zur Bearbeitung einer Protokolldatei angegeben.

### 4.3.2 Aktualisierung der Zugriffswahrscheinlichkeitstabelle

Die Aktualisierung der Zugriffswahrscheinlichkeitstabelle wird nach Ablauf des Aktualisierungsintervalls  $\Delta$  angestoßen (siehe Algorithmus 4.3). Dann ist im Zähler *anzahlAnfragen* die Anzahl der seit der letzten Aktualisierung beobachteten Anfragen enthalten. Ferner ist im Vektor  $c$  für jedes Informationsobjekt  $i$ , auf das seit der letzten Aktualisierung mindestens einmal innerhalb des Hoarding-Gebiets der Infostation zugegriffen wurde, ein Zähler  $r_{\Delta}(i)$  enthalten. Der Zähler  $r_{\Delta}(i)$  gibt an, wie oft seit der letzten Aktualisierung im Hoarding-Gebiet auf das Objekt  $i$  zugegriffen wurde.

Mit diesen Angaben kann die Infostation für jedes Informationsobjekt  $i$ , auf das seit der letzten Aktualisierung innerhalb des Hoarding-Gebiets mindestens einmal zugegriffen wurde, die Zugriffswahrscheinlichkeit  $a_{\Delta}(i)$  auf das Objekt  $i$  während des Aktualisierungsintervalls  $\Delta$  berechnen:

$$a_{\Delta}(i) = \frac{r_{\Delta}(i)}{\text{anzahlAnfragen}}$$

Ist für das Informationsobjekt  $i$  bereits ein Eintrag  $\langle i, a(i) \rangle$  in der Zugriffswahrscheinlichkeitstabelle enthalten, wird der Wert  $a(i)$  auf einen neuen Wert  $a'(i)$  gesetzt. Der neue Wert wird entsprechend der folgenden Formel berechnet:

$$a'(i) = \gamma \cdot a(i) + (1 - \gamma) \cdot a_{\Delta}(i)$$

Dabei ist  $0 \leq \gamma \leq 1$  ein Parameter, der bestimmt, wie stark Anfragen aus vorherigen

```

method aktualisiere()
for all  $r_{\Delta}(i) \in c$  do
     $a_{\Delta}(i) := \frac{r_{\Delta}(i)}{\text{anzahlAnfragen}};$ 
    if Eintrag  $\langle i, a(i) \rangle$  existiert in Zugriffswahrscheinlichkeitstabelle then
5:      $a(i) := \gamma \cdot a(i) + (1 - \gamma) \cdot a_{\Delta}(i);$ 
    else
        neuen Eintrag  $\langle i, a(i) \rangle$  anlegen;
         $a(i) := (1 - \gamma) \cdot a_{\Delta}(i);$ 
    end if
10: end for

if Tabellengröße > Schwellwert then
    Lösche (Tabellengröße – Schwellwert) Einträge mit den kleinsten  $a(i)$ ;
end if

```

**Algorithmus 4.3:** Methode der Zugriffswahrscheinlichkeitstabelle für ihre Aktualisierung.

Aktualisierungsintervallen im neuen Wert berücksichtigt werden sollen. Der optimale Wert des Parameters  $\gamma$  hängt von der Dynamik im Zugriffsverhalten der Benutzer ab. Wenn sich das Zugriffsverhalten schnell ändert, muss  $\gamma$  möglichst klein gewählt werden, damit sich die Änderungen schnell in der Zugriffswahrscheinlichkeitstabelle widerspiegeln. Bei einem sehr kleinen Wert von  $\gamma$  geht allerdings das Wissen aus vorherigen Aktualisierungsintervallen sehr schnell verloren und vorübergehende Änderungen im Zugriffsverhalten können die Zugriffswahrscheinlichkeitstabelle völlig verändern. Daher ist es bei relativ statischem Zugriffsverhalten besser, den Wert von  $\gamma$  größer zu wählen.

Wenn es ein Informationsobjekt  $i$  gibt, für das noch kein Eintrag in der Tabelle existiert, aber  $a_{\Delta}(i) > 1$  gilt, muss ein neuer Eintrag  $\langle i, a(i) \rangle$  zur Zugriffswahrscheinlichkeitstabelle hinzugefügt werden.  $a(i)$  wird dabei mit dem Wert  $(1 - \gamma) \cdot a_{\Delta}(i)$  initialisiert. Das heißt, dass für die Zugriffswahrscheinlichkeit in vorherigen Intervallen null angenommen wird.

Sobald die Größe der Zugriffswahrscheinlichkeitstabelle einen frei wählbaren Schwellwert überschreitet, werden die Einträge mit den kleinsten Zugriffswahrscheinlichkeiten aus der Tabelle gelöscht. Deshalb ist die obige Annahme einer Zugriffswahrscheinlichkeit von null für in der Tabelle nicht enthaltene Objekte nicht immer völlig exakt.

Schließlich wird noch der Zähler *anzahlAnfragen* auf null zurückgesetzt und die Zäh-

ler im Vektor  $c$  gelöscht.

### 4.3.3 Hoarding-Entscheidung

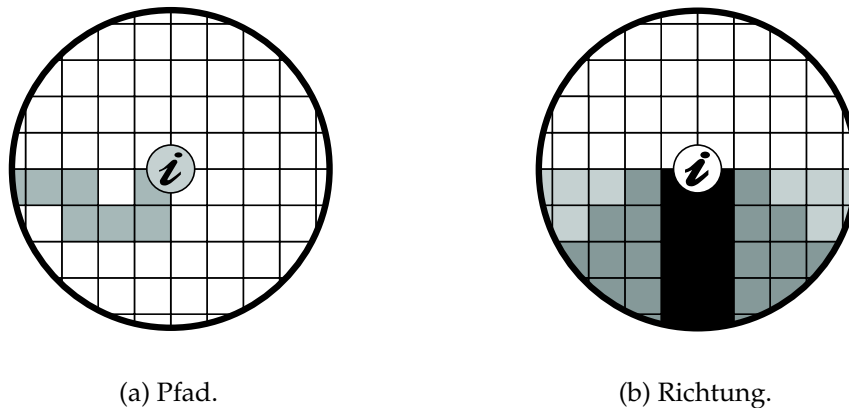
Auf Grundlage der Zugriffswahrscheinlichkeitstabelle kann die Hoarding-Entscheidung sehr einfach getroffen werden. Wenn ein Benutzer Informationen an einer Infostation vorab übertragen möchte und auf Grund der Randbedingungen, wie z.B. die maximale Transferzeit und der verfügbare Speicherplatz,  $m$  Informationsobjekte übertragen werden können, werden stets die  $m$  Objekte, die gemäß der aktuellen Zugriffswahrscheinlichkeitstabelle die höchsten Zugriffswahrscheinlichkeiten besitzen, auf das mobile Endgerät des Benutzers übertragen.

### 4.3.4 Diskussion

Der Nachteil des Basisverfahrens ist, dass die Ortsabhängigkeit der Informationszugriffe nur auf Ebene der Hoarding-Gebiete betrachtet wird. Daher weiß die Infostation nur, welche Informationsobjekte im gesamten Hoarding-Gebiet am häufigsten benötigt werden. Sie hat aber kein Wissen darüber, welche Objekte in kleineren Teilgebieten bevorzugt werden. Entsprechend kann Wissen über die zukünftige Bewegung eines Benutzers, das oftmals einfach verfügbar ist, nicht genutzt werden.

Ein Beispiel ist in Abbildung 4.4(a) angegeben. Wenn man annimmt, dass der Benutzer, nachdem er die Infostation verlassen hat, nur die schattierten Quadrate besucht, und die Infostation wüsste, welche Informationsobjekte in diesen Quadraten bevorzugt werden, würde es ausreichen, nur diese Informationsobjekte zu übertragen. Alle Objekte, die ausschließlich in den übrigen Quadraten bevorzugt werden, könnten von der Vorabübertragung ausgeschlossen werden, wodurch mehr Informationen übertragen werden könnten, die in den besuchten Quadraten benötigt werden.

Ein weiteres Beispiel ist, wenn eine Komponente zur Bewegungsvorhersage feststellt, dass sich der Benutzer vorzugsweise in eine bestimmte Richtung bewegt. In Abbildung 4.4(b) wird angenommen, dass sich der Benutzer von der Infostation weg nach Süden bewegt, beispielsweise weil er sich zuvor schon für längere Zeit in südlicher Richtung bewegt hat. Dann macht es Sinn, für die Gebiete nördlich der Infostation keine Informationen zu übertragen. Für die dunkelgrauen Gebiete sollten nur wenige



**Abbildung 4.4:** Wissen über die Bewegung eines Benutzers nach dem Verlassen einer Infostation.

Informationsobjekte übertragen werden, für die hellgrauen einige Objekte mehr und die meisten für die weißen Gebiete.

Bewegt sich ein Benutzer zu einem bestimmten Zielort, dann ist es unwahrscheinlich, dass er den gleichen Ort mehrmals besucht, es sei denn er hat sich verirrt. Daher wäre es in einem solchen Fall sinnvoll, alle Informationen, die ausschließlich in den bereits besuchten Gebieten bevorzugt werden, von der Vorabübertragung auszuschließen und damit mehr Ressourcen für die übrigen Informationsobjekte zur Verfügung zu haben.

Um solche und ähnliche externe Informationsquellen über die zukünftigen Aufenthaltsorte eines Benutzers ausnutzen zu können, wurde das Basisverfahren, wie im nächsten Abschnitt beschrieben, erweitert.

## 4.4 Erweitertes Verfahren

Das Ziel des erweiterten Verfahrens ist es, externe Informationsquellen, wie sie in den vorangehenden Beispielen beschrieben wurden, zu nutzen, um die Vorhersage der von einem Benutzer wahrscheinlich benötigten Informationen zu präzisieren. Da das Wissen, das von externen Informationsquellen geliefert wird, nicht durch das Hoarding-Verfahren selbst gewonnen wird, wird es als externes Wissen bezeichnet. Solches externes Wissen kann entweder von den Applikationen oder vom Benutzer selbst durch eine geeignete Benutzeroberfläche bereitgestellt werden.

**DEFINITION 4.5 (EXTERNES WISSEN)**

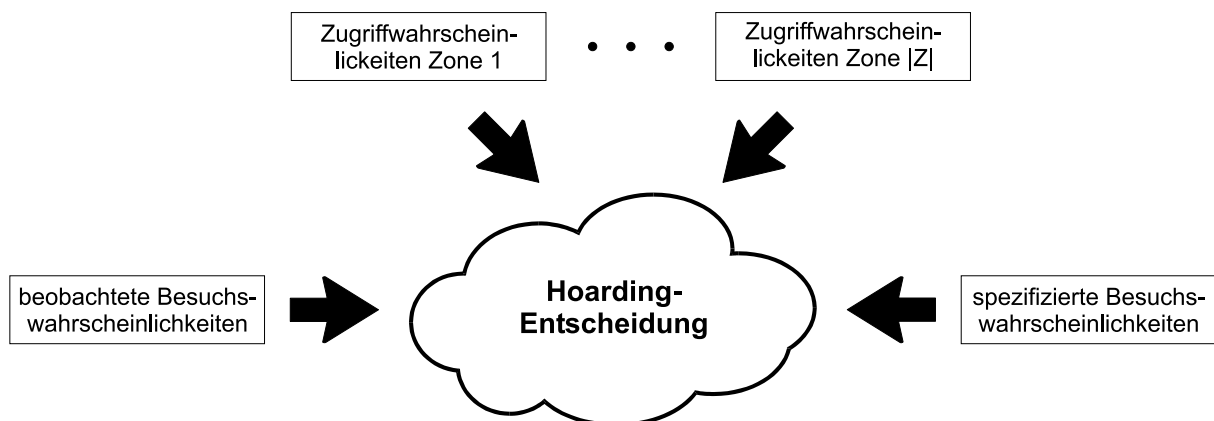
Wissen über die zukünftigen Aufenthaltsorte eines Benutzers, das nicht durch das Hoarding-Verfahren selbst gewonnen wird, sondern von Applikationen oder dem Benutzer bereitgestellt wird, wird als externes Wissen bezeichnet.

**4.4.1 Grundidee**

Die Grundidee des erweiterten Verfahrens ist es, jedes Hoarding-Gebiet in einzelne Zonen, die sich nicht überlappen, zu unterteilen. Diese Zonen können im einfachsten Fall gleich große Quadrate sein, wie in den Beispielen aus Abbildung 4.4. Sie können aber auch die Geometrie von Objekten aus der realen Welt, wie zum Beispiel Straßen oder Gebäuden, widerspiegeln.

**DEFINITION 4.6 (ZONE)**

Eine Zone ist ein geographisch begrenztes Teilgebiet eines Hoarding-Gebiets. Zwei verschiedene Zonen überlappen sich nicht.



**Abbildung 4.5:** Informationsquellen für die Hoarding-Entscheidung.

Für die Hoarding-Entscheidung werden im erweiterten Verfahren drei verschiedene Informationsquellen herangezogen (siehe Abbildung 4.5):

1. Jede Infostation beobachtet, getrennt für jede Zone in ihrem Hoarding-Gebiet, die jeweiligen lokalen Zugriffswahrscheinlichkeiten auf die einzelnen Informationsobjekte innerhalb der Zone.

2. Jede Infostation beobachtet für jede Zone  $z$  in ihrem Hoarding-Gebiet die Wahrscheinlichkeit mit der ein Benutzer, der das Hoarding-Gebiet der Infostation besucht, auch die Zone  $z$  besucht.
3. Der Benutzer und/oder die Anwendungen können ihr Wissen über die zukünftigen Aufenthaltsorte des Benutzers in die Hoarding-Entscheidung einbringen.

Im erweiterten Verfahren verwaltet jede Infostation eine separate Zugriffswahrscheinlichkeitstabelle für jede Zone innerhalb ihres Hoarding-Gebiets und eine so genannte Besuchswahrscheinlichkeitskarte (BWK), in der die Wahrscheinlichkeiten gespeichert sind, mit der ein durchschnittlicher Benutzer die einzelnen Zonen besucht. Das externe Wissen wird vom Benutzer oder den Anwendungen in externen Besuchswahrscheinlichkeitskarten spezifiziert.

Die Methode einer Infostation zur Verteilung einer empfangenen Protokolldatei (siehe Algorithmus 4.1) kann gegenüber dem Basisverfahren unverändert bleiben, während die eigentliche Bearbeitung einer Protokolldatei und die Aktualisierung der Zugriffswahrscheinlichkeitstabellen, wie nachfolgend beschrieben, modifiziert werden müssen.

### 4.4.2 Bearbeitung einer Protokolldatei

Wie im Basisverfahren wird bei der Verteilung der Protokolldateien sichergestellt, dass bei der Bearbeitung einer Protokolldatei durch die Infostationen, ausschließlich für die jeweilige Infostation relevante Einträge in der Protokolldatei enthalten sind.

Der prinzipielle Unterschied zur Bearbeitung einer Protokolldatei im Basisverfahren ist, dass nun die Anzahl der Zugriffe auf jedes Informationsobjekt  $i$  getrennt für jede Zone  $z$  im Hoarding-Gebiet der Infostation gezählt und in einem eigenen Zähler  $r_{\Delta}(i, z)$  festgehalten werden. Die Zähler  $r_{\Delta}(*, z)$  jeder Zone  $z$  werden in einem eigenen Vektor  $c(z)$  verwaltet. Zusätzlich wird noch die Anzahl der Besuche einer Zone  $z$  in einem Zähler  $n_{\Delta}(z)$  festgehalten. Es wird nur der jeweils erste Zugriff auf ein Informationsobjekt  $i$  bei einem Besuch einer Zone  $z$  gezählt, weil die Zähler letztlich dazu benutzt werden, die Wahrscheinlichkeit zu berechnen, mit der ein Besucher, der das Hoarding-Gebiet der Infostation besucht, auf ein bestimmtes Informationsobjekt zugreift. Die Häufigkeit, mit der er das tut, spielt dabei keine Rolle.

```

method bearbeiteProtokolldatei(in:Protokolldatei  $\mathcal{P}$ )
  bearbeiteteProtokolldateien := bearbeiteteProtokolldateien + 1;
  if  $\mathcal{P}$  direkt von mobilem Endgerät empfangen then
     $\mathcal{P}$  := this.verteileProtokolldatei( $\mathcal{P}$ );
5: end if
  for all  $\langle i, (x, y) \rangle \in \mathcal{P}$  do
     $z$  := this.bestimmeZone( $x, y$ );
    if bislang gab es in  $\mathcal{P}$  noch keinen Besuch von  $z$  then
       $n_{\Delta}(z)$  :=  $n_{\Delta}(z) + 1$ ;
10:   if  $i$  wurde während des aktuellen Besuchs von  $z$  noch nicht referenziert then
     if Zähler  $r_{\Delta}(i, z)$  existiert in Vektor  $c(z)$  then
        $r_{\Delta}(i, z)$  :=  $r_{\Delta}(i, z) + 1$ ;
     else
       Neuen Zähler  $r_{\Delta}(i, z)$  an Vektor  $c(z)$  anhängen;
15:    $r_{\Delta}(i, z)$  := 1;
     end if
   end if
  end if
  end for
20: if bearbeiteteProtokolldateien =  $|\Delta|$  then
  bearbeiteteProtokolldateien := 0;
  karte.aktualisiere();
  for all Zonen  $z$  in Hoarding-Gebiet do
    tabelle[ $z$ ].aktualisiere();
25: end for
  for all Zonen  $z$  in Hoarding-Gebiet do
     $n_{\Delta}(z)$  := 0;
    Lösche Inhalt von Vektor  $c(z)$ ;
  end for
30: end if

```

**Algorithmus 4.4:** Modifizierte Methode einer Infostation zur Bearbeitung einer Protokolldatei  $\mathcal{P}$ .

Wiederholte Besuche der gleichen Zone  $z$  werden nur in Ausnahmefällen auftreten, beispielsweise wenn sich ein Benutzer verirrt hat oder ziellos umherwandert. Da sich die Benutzer dann eventuell nicht repräsentativ verhalten, wird stets nur der erste Besuch jeder Zone  $z$  in einer Protokolldatei betrachtet. Durch diese Maßnahme bleiben lediglich wenige protokollierte Informationsanfragen unberücksichtigt, dafür wird aber sichergestellt, dass die berücksichtigten Anfragen möglichst repräsentativ sind.

Wenn das Aktualisierungsintervall  $\Delta$  abgelaufen ist, muss im erweiterten Verfahren

neben der Aktualisierung jeder Zugriffswahrscheinlichkeitstabelle auch die Aktualisierung der Besuchswahrscheinlichkeitskarte angestoßen werden. In Algorithmus 4.4 ist die modifizierte Methode zur Bearbeitung einer Protokolldatei angegeben.

### 4.4.3 Aktualisierung der Zugriffswahrscheinlichkeitstabellen

Die Aktualisierungsmethode der Zugriffswahrscheinlichkeitstabellen ist der im Basisverfahren verwendeten Methode sehr ähnlich (siehe Algorithmus 4.5). Allerdings muss die Zugriffswahrscheinlichkeitstabelle einer Zone  $z$  nun die lokalen Zugriffswahrscheinlichkeiten  $a(i, z)$ , wie sie für die Zone  $z$  beobachtet wurden, beinhalten. Entsprechend ändert sich die Berechnungsvorschrift für die Zugriffswahrscheinlichkeiten  $a_\Delta$ :

$$a_\Delta(i, z) = \frac{r_\Delta(i, z)}{n_\Delta(z)}$$

```

method aktualisiere()
  for all  $r_\Delta(i, z) \in c(z)$  do
     $a_\Delta(i, z) := \frac{r_\Delta(i, z)}{n_\Delta(z)}$ ;
    if Eintrag  $\langle i, a(i, z) \rangle$  existiert in Zugriffswahrscheinlichkeitstabelle then
5:    $a(i, z) := \gamma \cdot a(i, z) + (1 - \gamma) \cdot a_\Delta(i, z)$ ;
    else
      Neuen Eintrag  $\langle i, a(i, z) \rangle$  anlegen;
       $a(i, z) := (1 - \gamma) \cdot a_\Delta(i, z)$ ;
    end if
10: end for
    if Tabellengröße > Schwellwert then
      Lösche (Tabellengröße – Schwellwert) Einträge mit den kleinsten  $a(i, z)$ ;
    end if

```

**Algorithmus 4.5:** Modifizierte Methode der Zugriffswahrscheinlichkeitstabelle für ihre Aktualisierung.

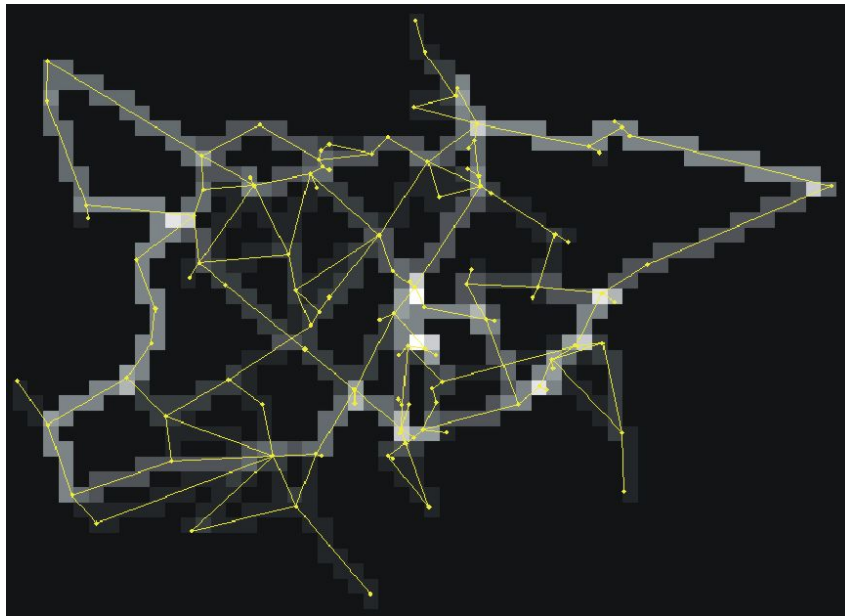
### 4.4.4 Besuchswahrscheinlichkeitskarten

Zusätzlich zu den Zugriffswahrscheinlichkeitstabellen der einzelnen Zonen verwaltet jede Infostation eine Besuchswahrscheinlichkeitskarte ihres Hoarding-Gebiets. Die Besuchswahrscheinlichkeitskarte eines Hoarding-Gebiets  $HG$  ist eine Funktion  $v : Z \mapsto$



$[0..1]$ , wobei  $Z$  die Menge aller Zonen innerhalb des Hoarding-Gebiets  $HG$  ist. Der Wert  $v(z)$  gibt die Wahrscheinlichkeit an, mit der ein Benutzer, der das Hoarding-Gebiet  $HG$  besucht, ebenfalls die Zone  $z$  besucht.

Aufgrund ihrer graphischen Darstellung werden diese Funktionen als Karten bezeichnet. Abbildung 4.6 zeigt die Darstellung einer solchen Karte für die Innenstadt von Stuttgart. Diese Karte wurde durch Simulation der Bewegungen von Touristen, die Stuttgart besuchen, erzeugt. In diesem Beispiel sind alle Zonen gleich große Quadrate. Je heller ein Quadrat dargestellt ist, umso höher ist die Besuchswahrscheinlichkeit der Zone.



**Abbildung 4.6:** Graphische Darstellung einer Besuchswahrscheinlichkeitskarte.

```

method aktualisiere()
  for all  $z \in Z$  do
     $v_{\Delta}(z) := \frac{n_{\Delta}(z)}{|\Delta|}$ ;
     $v(z) := \gamma \cdot v(z) + (1 - \gamma) \cdot v_{\Delta}(z)$ ;
  5: end for

```

**Algorithmus 4.6:** Methode der Besuchswahrscheinlichkeitskarte für ihre Aktualisierung.

Die Aktualisierung der Besuchswahrscheinlichkeitskarte funktioniert ähnlich wie die

Aktualisierung von Zugriffswahrscheinlichkeitstabellen (siehe Algorithmus 4.6). Aus der Anzahl  $n_{\Delta}(z)$  der während des Aktualisierungsintervalls gezählten Besuche einer Zone  $z$  und der Anzahl der bearbeiteten Protokolldateien kann die Wahrscheinlichkeit  $v_{\Delta}(z)$  berechnet werden, mit der ein Besucher, der während des Aktualisierungsintervalls das Hoarding-Gebiet der Infostation besuchte, auch die Zone  $z$  besuchte:

$$v_{\Delta}(z) = \frac{n_{\Delta}(z)}{|\Delta|}$$

Schließlich werden die bisher in der Besuchswahrscheinlichkeitskarte gespeicherten Werte  $v(z)$  auf neue Werte  $v'(z)$  gesetzt:

$$v'(z) = \gamma \cdot v(z) + (1 - \gamma) \cdot v_{\Delta}(z)$$

Dabei ist  $0 \leq \gamma \leq 1$  der gleiche Parameter wie bei der Aktualisierung der Zugriffswahrscheinlichkeitstabellen, der angibt wie stark Beobachtungen aus vorherigen Intervallen in den neuen Werten berücksichtigt werden.

### 4.4.5 Externes Wissen

Der Benutzer oder die Anwendung spezifizieren ihr Wissen in externen Besuchswahrscheinlichkeitskarten. Dabei werden zwei Arten von externem Wissen unterschieden. Bei der ersten Art handelt es sich um Wissen über konkrete Werte der Besuchswahrscheinlichkeiten einzelner Zonen. Solches Wissen wird in absoluten Besuchswahrscheinlichkeiten spezifiziert. Bei der anderen Art von Wissen ist nichts über konkrete Besuchswahrscheinlichkeiten bekannt, sondern lediglich, dass einige Gebiete mit größerer Wahrscheinlichkeit besucht werden als andere. Dieses Wissen wird durch relative Besuchswahrscheinlichkeiten dem Hoarding-Verfahren mitgeteilt.

Da der Benutzer oder die Anwendung unter Umständen nur Wissen über das Verhalten des Benutzers in bestimmten Teilgebieten hat, können externe Besuchswahrscheinlichkeitskarten im Gegensatz zur internen Karte der Infostation auch nur Teile eines Hoarding-Gebiets abdecken. Es ist auch möglich, dass mehrere externe Besuchswahrscheinlichkeitskarten angegeben werden, die sogar von verschiedenen externen

Informationsquellen stammen können. Um mögliche Inkonsistenzen durch die Spezifikation unterschiedlicher Besuchswahrscheinlichkeiten für die gleiche Zone auszuschließen, wird im Folgenden davon ausgegangen, dass zwei unterschiedliche externe Karten niemals die gleiche Zone abdecken. Wollte man diese Voraussetzung nicht machen, könnten eventuell auftretende Inkonsistenzen beispielsweise dadurch aufgelöst werden, dass immer der zuletzt spezifizierte Wert verwendet wird.

### **Absolute Besuchswahrscheinlichkeitskarten**

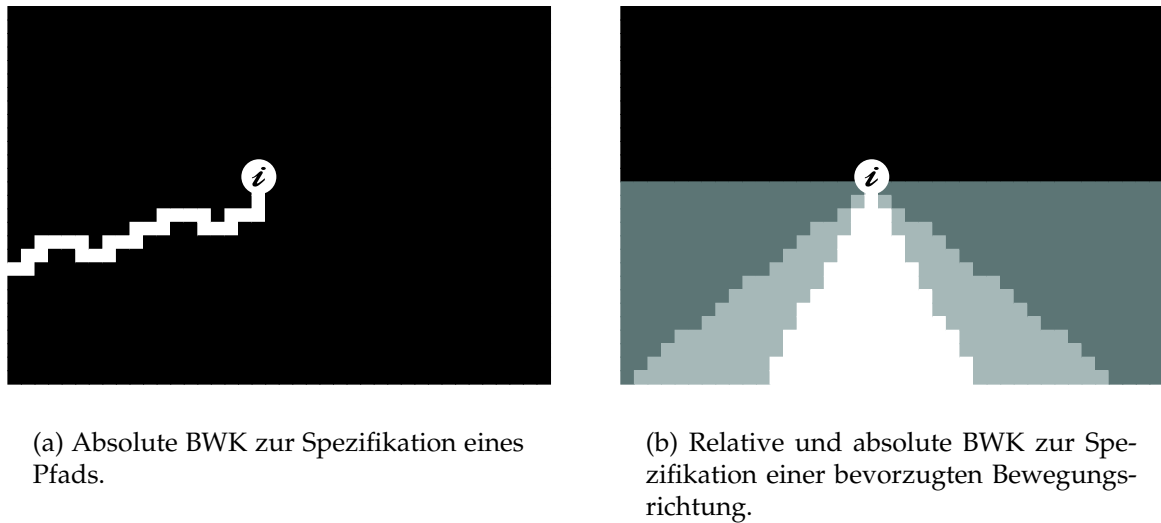
Eine absolute Besuchswahrscheinlichkeitskarte spezifiziert für einen Teil oder alle Zonen im Hoarding-Gebiet einer Infostation konkrete Werte für die Besuchswahrscheinlichkeiten der Zonen.

#### **DEFINITION 4.7 (ABSOLUTE BESUCHSWAHRSCHEINLICHKEITSKARTE)**

Eine absolute Besuchswahrscheinlichkeitskarte ist eine Funktion  $v_i : Z_i \rightarrow [0..1]$ , wobei  $Z_i \subseteq Z$  und  $Z_i \cap Z_j = \emptyset$ , wenn  $i \neq j$ .

Ein Beispiel für die Verwendung von absoluten Besuchswahrscheinlichkeitskarten ist die bereits erwähnte Navigationsanwendung, die den Pfad eines Benutzers angeben kann. Abbildung 4.7(a) zeigt eine absolute Besuchswahrscheinlichkeitskarte, die einen solchen Pfad spezifiziert. Sie weist den Zonen, die durch den Pfad abgedeckt werden eine Besuchswahrscheinlichkeit von eins zu. Allen übrigen Zonen wird die Besuchswahrscheinlichkeit null zugewiesen.

Ein mobiler Touristenführer kann mittels einer absoluten Besuchswahrscheinlichkeitskarte dem Hoarding-Verfahren mitteilen, dass der Benutzer ein bestimmtes Museum nicht besuchen wird, weil es zum gegenwärtigen Zeitpunkt geschlossen ist. Dazu muss lediglich für die vom Museum abgedeckten Zonen eine Besuchswahrscheinlichkeit von null spezifiziert werden. Auch der Ausschluss von Informationsobjekten, die zu bereits besuchten Zonen gehören, von der Vorabübertragung ist mit Hilfe einer externen absoluten Besuchswahrscheinlichkeitskarte möglich. In dieser Karte müssen nur die Besuchswahrscheinlichkeiten der bereits besuchten Zonen auf null gesetzt werden.



**Abbildung 4.7:** Externe Besuchswahrscheinlichkeitskarten.

### Relative Besuchswahrscheinlichkeitskarten

Eine relative Besuchswahrscheinlichkeitskarte weist einem Teil der Zonen oder allen Zonen im Hoarding-Gebiet einer Infostation relative Besuchswahrscheinlichkeiten zu.

#### DEFINITION 4.8 (RELATIVE BESUCHSWAHRSCHEINLICHKEITSKARTE)

Eine relative Besuchswahrscheinlichkeitskarte ist eine Funktion  $r_i : Z_i \rightarrow \mathbb{R}^+$ , wobei  $Z_i \subseteq Z$  und  $Z_i \cap Z_j = \emptyset$ , wenn  $i \neq j$ .

Wird einer Zone  $z_1$  die relative Besuchswahrscheinlichkeit  $r_i(z_1)$  zugeordnet und der Zone  $z_2$  die Wahrscheinlichkeit  $r_i(z_2)$ , dann bedeutet dies, dass die Zone  $z_1$   $\frac{r_i(z_1)}{r_i(z_2)}$  mal so häufig besucht wird wie die Zone  $z_2$ .

Ein Beispiel für die Verwendung von relativen Besuchswahrscheinlichkeitskarten ist in Abbildung 4.7(b) angegeben. Diese Karte spezifiziert, dass sich der Benutzer nach Verlassen der Infostation wahrscheinlich nach Süden bewegen wird. Dazu wird den dunkelgrauen Gebieten eine relative Besuchswahrscheinlichkeit von eins zugewiesen, den hellgrauen Gebieten wird eine relative Wahrscheinlichkeit von zwei zugewiesen und die weißen Gebiete schließlich erhalten eine relative Besuchswahrscheinlichkeit von drei. Das bedeutet, dass die hellgrauen Gebiete zweimal so wahrscheinlich besucht werden wie die dunkelgrauen Gebiete und dass die weißen Gebiete dreimal so

wahrscheinlich besucht werden wie die dunkelgrauen. Zusätzlich spezifiziert in diesem Beispiel eine absolute Besuchswahrscheinlichkeitskarte, dass die Gebiete nördlich der Infostation gar nicht besucht werden, indem sie ihnen eine absolute Besuchswahrscheinlichkeit von null zuweist. Das Beispiel zeigt auch, dass absolute und relative Karten sinnvoll kombiniert werden können. Auch dann wird davon ausgegangen, dass  $Z_i \cap Z_j = \emptyset$ , wenn  $i \neq j$ .

In einem elektronischen Stadtführer, der von einem Touristen benutzt wird, können relative Karten verwendet werden, um zu spezifizieren, dass der Benutzer Orte, an denen sich Sehenswürdigkeiten befinden, wahrscheinlicher besuchen wird als andere Orte.

Um für die Hoarding-Entscheidung verwendet werden zu können, müssen die relativen Besuchswahrscheinlichkeitskarten in absolute Besuchswahrscheinlichkeitskarten umgerechnet werden. Dabei gibt es für die absoluten Besuchswahrscheinlichkeiten  $v_i(z)$  jeweils prinzipiell unendlich viele Möglichkeiten, da lediglich die in der relativen Karte  $r_i$  spezifizierten Verhältnisse eingehalten werden müssen:

$$v_i(z) = c \cdot \frac{r_i(z)}{\sum_{j \in Z} r_i(j)}, c = \text{const.} \in \mathbb{R}^+$$

Da die relativen Karten jedoch eventuell nur Teile eines Hoarding-Gebiets abdecken, sollten sie die interne Besuchswahrscheinlichkeitskarte der Infostation nur soweit verändern, wie dies zur Einhaltung der spezifizierten Verhältnisse notwendig ist. Deckt beispielsweise eine relative Besuchswahrscheinlichkeitskarte ein Gebiet ab, das gemäß der Karte der Infostation nur sehr selten besucht wird, sollte die zu berechnende absolute Karte für dieses Gebiet nicht plötzlich sehr hohe Besuchswahrscheinlichkeiten angeben, die dann zwar die Verhältnisse der relativen Karte widerspiegeln aber im völligen Widerspruch zu den Beobachtungen der Infostation stehen. Daher wird bei der Umrechnung der relativen in absolute Wahrscheinlichkeiten versucht, die absoluten Wahrscheinlichkeiten so zu wählen, dass sich für sie der gleiche Durchschnittswert wie für die entsprechenden Wahrscheinlichkeiten in der Besuchswahrscheinlichkeitskarte der Infostation ergibt, d.h. es soll gelten:

$$\frac{\sum_{z \in Z_i} v_i(z)}{|Z_i|} = \frac{\sum_{z \in Z_i} v(z)}{|Z_i|}$$

$$\begin{aligned}
 &\Leftrightarrow \frac{\sum_{z \in Z_i} \left( c \cdot \frac{r_i(z)}{\sum_{j \in Z} r_i(j)} \right)}{|Z_i|} = \frac{\sum_{z \in Z_i} v(z)}{|Z_i|} \\
 &\Leftrightarrow c = \sum_{z \in Z_i} v(z) \\
 &\Leftrightarrow v_i(z) = \left( \sum_{z \in Z_i} v(z) \right) \cdot \frac{r_i(z)}{\sum_{j \in Z} r_i(j)}
 \end{aligned}$$

Allerdings können bei dieser Berechnung auch Werte  $v_i(z) > 1$  auftreten. Dann ist es nicht möglich, die in der relativen Karte angegebenen Verhältnisse einzuhalten und gleichzeitig den Durchschnittswert der absoluten Besuchswahrscheinlichkeiten, wie sie in der Karte der Infostation spezifiziert sind, beizubehalten. In diesem Fall müssen die Werte  $v_i(z)$  noch normiert werden. Die normierten Werte  $v'_i(z)$  werden wie folgt berechnet:

$$v'_i(z) = \frac{v_i(z)}{\max_{k \in Z} v_i(k)}$$

Danach ist sichergestellt, dass  $\forall z \in Z_i : v_i(z) \leq 1$ . Ferner liegt der Durchschnittswert der Werte  $v_i(z)$  so nah, wie unter Einhaltung der spezifizierten Verhältnisse möglich, am Durchschnittswert der entsprechenden Werte  $v(z)$  in der internen Besuchswahrscheinlichkeitskarte der Infostation.

An dem nachfolgenden Beispiel wird deutlich, dass der Durchschnittswert der Besuchswahrscheinlichkeiten in der Karte der Infostation nicht immer aufrecht erhalten werden kann: Eine externe relative Karte  $r_1$  spezifiziere, dass die Zone  $z_1$  zehn mal so häufig besucht werden wird, wie die Zonen  $z_2, \dots, z_{10}$ . Sei der Durchschnittswert der Besuchswahrscheinlichkeiten  $v(z_k)$  der Zonen  $z_1, \dots, z_{10}$  in der internen Karte der Infostation 0.5. Dann kann dieser Wert unter Einhaltung der in  $r_1$  spezifizierten Verhältnisse nicht erreicht werden. Der Zone  $z_1$  kann höchstens die Besuchswahrscheinlichkeit eins zugewiesen werden. Daraus ergibt sich für die Zonen  $z_2$  bis  $z_{10}$  eine Besuchswahrscheinlichkeit von 0.1 und folglich ein maximaler Durchschnittswert von  $\frac{1+9 \cdot 0.1}{10} = 0.19$ .

Der Durchschnittswert der in der Karte der Infostation gespeicherten Wahrscheinlichkeiten kann nur beibehalten werden, wenn keine Normierung erforderlich ist, d.h. es muss gelten:

$$\forall z \in Z_i : \left( \sum_{j \in Z_i} v(j) \right) \cdot \frac{r_i(z)}{\sum_{k \in Z_i} r_i(k)} \leq 1 \Leftrightarrow \forall z \in Z_i : r_i(z) \leq \frac{\sum_{j \in Z_i} r_i(j)}{\sum_{j \in Z_i} v(j)}$$

### 4.4.6 Integration

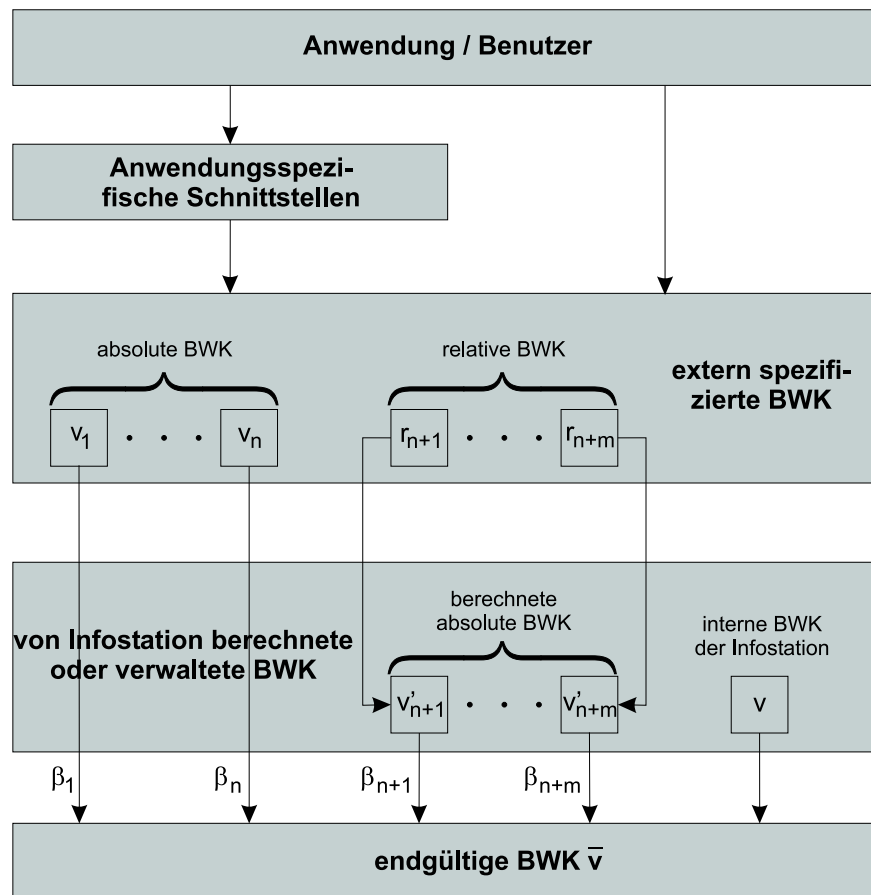
Die externen Besuchswahrscheinlichkeitskarten sind nur die einfachste, generische Schnittstelle zum vorgestellten Vorabübertragungsverfahren. Komfortablere, anwendungsspezifische Schnittstellen, die auf dieser einfachen Schnittstelle aufsetzen, sind ohne großen Aufwand realisierbar. So könnte eine Schnittstelle für Navigationsanwendungen aus der Angabe eines Pfads automatisch die entsprechende externe absolute Besuchswahrscheinlichkeitskarte berechnen. Eine weitere Schnittstelle könnte aus Richtungsangaben die entsprechenden relativen Karten berechnen. Die Anwendungen beziehungsweise die Benutzer müssen daher nichts über die Zonen und Hoarding-Gebiete wissen. In [Kubach & Rothermel 2000] wird ferner die Möglichkeit beschrieben, wie Besuchswahrscheinlichkeitskarten aus vorgefertigten Bausteinen (Hoarding-Patterns) zusammengesetzt werden können. Dieser Ansatz hat sich jedoch als zu unflexibel erwiesen.

Für die Hoarding-Entscheidung müssen alle externen Besuchswahrscheinlichkeitskarten und die Karte der Infostation zu einer konsistenten Sicht der Besuchswahrscheinlichkeiten integriert werden. Abbildung 4.8 gibt einen Überblick über diesen Integrationsprozess. Wie zuvor beschrieben wurde, spezifizieren die Benutzer oder die Anwendungen ihr Wissen über externe Besuchswahrscheinlichkeitskarten. Dies kann entweder direkt oder über anwendungsspezifische Schnittstellen erfolgen. Die relativen Besuchswahrscheinlichkeitskarten werden anschließend, wie im vorherigen Abschnitt beschrieben, in absolute Karten umgerechnet. Schließlich müssen alle absoluten Karten in eine endgültige Karte  $\bar{v}$  integriert werden. Dies geschieht nach der folgenden Vorschrift:

$$\forall z \in Z : \bar{v}(z) = \begin{cases} \beta_i \cdot v_i(z) + (1 - \beta_i) \cdot v(z), & \text{falls } z \in Z_i \\ v(z), & \text{falls } z \notin \bigcup_{\forall i} Z_i \end{cases}$$

Diese Vorschrift besagt, dass in den Zonen, für die kein externes Wissen spezifiziert wurde, die Karte der Infostation verwendet wird. In den Zonen, für die externes Wissen angegeben wurde, bestimmt der Parameter  $\beta_i$  wie sehr dem externen Wissen vertraut wird. Wenn  $\beta_i = 1$  wird der entsprechenden externen Karte voll vertraut, ist hingegen  $\beta_i = 0$  wird die zugehörige Karte gar nicht berücksichtigt.

Sonderfälle bilden die Zonen, für die die interne Karte der Infostation eine Besuchswahrscheinlichkeit von eins oder null angibt. In diesen Fällen werden die zugehöri-



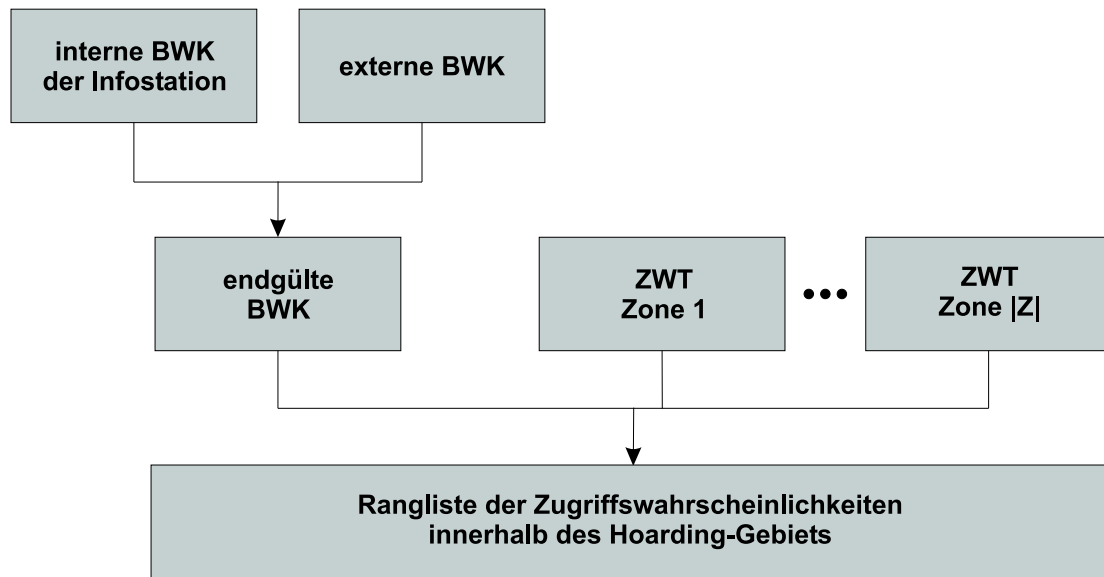
**Abbildung 4.8:** Integration aller verfügbaren BWK in eine endgültige BWK.

gen Zonen immer beziehungsweise niemals besucht. Dies lässt auf Beschränkungen in der Infrastruktur schließen, die den Benutzer zwingen, eine Zone immer zu besuchen, beispielsweise weil alle Straßen innerhalb des Hoarding-Gebiets durch diese Zone führen, beziehungsweise eine Zone für den Benutzer unerreichbar machen, beispielsweise weil keine Wege in diese Zone führen. Weil die externen Informationsquellen häufig kein Wissen über die Infrastruktur haben und daher solche Beschränkungen nicht kennen, werden in diesen Sonderfällen abweichend von der obigen Vorschrift immer unverändert die Werte aus der internen Karte der Infostation in die endgültige Karte  $\bar{v}$  übernommen.



### 4.4.7 Hoarding-Entscheidung

Bislang wurde erläutert wie aus den externen und der internen Karte der Infostation die endgültige Besuchswahrscheinlichkeitskarte berechnet wird. Auf Basis dieser endgültigen Karte und den Zugriffswahrscheinlichkeitstabellen aller Zonen im Hoarding-Gebiet wird schließlich die Hoarding-Entscheidung getroffen (siehe Abbildung 4.9).



**Abbildung 4.9:** Berechnung der Zugriffswahrscheinlichkeiten innerhalb des Hoarding-Gebiets.

Aus der endgültigen Besuchswahrscheinlichkeitskarte und den Zugriffswahrscheinlichkeitstabellen kann die Wahrscheinlichkeit berechnet werden, mit der innerhalb des Hoarding-Gebiets auf jedes Informationsobjekt  $i$  zugegriffen wird. Dazu wird zunächst die Wahrscheinlichkeit  $\bar{v}(z) \cdot a(i, z)$  berechnet, mit der  $i$  in einer bestimmten Zone  $z$  referenziert wird. Die komplementäre Wahrscheinlichkeit  $1 - \bar{v}(z) \cdot a(i, z)$  ist entsprechend die Wahrscheinlichkeit, dass in der Zone  $z$  nicht auf das Informationsobjekt  $i$  zugegriffen wird. Indem diese komplementären Wahrscheinlichkeiten für alle Zonen  $z$  im Hoarding-Gebiet multipliziert werden, erhält man die Wahrscheinlichkeit  $q(i)$ , dass im gesamten Hoarding-Gebiet nicht auf  $i$  zugegriffen wird:

$$q(i) = \prod_{z \in Z} (1 - \bar{v}(z) \cdot a(i, z))$$

Durch nochmaliges Komplementieren erhält man schließlich die Wahrscheinlichkeit  $a(i)$ , dass innerhalb des Hoarding-Gebiets mindestens einmal auf das Informationsobjekt zugegriffen wird:

$$a(i) = 1 - q(i) = 1 - \prod_{z \in Z} (1 - \bar{v}(z) \cdot a(i, z))$$

Für die Vorübertragung werden schließlich wiederum die Informationsobjekte ausgewählt, für die die höchsten Zugriffswahrscheinlichkeiten  $a(i)$  berechnet wurden. Es sei noch angemerkt, dass die Berechnung der  $a(i)$  wie beschrieben durchgeführt werden muss. Da die Besuchereignisse  $\bar{v}(z)$  nicht paarweise disjunkt sind, darf die totale Wahrscheinlichkeit  $a(i) = \sum_{z \in Z} \bar{v}(z) \cdot a(i, z)$  nicht verwendet werden.

Wenn im erweiterten Verfahren keinerlei externes Wissen zur Verfügung steht, sind die erzielten Ergebnisse annähernd identisch mit denen, die mit dem Basisverfahren erzielt werden. Dann werden im externen Verfahren die Wahrscheinlichkeiten, die im Basisverfahren direkt in der Zugriffswahrscheinlichkeitstabelle verwaltet werden, über den Umweg der Besuchswahrscheinlichkeitskarte und der einzelnen Zugriffswahrscheinlichkeitstabellen berechnet. Kleinere Unterschiede können sich aus der Tatsache ergeben, dass im erweiterten Verfahren jeweils nur der erste Besuch einer Zone betrachtet wird und wiederholte Zugriffe auf das gleiche Objekt während eines Besuchs einer bestimmten Zone nicht berücksichtigt werden.

## 4.5 Diskussion

In diesem Abschnitt werden zunächst die gemeinsamen Eigenschaften des Basisverfahrens und des erweiterten Verfahrens beschrieben. Anschließend wird der Einfluss der Größe der Hoarding-Gebiete und der Zonen auf die Trefferraten und die Kosten der Verfahren diskutiert.

### 4.5.1 Gemeinsame Eigenschaften beider Varianten

#### Adaptivität

Ein Vorteil beider Varianten des vorgestellten Verfahrens ist, dass keine feste Beziehung zwischen den Orten und den dort jeweils bevorzugten Informationsobjekten vorliegen muss. Vielmehr passen sich beide Varianten dynamisch an Veränderungen im

Zugriffsverhalten der Benutzer an, indem sie kontinuierlich beobachten, welche Informationsobjekte jeweils bevorzugt werden.

So könnten beispielsweise die Benutzer eines mobilen Stadtführers tagsüber Informationen über Geschäfte und Einkaufszentren bevorzugen, während sie abends jedoch am gleichen Ort verstärkt auf Informationen über Restaurants, Bars und Kinos zugreifen. Eine solche Veränderung würde sich in den Zugriffswahrscheinlichkeitstabellen widerspiegeln und die Benutzer würden stets primär mit den zur Zeit bevorzugten Informationen versorgt.

### **Skalierbarkeit**

Ein weiterer Vorteil sind die guten Skalierungseigenschaften beider Varianten. Sollte eine Infostation überlastet sein, kann dieses Problem einfach behoben werden, indem in der Nähe der überlasteten Infostation eine oder mehrere neue Infostationen installiert werden. Es ist sogar möglich, dass die Hoarding-Gebiete zweier verschiedener Infostationen das gleiche geographische Gebiet abdecken. Die in jedem beliebigen Gebiet auftretende Last kann also auf beliebig viele Infostationen verteilt werden. Durch die Installation neuer Infostationen kann auch das gesamte vom Informationssystem versorgte Gebiet vergrößert werden.

Lediglich der Kommunikationsaufwand zwischen den Infostationen wächst mit einer zunehmenden Anzahl von Infostationen, weil eine Infostation eine empfangene Protokolldatei eventuell zu einer größeren Anzahl von Infostationen weiterleiten muss. Allerdings erfolgt die Verteilung der Protokolldateien über ein Festnetz und die erforderliche Kommunikation bleibt lokal sehr begrenzt, da eine vom mobilen Endgerät zur Infostation übertragene Protokolldatei üblicherweise nur Zugriffe enthält, die im Hoarding-Gebiet der empfangenden Infostation oder benachbarter Infostationen aufgetreten sind. Die einzige Ausnahme ist, wenn der Benutzer an sehr vielen Infostationen vorbei geht, ohne diese zu besuchen. Dies sollte jedoch allein schon deshalb vermieden werden, um zu verhindern dass der Benutzer nur Informationen auf seinem Endgerät hat, die nicht mehr zu seiner aktuellen Umgebung gehören.

Sollte dennoch der Kommunikationsaufwand zu groß werden, kann eine Infostation darauf verzichten, einen Teil der empfangenen Protokolldateien weiterzuleiten, ohne dass dadurch ein größerer Schaden entstehen würde. Denn wenn zwischen den Infostationen ein hohes Kommunikationsaufkommen herrscht, bedeutet dies, dass sehr

viele Informationen über das Benutzerverhalten zwischen den Infostationen ausgetauscht werden. Entsprechend kann auch noch mit einem Teil dieser Informationen ein gutes Bild vom Benutzerverhalten gewonnen werden.

### Datenschutz

In beiden Varianten des Verfahrens wird eine Reihe von Informationen über das Verhalten eines Benutzers im System gesammelt. Im Rahmen dieser Arbeit wird davon ausgegangen, dass die Benutzer dem Betreiber der Infostationsinfrastruktur vertrauen, wie dies beispielsweise auch in Mobilfunksystemen der Fall ist. Daher werden keine weiteren Maßnahmen getroffen, um den Missbrauch der gesammelten Daten zu verhindern.

Kann ein solches Vertrauen nicht vorausgesetzt werden, ergibt sich das Problem, dass Informationen über das Verhalten des Benutzers mit der Adresse seines Endgeräts in Verbindung gebracht werden können, da die Geräteadresse für die Übertragung der Protokolldateien bekannt sein muss. Daher müssen in diesem Fall Maßnahmen ergriffen werden, die die Geräteadresse vor dem Betreiber der Infostationsinfrastruktur verbergen, wie zum Beispiel anonymisierende Proxies<sup>1</sup>, Crowds<sup>2</sup> oder Onion Routing<sup>3</sup>. Ein Überblick über diese Techniken gibt [Federrath & Pfitzmann 1998].

### 4.5.2 Größe von Hoarding-Gebieten und Zonen

Bevor im Folgenden der Einfluss der Größe von Hoarding-Gebieten und Zonen auf die mit dem Vorübertragungsverfahren erzielten Trefferraten diskutiert wird, werden zunächst einige Definitionen eingeführt, die die Beschreibung wichtiger Parameter erleichtern.

#### Definitionen

##### DEFINITION 4.9 (ZUGEHÖRIGKEIT BEZÜGLICH $t$ )

Ein Informationsobjekt gehört bezüglich eines festen Schwellwerts  $t$  zu einem geographisch begrenzten Gebiet, wenn die Zugriffswahrscheinlichkeit auf dieses Objekt innerhalb des Gebiets den Schwellwert  $t$  übersteigt.

---

<sup>1</sup><http://www.anonymizer.com>

<sup>2</sup><http://www.research.att.com/projects/crowds>

<sup>3</sup><http://www.onion-router.net>

In einem nicht inhärent ortsbezogenen Informationssystem sollte  $t > 0$  gewählt werden, da bezüglich des Schwellwerts  $t = 0$  in solchen Systemen in der Regel alle Objekte des Informationsraums zu jedem beliebigen Gebiet gehören. Für inhärent ortsbezogene Systeme macht  $t = 0$  jedoch durchaus Sinn, da hier in jedem Gebiet nur eine Teilmenge aller Objekte des Informationsraums referenzierbar ist. Entsprechend kann man für solche Systeme einen generellen Zugehörigkeitsbegriff definieren:

**DEFINITION 4.10 (ZUGEHÖRIGKEIT)**

In einem inhärent ortsbezogenen Informationssystem gehört ein Informationsobjekt zu einem geographisch begrenzten Gebiet, wenn die Zugriffswahrscheinlichkeit auf dieses Objekt innerhalb des Gebiets größer als null ist.

Mittels der Zugehörigkeit bezüglich eines Schwellwerts kann die mittlere Informationsdichte bezüglich eines Schwellwerts definiert werden:

**DEFINITION 4.11 (MITTLERE INFORMATIONSDICHT BEZÜGLICH  $t$ )**

Die mittlere Informationsdichte  $\rho_t$  bezüglich eines festen Schwellwerts  $t$  in einem geographisch begrenzten Gebiet ist die Zahl der bezüglich des Schwellwerts  $t$  zu dem Gebiet gehörenden Informationsobjekte, dividiert durch die Größe des Gebiets.

Entsprechend wird die mittlere Informationsdichte für inhärent ortsbezogene Systeme folgendermaßen definiert:

**DEFINITION 4.12 (MITTLERE INFORMATIONSDICHT BEZÜGLICH  $t$ )**

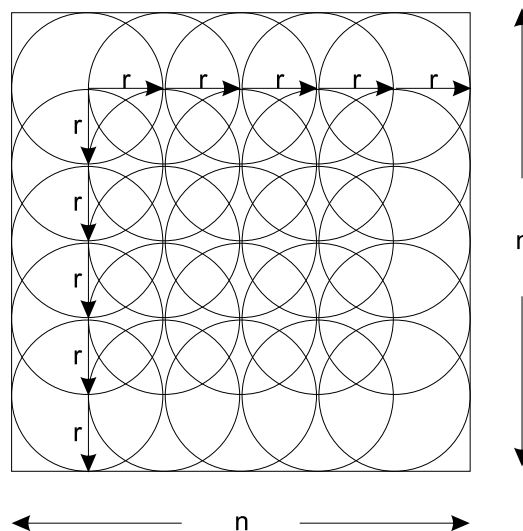
In einem inhärent ortsbezogenen Informationssystem ist die mittlere Informationsdichte  $\rho$  in einem geographisch begrenzten Gebiet die Zahl der zu dem Gebiet gehörenden Informationsobjekte, dividiert durch die Größe des Gebiets.

## **Hoarding-Gebiete**

Die Größe der Hoarding-Gebiete hat einen entscheidenden Einfluss auf die mit beiden Verfahren erzielten Trefferraten. Geht man davon aus, dass die Anzahl der pro Flächeneinheit benötigten Informationsobjekte konstant ist, so nimmt die Anzahl der

innerhalb eines Hoarding-Gebiets benötigten Informationsobjekte linear mit der Größe des Gebiets zu. Entsprechend kann mit den gleichen Ressourcen nur ein kleinerer Teil der benötigten Informationen vorab übertragen werden. Deshalb führen größere Hoarding-Gebiete im Allgemeinen zu niedrigeren Trefferraten.

Aus zweierlei Gründen werden bei kleineren Hoarding-Gebieten leider mehr Infostationen benötigt. Zum einen muss jede Position, von der aus auf das Informationssystem zugegriffen werden kann, von mindestens einem Hoarding-Gebiet abgedeckt werden. Ferner müssen sich die Hoarding-Gebiete, wie in Abschnitt 4.2 beschrieben, überlappen, um eine lückenlose Versorgung des Benutzers mit Informationen gewährleisten zu können. Diese Restriktionen führen dazu, dass sich die Anzahl der für die Versorgung eines bestimmten Gebiets benötigten Infostationen umgekehrt proportional zur Größe der Hoarding-Gebiete verhält. Das heißt, dass beispielsweise für eine Halbierung der von einem Hoarding-Gebiet abgedeckten Fläche, die Zahl der Infostationen verdoppelt werden muss. Abbildung 4.10 veranschaulicht, dass zur Abdeckung eines  $n^2$  Flächeneinheiten umfassenden Gebiets  $\frac{n^2}{r^2}$  Infostationen notwendig sind, wenn die Hoarding-Gebiete jeweils Kreise mit einem Radius von  $r$  Längeneinheiten sind.



**Abbildung 4.10:** Abdeckung eines Gebiets durch einzelne Hoarding-Gebiete.

Für die weitere Diskussion der Größe von Hoarding-Gebieten wird ein inhärent ortsbezogenes Informationssystem betrachtet. Es wird angenommen, dass die Informati-

onsdichte in jedem beliebigen Teilgebiet des vom Informationssystem versorgten Gesamtgebiets gleich ist und dass die vom Vorabübertragungsverfahren beobachteten Zugriffswahrscheinlichkeiten denen des betrachteten Benutzers exakt entsprechen.

Unter diesen Annahmen kann die mit dem Basisverfahren zu erwartende Trefferrate  $ET_B$  innerhalb eines Hoarding-Gebiets berechnet werden als die Summe der Zugriffswahrscheinlichkeiten auf die einzelnen vorab übertragenen Informationsobjekte:

$$ET_B = \sum_{i=1}^S p(i),$$

wobei  $p(i)$  die Zugriffswahrscheinlichkeit auf das  $i$ -te vorab übertragene Informationsobjekt ist und insgesamt  $S$  Objekte vorab übertragen werden.

Bei einer Gleichverteilung der Zugriffswahrscheinlichkeiten und einer Flächengröße  $A$  des Hoarding-Gebiets ergibt sich für  $p(i) = \frac{1}{\text{Anzahl zum Hoarding-Gebiet gehörender Objekte}} = \frac{1}{\rho \cdot A}$  und für die zu erwartende Trefferrate  $ET_B$ :

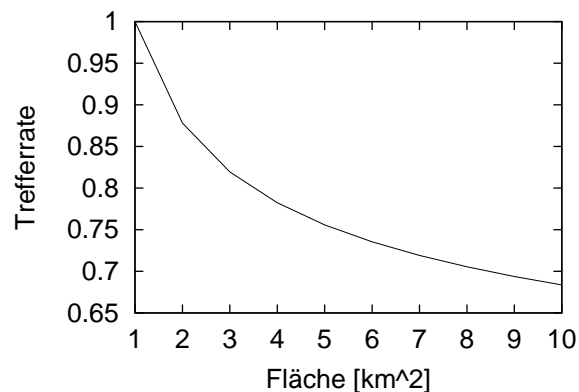
$$ET_B = \sum_{i=1}^S p(i) = \frac{S}{\rho \cdot A}$$

Mit der realistischeren Zipf-Verteilung (siehe Abschnitt 6.2.2) erhält man  $p(i) = \frac{1}{i} \cdot \frac{1}{\sum_{j=1}^S \frac{1}{j}}$  und für  $ET_B$ :

$$ET_B = \frac{1}{\sum_{i=1}^S \frac{1}{i}} \cdot \sum_{i=1}^S \frac{1}{i}$$

In Abbildung 4.11 ist diese zu erwartende Trefferrate für eine Größe des Hoarding-Gebiets von 1 km<sup>2</sup> bis 10 km<sup>2</sup> dargestellt. Dabei wurde für  $\rho$  ein Wert von 80 Informationsobjekten pro km<sup>2</sup> und für  $S$  ebenfalls ein Wert von 80 Objekten angenommen. Bei den Berechnungen der zu erwartenden Trefferrate wurde eine mögliche Überlappung der jeder Flächeneinheit zugeordneten Informationsobjekte vernachlässigt.

Im Allgemeinen sollte für die Wahl der Größe eines Hoarding-Gebiets die jeweilige Informationsdichte  $\rho_t$  bezüglich eines festen Schwellwerts  $t$  betrachtet werden. In Gegenden mit hoher Informationsdichte sollten die Hoarding-Gebiete kleiner gewählt werden als in Gegenden mit kleiner Informationsdichte. Idealerweise sollte die Menge an Informationsobjekten, die zu jedem Hoarding-Gebiet gehört, möglichst konstant



**Abbildung 4.11:** Trefferrate in Abhängigkeit von der Größe des Hoarding-Gebiets.

sein, so dass die Benutzer in allen Gebieten in etwa die gleichen Trefferraten erwarten können.

### Zonen

Für die mit dem erweiterten Verfahren erzielten Trefferraten spielt die Anzahl beziehungsweise die Größe der Zonen, in die ein Hoarding-Gebiet unterteilt ist, ebenfalls eine wichtige Rolle. Im Allgemeinen gilt, dass kleinere Zonen eine höhere Granularität in der Hoarding-Entscheidung ermöglichen und somit zu höheren Trefferraten führen.

In diesem Sinne ist der Idealfall, dass jeder Ort, der einen eigenen Kontext besitzt, zu einer separaten Zone gehört. Beispielsweise sollte in einem mobilen Touristenführer jeder Ort, an dem sich eine Sehenswürdigkeit befindet, einer getrennten Zone zugeordnet sein. Sobald zwei Orte in der gleichen Zone liegen, kann das Vorübertragungsverfahren nicht mehr zwischen ihnen unterscheiden. Der geringste Nutzen aus dem erweiterten Verfahren ergibt sich, wenn das gesamte Hoarding-Gebiet aus nur einer Zone besteht. Dann werden mit dem erweiterten Verfahren keine höheren Trefferraten erzielt als mit dem Basisverfahren. Leider erhöhen sich mit einer steigenden Anzahl von Zonen die Kosten des Verfahrens (siehe Abschnitt 4.5.3).

Zur Veranschaulichung des Einflusses der Anzahl von Zonen wird nachfolgend für den Fall eines inhärent ortsbezogenen Informationssystems die mit dem erweiterten Verfahren zu erwartende Trefferrate  $ET_E$  in Abhängigkeit von der Anzahl der Zonen berechnet. Sei das Hoarding-Gebiet einer Infostation mit der Fläche  $A$  zunächst in  $|Z|$



Zonen unterteilt. Sei die Informationsdichte im betrachteten Hoarding-Gebiet  $\rho$  und die Besuchswahrscheinlichkeit für alle Zonen  $v$ . Weiter wird angenommen, dass die Bewegungspfade der Benutzer stets korrekt spezifiziert werden.

Dann kann die mittlere Anzahl  $n$  der Objekte, auf die ein Benutzer zugreifen kann, wie folgt berechnet werden:

$$n = \# \text{ besuchter Zonen} \cdot \# \text{ verfügbarer Objekte pro Zone} = |Z| \cdot v \cdot \frac{\rho \cdot A}{|Z|} = v \cdot \rho \cdot A$$

Damit kann die Wahrscheinlichkeit  $p_H$ , dass ein referenziertes Informationsobjekt vorab übertragen wurde, folgendermaßen berechnet werden:

$$p_H = \frac{S}{n} = \frac{S}{v \cdot \rho \cdot A},$$

wobei  $S$  wiederum die Anzahl vorab übertragener Objekte ist und eine Gleichverteilung der Zugriffswahrscheinlichkeiten angenommen wird.

Schließlich kann die zu erwartende Trefferrate  $ET_E$  berechnet werden:

$$ET_E = \frac{N \cdot p_H}{N} = p_H = \frac{S}{v \cdot \rho \cdot A},$$

wobei  $N$  die Anzahl der referenzierten Objekte ist.

Werden nun je  $m$  der ursprünglichen Zonen zu einer Zone zusammengefasst, ergibt sich für die Besuchswahrscheinlichkeiten  $v'$  der neuen Zonen:

$$v' = 1 - (1 - v)^m$$

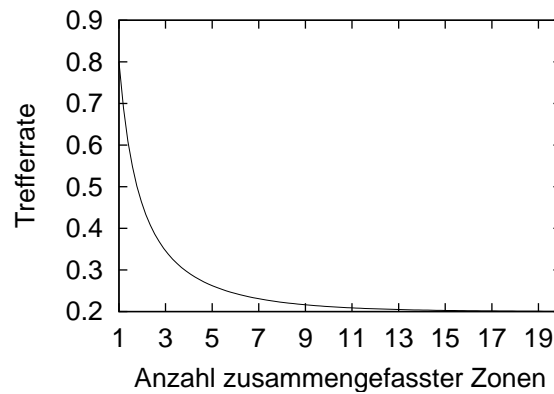
Zu jeder neuen Zone gehören dann  $\frac{\rho \cdot A \cdot m}{|Z|}$  Objekte, folglich ergibt sich für die mittlere Anzahl  $n'$  der Objekte, auf die ein Benutzer zugreifen kann, jetzt:

$$n' = \frac{|Z|}{m} \cdot v' \cdot \frac{\rho \cdot A \cdot m}{|Z|} = v' \cdot \rho \cdot A$$

Für die Werte  $p'_H$  und  $ET'_E$  gilt folglich:

$$ET'_E = p'_H = \frac{S}{v' \cdot \rho \cdot A}$$

Abbildung 4.12 zeigt, wie sich die Trefferrate in Abhängigkeit von der Zahl der zu einer Zone zusammengefassten ursprünglichen Zonen entwickelt. Für die Parameter wurden folgende Werte gewählt:  $v = 0.25$ ,  $\rho = 80 \frac{\text{Objekte}}{\text{km}^2}$ ,  $A = 5 \text{ km}^2$ ,  $S = 80$  Objekte.



**Abbildung 4.12:** Trefferrate in Abhängigkeit von der Anzahl zusammengefasster Zonen.

### 4.5.3 Kosten

#### Kommunikationskosten

In beiden Varianten des vorgestellten Verfahrens ergeben sich die gleichen Kommunikationskosten. Sie können in die Kosten für die drahtlose Kommunikation  $K_d$  und die Festnetzkommunikation  $K_f$  aufgeteilt werden.

Die Kosten für die drahtlose Kommunikation entstehen zum einen durch den Transfer der Protokolldateien von den mobilen Endgeräten zu den Infostationen und zum anderen natürlich durch die Übertragung der für die Vorübertragung ausgewählten Informationsobjekte von der Infostation zum mobilen Endgerät.

Der Aufwand für den Transfer der Protokolldateien wird durch deren Größe  $G_P$  bestimmt. Für  $G_P$  gilt:

$$G_P = r \cdot (G_{OB} + G_{KP}),$$

wobei  $r$  die Anzahl der Informationsanfragen ist, die ein Benutzer auf seinem Weg von einer Infostation zur nächsten stellt,  $G_{OB}$  die Größe eines Objektbezeichners und  $G_{KP}$  die Größe eines Koordinatenpaares angibt.

Werden URLs als Objektbezeichner verwendet und geht man von einer maximalen Länge von 255 Zeichen aus, ergibt sich für  $G_{OB}$  ein maximaler Wert von 255 byte. Bei der Verwendung von Koordinaten einfacher Genauigkeit erhält man für  $G_{KP}$  einen Wert von acht Bytes. Geht man von  $r = 100$  Anfragen aus, ergibt sich in diesem Fall für die Größe der Protokolldatei ein maximaler Wert von  $G_P = 100 \cdot (255 \text{ byte} + 8 \text{ byte}) \approx 25.68 \text{ kbyte}$ .

In einem proprietären Informationssystem mit einer festen Anzahl an Informationsobjekten kann für die Länge der Objektbezeichner ein deutlich kleinerer Wert angenommen werden. Bei einer fortlaufenden Nummerierung können beispielsweise mit nur drei Bytes schon mehr als 16 Millionen Informationsobjekte bezeichnet werden. Dann erhält man selbst bei  $r = 1000$  Anfragen nur vergleichsweise kleine Protokolldateien mit  $G_P = 1000 \cdot (3 + 8) \text{ byte} \approx 10.74 \text{ kbyte}$ .

Die für die Übertragung der ausgewählten Informationsobjekte zu transferierende Datenmenge wird, wie bereits in Abschnitt 4.1 diskutiert wurde, im Bereich weniger MBytes (typisch zwei bis acht MBytes) liegen. Demgegenüber ist der Kommunikationsaufwand für die Übertragung der Protokolldateien über das drahtlose LAN einer Infostation vernachlässigbar. Die Kommunikationkosten  $K_d$  werden also im Wesentlichen durch die Anzahl der vorabübertragenen Informationsobjekte bestimmt.

Der Kommunikationsaufwand  $K_f$  im Festnetz setzt sich aus den Kosten für den Transfer der Informationsobjekte vom Informationssystem zum Zugangspunkt des drahtlosen Netzes der Infostation und den Kosten für die Verteilung der Protokolldateien unter den Infostationen zusammen. Der Aufwand für den Transfer der Informationsobjekte wird nachfolgend nicht mehr weiter betrachtet, weil er einerseits durch Caching und Replikation der Objekte auf dem Infostation-Server stark reduziert werden kann und andererseits dem Aufwand in klassischen Proxy-basierten Informationssystemen entspricht.

Der Aufwand für die Verteilung einer Protokolldatei hängt wiederum von deren Größe  $G_P$  ab. Ferner muss beachtet werden, zu wievielen Infostationen die Protokolldatei gesendet werden muss. Für den Aufwand  $K_P$  zum Verteilen einer Protokolldatei gilt:

$$K_P = m \cdot c \cdot G_P + K_V,$$

wobei  $m$  die Anzahl der Infostationen ist, für die Teile der zu verteilenden Protokolldatei relevant sind.  $c$  ist der Anteil der Einträge in der Protokolldatei, die im Mittel

weitergeleitet werden müssen.  $K_V$  sind schließlich die Kosten für die Kommunikation mit dem Verzeichnisdienst, die notwendig ist, um festzustellen für welche Infostationen ein Eintrag in der Protokolldatei relevant ist. Auch diese Kosten können durch die Verwendung geeigneter Caches so stark reduziert werden, dass sie nicht weiter betrachtet werden müssen.

Um das gesamte Transfervolumen zu bestimmen, das von einer Infostation in einer bestimmten Zeiteinheit ausgeht, um alle Protokolldateien zu verteilen, muss die Anzahl der im Mittel pro Zeiteinheit von mobilen Endgeräten empfangenen Protokolldateien betrachtet werden. Sie entspricht der mittleren Anzahl  $b$  der Besuche der Infostation pro Zeiteinheit. Für das gesamte Transfervolumen einer Infostation  $K_{P,ges}$  pro Zeiteinheit erhält man dann  $K_{P,ges} = b \cdot K_P$ .

Ist  $|I|$  die Gesamtzahl der Infostationen, dann kann der gesamte zur Verteilung der Protokolldateien benötigte Kommunikationsaufwand im Festnetz pro Zeiteinheit mit

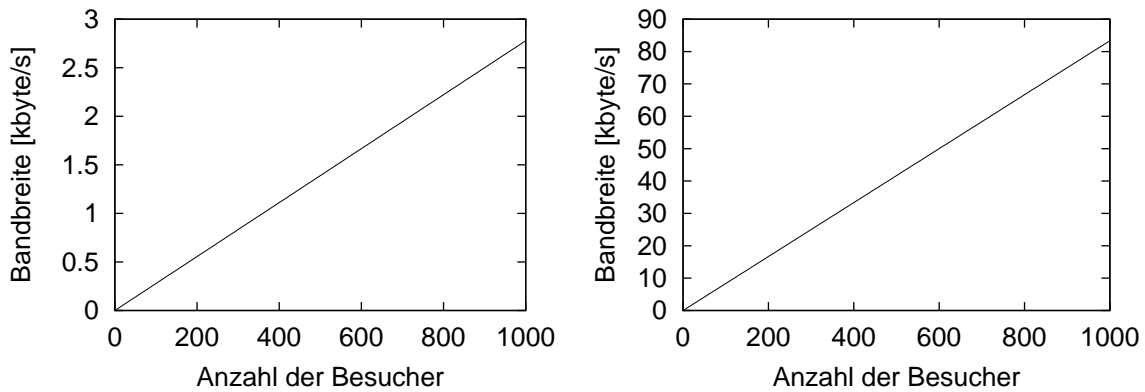
$$K_f = |I| \cdot K_{P,ges} = |I| \cdot b \cdot K_P = |I| \cdot b \cdot m \cdot c \cdot G_P$$

angegeben werden.

Um den Kommunikationsaufwand zu veranschaulichen, werden folgende konkrete Werte für die einzelnen Parameter angenommen: Für  $G_P$  wird ein Wert von 20 kbyte angenommen. Es wird davon ausgegangen, dass die Hälfte aller Einträge in jeder Protokolldatei zu drei Infostationen weiter verteilt werden muss und dass insgesamt zehn Infostationen vorhanden sind. Abbildung 4.13(a) zeigt die zwischen jedem Paar von Infostationen benötigte Bandbreite, während in Abbildung 4.13(b) das gesamte Datenvolumen, das pro Zeiteinheit für die Verteilung aller Protokolldateien transferiert werden muss, dargestellt ist. Die Anzahl der Besucher bezieht sich jeweils auf eine Stunde und eine Infostation. Die einzelnen Parameterwerte sind eher pessimistisch gewählt, das heißt der tatsächliche Aufwand könnte durchaus geringer sein.

### Speicherkosten

Im Basisverfahren muss lediglich eine Zugriffswahrscheinlichkeitstabelle auf jedem Infostation-Server gespeichert werden. Die Kosten dafür berechnen sich als Produkt der Zahl der Einträge und der Größe eines Eintrags  $e$ . Die Größe eines Eintrags ergibt sich als Summe der Länge eines Objektbezeichners und der Speicherkosten für



(a) Zwischen jedem Paar von Infostationen benötigte Bandbreite.

(b) Insgesamt transferiertes Datenvolumen.

**Abbildung 4.13:** Kommunikationsaufwand zur Verteilung von Protokolldateien.

die Zugriffswahrscheinlichkeit, die bei Verwendung von Gleitkommazahlen einfacher Genauigkeit vier Bytes betragen. Da die Zahl der Einträge in einer Zugriffswahrscheinlichkeitstabelle auf einen konfigurierbaren Maximalwert  $max$  beschränkt ist, erhält man für die maximale Größe einer Zugriffswahrscheinlichkeitstabelle:

$$S_{ZWT} = max \cdot e$$

Bei Verwendung von URLs mit maximal 255 Zeichen und den erwähnten Gleitkommazahlen einfacher Genauigkeit ergibt sich für  $e$  eine Größe von 259 byte. Der Wert  $max$  sollte in Abhängigkeit von der Anzahl der Informationsobjekte, die im Mittel vorab übertragen werden können, gewählt werden. Können beispielsweise nur 100 Informationsobjekte im Mittel übertragen werden, ist es sicherlich mehr als ausreichend die jeweils 200 beliebtesten Informationsobjekte in der Zugriffswahrscheinlichkeitstabelle zu speichern, d.h.  $max = 200$ . Mit diesen typischen Werten ergibt sich für die maximale Tabellengröße  $S_{ZWT} = 200 \cdot 259 \text{ byte} \approx 50.58 \text{ kbyte}$ .

Im erweiterten Verfahren muss für jede Zone im Hoarding-Gebiet eine eigene Zugriffswahrscheinlichkeitstabelle gespeichert werden. Dies bedeutet jedoch gegenüber dem Basisverfahren nicht zwangsläufig einen  $|Z|$ -fachen Speicheraufwand bei  $|Z|$  Zonen. Denn der Wert  $max$  für die maximale Tabellengröße kann in der Regel kleiner gewählt werden, sobald die Benutzer im Mittel mehr als eine Zone besuchen, da dann der auf

dem Endgerät für die Vorübertragung verfügbare Speicherplatz auf die einzelnen Zonen verteilt wird. Entsprechend werden pro Zone weniger Informationsobjekte übertragen als für das gesamte Hoarding-Gebiet im Basisverfahren. Daher reicht es aus, für weniger Objekte die Zugriffswahrscheinlichkeiten in der Zugriffswahrscheinlichkeitstabelle jeder Zone vorzuhalten.

Mit Hilfe der Informationsdichte  $\rho_t$  kann ausgerechnet werden, welcher Speicheraufwand sich ergäbe, wenn der Maximalwert  $max$  für die Größe der Zugriffswahrscheinlichkeitstabellen so gewählt würde, dass stets alle bezüglich des Schwellwerts  $t$  zur jeweiligen Zone gehörenden Informationsobjekte in jeder Zugriffswahrscheinlichkeitstabelle enthalten sind. Sei  $\rho_t$  konstant für das gesamte Hoarding-Gebiet und das Hoarding-Gebiet mit der Fläche  $A$  in  $|Z|$  Zonen unterteilt, dann müssen in der Tabelle jeder Zone  $\frac{A}{|Z|} \cdot \rho_t$  Einträge gespeichert werden. Daraus ergibt sich ein Gesamtspeicheraufwand  $S_{ZWT,ges}$  für alle Tabellen von:

$$S_{ZWT,ges} = |Z| \cdot \frac{A}{|Z|} \cdot \rho_t \cdot e = A \cdot \rho_t \cdot e.$$

Die Größe  $e$  eines Eintrags bleibt dabei gegenüber dem Basisverfahren unverändert. Interessant ist die Beobachtung, dass der Gesamtspeicheraufwand dann nicht mehr von der Zahl der Zonen abhängt. Für typische Werte ( $A = 5km^2$ ,  $\rho_t = 80 \frac{Objekte}{km^2}$ ,  $e = 259$  byte) erhält man  $S_{ZWT,ges} \approx 101$  kbyte.

Dazu kommen noch die Kosten  $S_{BWK}$  zur Speicherung der Besuchswahrscheinlichkeitskarte. Diese hängen linear von der Zahl  $|Z|$  der Zonen im Hoarding-Gebiet ab, weil für jede Zone die jeweilige Besuchswahrscheinlichkeit gespeichert werden muss, d.h.  $S_{BWK} = |Z| \cdot f$ , wobei  $f$  der Speicherbedarf für eine Besuchswahrscheinlichkeit ist. Für die Angabe von Besuchswahrscheinlichkeiten sind Gleitkommazahlen einfacher Genauigkeit ausreichend, was für  $f$  einen Wert von vier Bytes bedeutet.

### Rechenaufwand

Der Rechenaufwand setzt sich im Basisverfahren aus dem Aufwand für die Bearbeitung der Protokolldateien, dem Aufwand zur Aktualisierung der Zugriffswahrscheinlichkeitstabelle sowie dem Aufwand zur Berechnung der Hoarding-Entscheidung zusammen.

Während der Bearbeitung einer Protokolldatei muss die Protokolldatei unter Umständen zunächst an andere Infostationen verteilt werden. Der Aufwand dafür wächst linear mit der Länge der Protokolldatei, da für jeden Eintrag die Infostationen, für die er relevant ist, beim Verzeichnisdienst erfragt werden müssen. Danach muss für jeden Eintrag in der Protokolldatei geprüft werden, ob der entsprechende Zugriffszähler bereits angelegt wurde, gegebenenfalls muss ein neuer Zähler angelegt werden und schließlich muss der Zähler noch inkrementiert werden. Die Bearbeitung eines einzelnen Eintrags kann daher bei Verwendung geeigneter Hashing-Verfahren zum Auffinden der Zähler in annähernd konstanter Rechenzeit erfolgen. Der Rechenaufwand  $R_P$  für die Bearbeitung einer Protokolldatei hängt also linear von deren Länge ab. Für jeden Benutzer, der an einer Infostation ankommt, muss eine Protokolldatei bearbeitet werden. Zusätzlich müssen die von anderen Infostationen empfangenen Protokolldateien bearbeitet werden. Daher kann der gesamte Aufwand  $R_{P,ges}$ , der im Mittel pro Zeiteinheit für die Bearbeitung von Protokolldateien von einer Infostation aufgebracht werden muss, mit

$$R_{P,ges} = (a + b) \cdot \mathcal{O}(l)$$

abgeschätzt werden. Dabei ist  $b$  wiederum die mittlere Zahl der Besucher, die die Infostation pro Zeiteinheit besuchen und  $l$  die mittlere Länge der Protokolldateien.  $a$  gibt die mittlere Anzahl der pro Zeiteinheit von anderen Infostationen empfangenen Protokolldateien an.

Während eines Aktualisierungsintervalls  $\Delta$  wird für jedes Informationsobjekt, das innerhalb des Hoarding-Gebiets mindestens einmal angefordert wird, ein Zugriffszähler  $r_\Delta(i)$  angelegt. Bei der Aktualisierung der Zugriffswahrscheinlichkeitstabelle muss für jeden dieser Zähler der zugehörige Wert  $a_\Delta(i)$  berechnet werden und der entsprechende Wert  $a(i)$  in der Zugriffswahrscheinlichkeitstabelle aktualisiert werden. Der Aufwand für diese Berechnungen ist, wiederum bei Verwendung von Hash-Verfahren zum Auffinden der Tabelleneinträge, annähernd konstant. Daher erhält man für den Gesamtaufwand  $R_{A,ges}$ , der pro Zeiteinheit von einer Infostation für die Aktualisierung erbracht werden muss:

$$R_{A,ges} = \frac{|\Delta|}{a + b} \cdot \mathcal{O}(|c|),$$

wobei  $|c|$  die Anzahl der Zähler ist, die während des Intervalls  $\Delta$  angelegt wurden und  $|\Delta|$  die Länge des Intervalls  $\Delta$  angibt. Der Wert  $|c|$  entspricht der Anzahl der verschie-

denen Informationsobjekte, auf die während  $\Delta$  innerhalb des Hoarding-Gebiets der Infostation zugegriffen wurde.

Für die eigentliche Hoarding-Entscheidung müssen nur noch die Einträge in der Zugriffswahrscheinlichkeitstabelle nach den Zugriffswahrscheinlichkeiten sortiert werden und danach die beliebtesten Objekte ausgewählt werden. Da für jeden Besucher eine Hoarding-Entscheidung getroffen wird und der Aufwand für das Sortieren im Mittel  $\mathcal{O}(l \cdot \log l)$  beträgt, ergibt sich für den pro Zeiteinheit für die Hoarding-Entscheidung zu erbringenden mittleren Aufwand  $R_{H,ges}$ :

$$R_{H,ges} = b \cdot \mathcal{O}(l \cdot \log l)$$

Im Basisverfahren ist es nicht unbedingt erforderlich, für jeden Benutzer eine individuelle Hoarding-Entscheidung zu treffen, da der jeweilige Benutzer ohnehin keinen Einfluss auf die Hoarding-Entscheidung hat. Daher könnte die Hoarding-Entscheidung auch nur periodisch berechnet werden, wodurch gegenüber einer erneuten Berechnung für jeden Benutzer lediglich die Aktualität der Entscheidungen leidet, aber Rechenaufwand eingespart wird.

Im erweiterten Verfahren muss zusätzlich noch die Besuchswahrscheinlichkeitskarte verwaltet werden. Bei der Bearbeitung einer Protokolldatei ergibt sich dadurch ein zusätzlicher Aufwand, da die Zahl der Besuche einer jeden Zone gezählt werden muss. Mit einer ähnlichen Argumentation wie für das Basisverfahren kann jedoch gezeigt werden, dass dieser Aufwand und der restliche Aufwand, der pro Eintrag erbracht werden muss, wiederum annähernd konstant ist. Daher gilt die Abschätzung für  $R_{P,ges}$  aus dem Basisverfahren auch für das erweiterte Verfahren.

Auch für den Rechenaufwand zur Aktualisierung einer Zugriffswahrscheinlichkeitstabelle kann analog zum Basisverfahren argumentiert werden. Allerdings müssen nun  $|Z|$  Zugriffswahrscheinlichkeitstabellen aktualisiert werden und die Anzahl der zu betrachtenden Zähler  $|c|$  wird nicht mehr durch die Anzahl verschiedener Informationsobjekte, die im Hoarding-Gebiet angefragt werden bestimmt, sondern nur noch durch die, die innerhalb der zur Tabelle gehörenden Zone angefragt werden. Daher erhält man für  $R_{A,ges}$  nun:

$$R_{A,ges} = \frac{|\Delta|}{a+b} \cdot |Z| \cdot \mathcal{O}(|c|)$$

Für die Hoarding-Entscheidung müssen im erweiterten Verfahren zunächst alle verfügbaren Besuchswahrscheinlichkeitskarten zur endgültigen Karte  $\bar{v}$  integriert wer-



den. Der größte Aufwand entsteht dabei, wenn für jede Zone im Hoarding-Gebiet eine relative Besuchswahrscheinlichkeit spezifiziert wird, da dann der Aufwand für die Umrechnung in absolute Wahrscheinlichkeiten und der Aufwand für die eigentliche Integration anfällt. Es spielt dabei keine Rolle, ob die relativen Wahrscheinlichkeiten in einer Karte oder in mehreren Karten spezifiziert werden. Der Aufwand für die Umrechnung der relativen Karten in absolute Karten wächst linear mit der Zahl der Zonen im Hoarding-Gebiet. Das Gleiche gilt für den Aufwand für die unter Umständen erforderliche Normierung, sowie die eigentliche Integration. Daher wächst der Gesamtaufwand  $R_{I,ges}$  für die Integration der Besuchswahrscheinlichkeitskarten linear mit der Zahl der Zonen:  $R_{I,ges} = \mathcal{O}(|Z|)$ .

Für die Hoarding-Entscheidung muss im schlechtesten Fall für jeden Eintrag  $\langle i, a(i, z) \rangle$  in jeder der  $|Z|$  Zugriffswahrscheinlichkeitstabellen die entsprechende Zugriffswahrscheinlichkeit  $a(i)$  berechnet werden. Da bereits der Aufwand für die Berechnung eines einzelnen Wertes  $a(i)$  linear mit der Zahl der Zonen wächst, ergibt sich für den maximalen Aufwand  $R_{B,ges}$  zur Berechnung aller  $a(i)$ :  $R_{B,ges} = max \cdot \mathcal{O}(|Z|^2)$ , wobei  $max$  wieder die maximale Tabellengröße ist.

Schließlich müssen die  $a(i)$  wieder sortiert und die Objekte mit den größten Zugriffswahrscheinlichkeiten selektiert werden. Dies erfordert einen mittleren Aufwand von  $R_S = \mathcal{O}(max \cdot |Z| \cdot \log(max \cdot |Z|))$ .

Im erweiterten Verfahren muss im Allgemeinen für jeden Besucher einer Infostation eine individuelle Hoarding-Entscheidung getroffen werden, weil jeder Besucher andere externe Besuchswahrscheinlichkeitskarten spezifizieren kann. Entsprechend gilt für den Gesamtaufwand pro Zeiteinheit für die Hoarding-Entscheidung:

$$R_{H,ges} = b \cdot (R_{I,ges} + R_{B,ges} + R_S) = b \cdot \mathcal{O}(|Z|^2)$$

## 4.6 Erweiterungen

In diesem Abschnitt werden zwei Konzepte zur Erweiterung des bisher vorgestellten Verfahrens vorgeschlagen, die es ermöglichen die Spezifität der Hoarding-Entscheidungen weiter zu erhöhen. Diese Konzepte können sowohl im Basisverfahren als auch im erweiterten Verfahren eingesetzt werden.

Das erste Konzept sind die so genannten Informationskanäle, die es erlauben das Profil von Anwendungen und Benutzern bei der Hoarding-Entscheidung zu berücksichtigen. Neben der Berücksichtigung von Profilen ist es häufig auch wichtig, ganz individuelle Eigenschaften des Benutzers oder seines Endgeräts bei der Hoarding-Entscheidung in Betracht zu ziehen. Dies wird durch ein Filterkonzept geleistet.

### 4.6.1 Informationskanäle

Wenn verschiedene Anwendungen auf den gleichen Informationsraum zugreifen, wie das beispielsweise in der neXus-Plattform der Fall sein wird, dann werden sich diese Anwendungen in den bevorzugten Informationsobjekten unterscheiden. So wird zum Beispiel eine Navigationsanwendung primär auf Karteninformationen und Verkehrsnachrichten zugreifen, während ein Restaurantführer primär Informationen über Restaurants benötigt. Zusätzlich werden die Benutzer verschiedener Anwendungen unter Umständen verschiedene Orte vorzugsweise besuchen. Um solche Unterschiede in der Hoarding-Entscheidung berücksichtigen zu können, muss das Wissen über die bevorzugten Informationsobjekte und die Besuchswahrscheinlichkeiten getrennt für jede unterstützte Anwendung gesammelt werden.

Ähnlich verhält es sich mit unterschiedlichen Benutzertypen der gleichen Anwendung. Ein Tourist der einen elektronischen Stadtführer benutzt, wird auf andere Informationen zugreifen und andere Orte besuchen, als ein Geschäftsreisender, der den gleichen Stadtführer benutzt. Deshalb sollte auch in solchen Fällen die Möglichkeit bestehen, die Informationen über bevorzugt referenzierte Informationsobjekte und bevorzugt besuchte Orte für typische Benutzerprofile getrennt zu sammeln.

Die Grundidee des Kanalkonzepts ist es, für jedes typische Benutzer- beziehungsweise Anwendungsprofil einen getrennten Informationskanal bereitzustellen, der den Benutzer beziehungsweise die Anwendung mit auf das jeweilige Profil zugeschnittenen Informationen versorgt. Dazu wird im Basisverfahren für jeden Kanal  $k$  eine eigene kanalspezifische Zugriffswahrscheinlichkeitstabelle  $ZWT_k$  und im erweiterten Verfahren  $|Z|$  kanalspezifische Zugriffswahrscheinlichkeitstabellen  $ZWT_{k,1}$  bis  $ZWT_{k,|Z|}$  sowie eine kanalspezifische Besuchswahrscheinlichkeitskarte  $BWK_k$  verwaltet. Die Tabellen und Karten enthalten jeweils spezifisches Wissen über das Verhalten der Anwendung beziehungsweise des Benutzers des jeweiligen Profils.

Anwendungen und Benutzer können eventuell Informationen unterschiedlicher Profile benötigen. Ein elektronischer Stadtführer benötigt zum Beispiel sowohl Navigationsinformation als auch Informationen über Restaurants, ein Tourist, der unterwegs auch einige Einkäufe erledigen muss, dürfte Informationen über Sehenswürdigkeiten und Einkaufsmöglichkeiten benötigen. Daher wäre es zu restriktiv, wenn jede Anwendung und jeder Benutzer genau einem Kanal zugeordnet wäre. Daher gibt es im hier vorgestellten Informationskanal-Konzept die Möglichkeit, mehrere Kanäle zu abonnieren.

### Abonnieren eines Kanals

Der Benutzer beziehungsweise die Anwendung abonnieren einen oder mehrere Informationskanäle. Dabei wird das Interesse in jeden abonnierten Kanal  $k$  durch die Vergabe eines Gewichtes  $g(k) \in \mathbb{R}^+$  näher spezifiziert. Jedem nicht abonnierten Kanal  $k$  in der Menge der unterstützten Kanäle  $K$  wird das Gewicht  $g(k) = 0$  zugewiesen. Aus den Gewichten  $g(k)$  wird für jeden Kanal  $k \in K$  ein relatives Gewicht  $\bar{g}(k)$  berechnet:

$$\bar{g}(k) = \frac{g(k)}{\sum_{i \in K} g(i)}$$

### Verwaltung der Tabellen und Karten

Im Wesentlichen funktioniert die Verwaltung der Zugriffswahrscheinlichkeitstabellen und Besuchswahrscheinlichkeitskarten wie bisher. Allerdings müssen nun bereits im Basisverfahren  $|K|$  Zugriffswahrscheinlichkeitstabellen an jeder Infostation für ihr Hoarding-Gebiet verwaltet werden. Im erweiterten Verfahren muss jede Infostation  $|K| \cdot |Z|$  Zugriffswahrscheinlichkeitstabellen sowie  $|K|$  Besuchswahrscheinlichkeitskarten verwalten. Ohne weitere Optimierungen ergibt sich also in jeder der beiden Varianten des Verfahrens ein  $|K|$ -facher Speicheraufwand bei der Verwendung von  $|K|$  Informationskanälen. Eine einfache Optimierung, die die Mehrfachspeicherung der Objektbezeichner vermeidet, ist es, jeweils nur eine Zugriffswahrscheinlichkeitstabelle zu verwalten und dafür die Einträge zu Einträgen der Form  $\langle i, a_1(i), \dots, a_{|K|}(i) \rangle$  beziehungsweise  $\langle i, a_1(i, z), \dots, a_{|K|}(i, z) \rangle$  zu erweitern, wobei  $a_k$  jeweils die für den Kanal  $k$  beobachtete Zugriffswahrscheinlichkeit ist.

Ferner müssen im Basisverfahren für jeden Kanal  $k$  eigene Zugriffszähler  $r_{\Delta,k}(i)$  und ein eigener Zähler  $anzahlAnfragen_k$  für die Anzahl der berücksichtigten Anfragen verwaltet werden. Im erweiterten Verfahren werden für jeden Kanal  $k$ , getrennte Zugriffszähler  $r_{\Delta,k}(i, z)$ , Besuchszähler  $n_{\Delta,k}$  sowie je ein Zähler  $anzahlDateien_k$  für die Anzahl der berücksichtigten Protokolldateien benötigt. Bei der Bearbeitung einer Protokolldatei müssen die Gewichte, mit denen die einzelnen Kanäle abonniert wurden, berücksichtigt werden. Im einfachsten Fall wurde nur ein Kanal  $k$  abonniert. Dann müssen die protokollierten Zugriffe, sowie im erweiterten Verfahren die besuchten Zonen, nur in den Zugriffszählern  $r_{\Delta,k}$  und gegebenenfalls in den Besuchszählern  $n_{\Delta,k}$  berücksichtigt werden. Die Zähler der übrigen Kanäle bleiben dann von den Informationen in der Protokolldatei gänzlich unberührt.

Wurden jedoch mehrere Kanäle abonniert, stellt sich das Problem, dass nicht festgestellt werden kann, unter welchem Profil ein bestimmter Informationszugriff erfolgte. Hat beispielsweise ein Tourist den Kulturkanal und den Architekturkanal abonniert, dann kann grundsätzlich nicht festgestellt werden, ob der Benutzer auf eine bestimmte Information als Kultur- oder als Architektur-Interessierter zugegriffen hat. Daher werden bei der Bearbeitung einer Protokolldatei die Zähler der einzelnen Kanäle entsprechend dem Gewicht, mit dem die Kanäle während der Aufzeichnung der Protokolldatei abonniert waren, inkrementiert. Befindet sich also in einer Protokolldatei ein Eintrag  $\langle i, (x, y) \rangle$ , dann werden im Basisverfahren die Zugriffszähler  $r_{\Delta,k}(i)$  und die Zähler für die Anzahl der Anfragen wie folgt inkrementiert:

$$\begin{aligned} r_{\Delta,k}(i) &= r_{\Delta,k}(i) + \bar{g}(k) \\ anzahlAnfragen_k &= anzahlAnfragen_k + \bar{g}(k) \end{aligned}$$

Entsprechend wird mit den Zählern im erweiterten Verfahren vorgegangen:

$$\begin{aligned} r_{\Delta,k}(i, z) &= r_{\Delta,k}(i, z) + \bar{g}(k) \\ n_{\Delta,k}(z) &= n_{\Delta,k}(z) + \bar{g}(k) \\ anzahlDateien_k &= anzahlDateien_k + \bar{g}(k) \end{aligned}$$

Damit können am Ende eines Aktualisierungsintervalls im Basisverfahren die Werte  $a_{\Delta,k}(i) = \frac{r_{\Delta,k}(i)}{anzahlAnfragen_k}$  berechnet werden. Analog gilt im erweiterten Verfahren  $a_{\Delta,k}(i, z) = \frac{r_{\Delta,k}(i, z)}{n_{\Delta,k}(z)}$  und  $v_{\Delta,k}(z) = \frac{n_{\Delta,k}(z)}{anzahlDateien_k}$ . Die Werte  $a_k(i)$  beziehungsweise  $a_k(i, z)$  und  $v_k(z)$  werden unter Berücksichtigung der Historie wie zuvor berechnet.

## Hoarding-Entscheidung

Im Basisverfahren werden die kanalspezifischen Zugriffswahrscheinlichkeiten  $a_k$  aus den Tabellen der einzelnen Kanäle entsprechend folgender Vorschrift zu einer kanalübergreifenden Zugriffswahrscheinlichkeit  $a$  zusammengefasst:

$$a(i) = \sum_{k \in K} \bar{g}(k) \cdot a_k(i)$$

Analog zur bisherigen Hoarding-Entscheidung werden die Informationsobjekte mit den höchsten globalen Zugriffswahrscheinlichkeiten  $a$  für die Vorabübertragung ausgewählt.

Im erweiterten Verfahren bieten sich prinzipiell zwei Möglichkeiten, wie die kanalspezifischen Informationen in die Hoarding-Entscheidung einfließen können:

1. Bei der ersten Möglichkeit werden zunächst die kanalspezifischen Besuchswahrscheinlichkeitskarten  $BWK_k$  der einzelnen Kanäle entsprechend den Gewichten, mit denen der Benutzer die Kanäle abonniert hat, zu einer kanalübergreifenden Besuchswahrscheinlichkeitskarte  $BWK$  zusammengefasst:

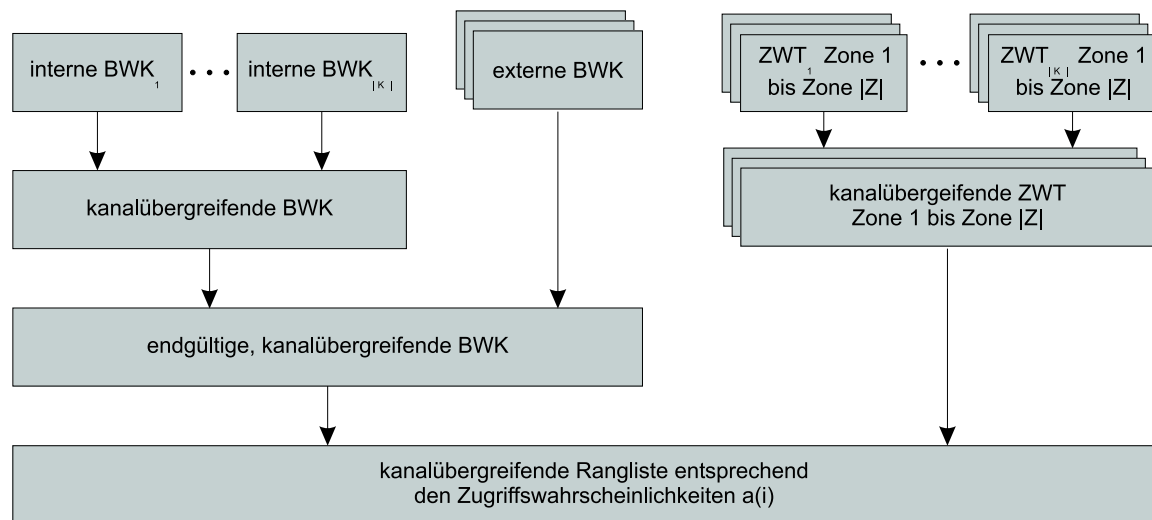
$$v(z) = \sum_{k \in K} \bar{g}(k) \cdot v_k(z),$$

wobei  $v_k(z)$  die kanalspezifische Besuchswahrscheinlichkeit der Zone  $z$  gemäß der Karte  $BWK_k$  ist und  $v(z)$  die kanalübergreifende Besuchswahrscheinlichkeit von  $z$  ist.

Die kanalspezifischen Zugriffswahrscheinlichkeitstabellen  $ZWT_{k,z}$  der einzelnen Kanäle, die jeweils zur gleichen Zone  $z$  gehören, können analog zu einer kanalübergreifenden Zugriffswahrscheinlichkeitstabelle  $ZWT_z$  der Zone  $z$  zusammengefasst werden:

$$a(i, z) = \sum_{k \in K} \bar{g}(k) \cdot a_k(i, z)$$

Damit erhält man wie im bisherigen erweiterten Verfahren eine Besuchswahrscheinlichkeitskarte  $BWK$  und für jede Zone  $z \in Z$  eine Zugriffswahrscheinlichkeitstabelle  $ZWT_z$ . Davon ausgehend kann die restliche Hoarding-Entscheidung dann genau so wie im bisherigen erweiterten Verfahren ablaufen (siehe Abbildung 4.14).



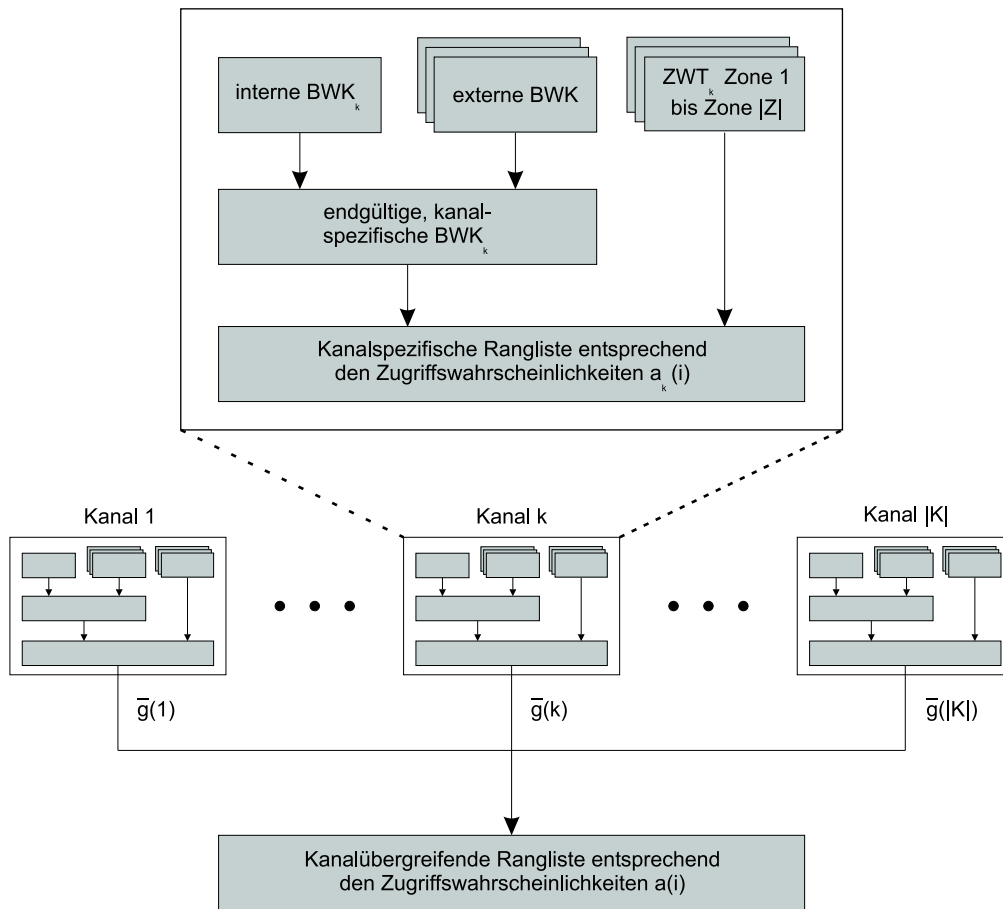
**Abbildung 4.14:** Alternative 1 zur Berücksichtigung von Informationskanälen in der Hoarding-Entscheidung.

2. Alternativ (siehe Abbildung 4.15) kann man auch zunächst für jeden Kanal getrennt aus den spezifischen Zugriffswahrscheinlichkeitstabellen  $ZWT_{k,z}$  und der Besuchswahrscheinlichkeitskarte  $BWK_k$  des Kanals die kanalspezifische Wahrscheinlichkeit  $a_k(i)$  berechnen, mit der die Abonnenten des Kanals innerhalb des Hoarding-Gebiets auf das Informationsobjekt  $i$  zugreifen. Diese Berechnung verläuft unter Einbeziehung eventuell angegebener externer Karten völlig analog zur Berechnung der  $a(i)$  im bisherigen erweiterten Verfahren, nur werden die kanalspezifischen Zugriffswahrscheinlichkeitstabellen mit den Wahrscheinlichkeiten  $a_k(i, z)$  und die kanalspezifische Besuchswahrscheinlichkeitskarte mit den Wahrscheinlichkeiten  $v_k(z)$  des jeweiligen Kanals verwendet.

Dadurch erhält man für jedes Informationsobjekt  $i$  und jeden Kanal  $k$  eine kanalspezifische Zugriffswahrscheinlichkeit  $a_k(i)$ . Diese Zugriffswahrscheinlichkeiten müssen dann noch entsprechend den Gewichten, mit denen die Kanäle abonniert wurden, zu einer kanalübergreifenden Zugriffswahrscheinlichkeit  $a(i)$  zusammengefasst werden:

$$a(i) = \sum_{k \in K} \bar{g}(k) \cdot a_k(i)$$

Aufgrund der klareren Semantik wurde in dieser Arbeit die zweite Alternative ge-



**Abbildung 4.15:** Alternative 2 zur Berücksichtigung von Informationskanälen in der Hoarding-Entscheidung.

wählt. In der ersten Möglichkeit geht die Beziehung zwischen den kanalspezifischen Zugriffswahrscheinlichkeiten  $a_k(i, z)$  in einer Zone  $z$  und der kanalspezifischen Besuchswahrscheinlichkeit der entsprechenden Zone  $z$  verloren, da vor der Multiplikation der Zugriffswahrscheinlichkeiten mit den Besuchswahrscheinlichkeiten die gewichteten Mittelwerte gebildet werden. Dies wird an folgendem Beispiel deutlich:

Man betrachte eine Zone  $z$ , ein Informationsobjekt  $i$  sowie zwei Kanäle  $k_1$  und  $k_2$ . Die Besuchswahrscheinlichkeitskarte des Kanals  $k_1$  weise für die Zone  $z$  eine kanalspezifische Besuchswahrscheinlichkeit  $v_{k_1}(z) = 0.01$  aus, entsprechend sei  $v_{k_2}(z) = 1$ . Sei die kanalspezifische Zugriffswahrscheinlichkeit auf das Informationsobjekt  $i$  in der Zone  $z$  für Kanal  $k_1$   $a_{k_1}(i, z) = 1$  und für Kanal  $k_2$   $a_{k_2}(i, z) = 0.01$ . Ferner habe die Anwendung beziehungsweise der Benutzer beide Kanäle mit dem gleichen Gewicht

$g(k_1) = g(k_2) = 0.5$  abonniert und keine externen Besuchswahrscheinlichkeitskarten spezifiziert.

Dann folgt mit dem ersten Ansatz für die kanalübergreifende Besuchswahrscheinlichkeit der Zone  $z$ :  $v(z) = g(k_1) \cdot v_{k_1}(z) + g(k_2) \cdot v_{k_2}(z) = 0.5 \cdot 0.01 + 0.5 \cdot 1 = 0.505$  und für die kanalübergreifende Zugriffswahrscheinlichkeit auf das Informationsobjekt  $i$  in der Zone  $z$ :  $a(i, z) = g(k_1) \cdot a_{k_1}(i, z) + g(k_2) \cdot a_{k_2}(i, z) = 0.5 \cdot 1 + 0.5 \cdot 0.01 = 0.505$ . Wenn  $z$  die einzige Zone im Hoarding-Gebiet ist, erhält man für die kanalübergreifende Zugriffswahrscheinlichkeit  $a(i)$  auf das Informationsobjekt  $i$ :  $a(i) = v(z) \cdot a(i, z) = 0.505 \cdot 0.505 = 0.255025$ .

Im Gegensatz dazu ergibt sich in der zweiten Variante zunächst, dass Abonnenten des Kanals  $k_1$  mit der kanalspezifischen Wahrscheinlichkeit  $a_{k_1}(i) = v_{k_1}(z) \cdot a_{k_1}(i, z) = 0.01 \cdot 1 = 0.01$  auf das Informationsobjekt  $i$  zugreifen. Der gleiche Wert ergibt sich für Abonnenten des zweiten Kanals:  $a_{k_2}(i) = v_{k_2}(z) \cdot a_{k_2}(i, z) = 1 \cdot 0.01 = 0.01$ . Wie erwartet erhält man für die kanalübergreifende Zugriffswahrscheinlichkeit  $a(i)$  auf das Informationsobjekt  $i$ :  $a(i) = g(k_1) \cdot a_{k_1}(i) + g(k_2) \cdot a_{k_2}(i) = 0.5 \cdot 0.01 + 0.5 \cdot 0.01 = 0.01$ .

Der Aufwand für die Berechnung der  $a(i)$  setzt sich im ersten Ansatz zusammen aus dem Aufwand für die Berechnung der kanalübergreifenden Besuchswahrscheinlichkeiten  $v(z)$  und der kanalübergreifenden Zugriffswahrscheinlichkeiten  $a(i, z)$ . Für den Aufwand zur Berechnung der Besuchswahrscheinlichkeiten ergibt sich  $|Z| \cdot \mathcal{O}(|K|)$ , während der Aufwand für die Zugriffswahrscheinlichkeiten mit  $|Z| \cdot \mathcal{O}(|K|^2)$  abgeschätzt werden kann. Hinzukommt der Aufwand  $R_{I,ges}$  für die Integration externer Besuchswahrscheinlichkeitskarten sowie der Aufwand  $R_{B,ges}$  für die Berechnung der globalen Zugriffswahrscheinlichkeiten  $a(i)$  (siehe Abschnitt 4.5.3). Damit ergibt sich ein Gesamtaufwand von  $|Z| \cdot \mathcal{O}(|K|) + |Z| \cdot \mathcal{O}(|K|^2) + \mathcal{O}(|Z|) + max \cdot \mathcal{O}(|Z|^2)$ . Der Aufwand hängt also quadratisch von der Anzahl der Zonen und der Anzahl der Kanäle ab.

Für die zweite Alternative ergibt sich in der Regel ein etwas höherer absoluter Berechnungsaufwand, weil eventuell vorhandene externe Karten in die kanalspezifische Besuchswahrscheinlichkeitskarte jeden abonnierten Kanals integriert werden müssen und für jeden abonnierten Kanal die Berechnung der  $a_k(i)$  durchgeführt werden muss. Dies bedeutet bereits einen Aufwand von  $|K| \cdot (R_{I,ges} + R_{B,ges})$ . Hinzu kommt noch der Aufwand für die Berechnung der kanalübergreifenden Wahrscheinlichkeiten  $a(i)$  aus den kanalspezifischen Wahrscheinlichkeiten  $a_k(i)$ . Er kann mit  $|Z| \cdot max \cdot \mathcal{O}(|K|^2)$



abgeschätzt werden. Damit ergibt sich für den Gesamtaufwand bei Verwendung der zweiten Alternative:  $|K| \cdot \mathcal{O}(|Z|) + |K| \cdot \max \cdot \mathcal{O}(|Z|^2) + |Z| \cdot \max \cdot \mathcal{O}(|K|^2)$ . Also hängt auch hier der Berechnungsaufwand quadratisch von der Anzahl der Zonen und der Anzahl der Kanäle ab. Daher ist die Verwendung der zweiten Alternative auch bezüglich des Rechenaufwands vertretbar.

### 4.6.2 Filter

Unabhängig davon ob Informationskanäle eingesetzt werden oder nicht, basiert die Hoarding-Entscheidung sowohl im Basisverfahren als auch im erweiterten Verfahren auf einer Rangliste der Informationsobjekte, die nach deren Zugriffswahrscheinlichkeit geordnet ist.

Dabei kann es vorkommen, dass bestimmte Informationsobjekte auf ein bestimmtes Endgerät nicht übertragen werden sollten, obwohl sie eine sehr hohe Zugriffswahrscheinlichkeit haben. So würde es zum Beispiel keinen Sinn machen, ein hochauflösendes Bild auf einen PDA, dessen Display typischerweise eine Auflösung von 160 x 160 Pixel hat, zu übertragen. Ebenso macht es wenig Sinn sehr detaillierte Straßenkarten zu einer Navigationsanwendung zu übertragen, deren Benutzer sich mit hoher Geschwindigkeit auf einer Autobahn bewegt [Ye et al. 1998].

Ziel der Filter ist es, solche in einem bestimmten Kontext nicht benötigte Informationsobjekte zu identifizieren und von der Vorabübertragung auszuschließen. Prinzipiell könnte man dies auch mit den Informationskanälen erreichen, indem man für jeden Kontext, der eine spezifische Filterung benötigt, einen separaten Informationskanal verwaltet. Dies würde allerdings in einer sehr hohen Zahl von Informationskanälen resultieren, da man für jede Kombination möglicher Ausprägungen der Kontextkriterien einen eigenen Kanal anlegen müsste. Bei  $n$  Kriterien, die jeweils in  $m$  Ausprägungen auftreten können, würde dies  $m^n$  Kanäle bedeuten. Wollte man beispielsweise nur zwischen hoch und niedrig auflösenden Displays sowie sich schnell und langsam bewegend Benutzern unterscheiden, müssten bereits vier Kanäle angelegt werden, um alle Kombinationen aus Auflösung und Bewegungsgeschwindigkeit abzudecken. Dies ist natürlich aus Speicherplatzgründen und wegen des Rechenaufwands nicht vertretbar.

Daher werden Filter zur Nachbearbeitung der berechneten Ranglisten eingesetzt. Die Filter sind als Methoden implementiert, die als einen Parameter die bisherige Rang-

liste erhalten und als Resultat eine reduzierte Rangliste zurückliefern, in der die nicht benötigten Informationsobjekte nicht mehr enthalten sind.

Je nachdem welches Informationssystem unterstützt wird, können die Infostationen eine Reihe von Standardfiltern bereitstellen, die dann nur noch vom Benutzer oder der Anwendung parametrisiert werden müssen. Ein Beispiel für einen solchen Standardfilter ist ein Filter, der Grafiken ausfiltert, deren Auflösung nicht auf dem Endgerät des Benutzers dargestellt werden kann. Die im Rahmen dieser Arbeit realisierte Architektur 5.1 des Vorübertragungssystems sieht es zusätzlich vor, dass der Benutzer oder die Anwendung sehr spezifische Filter zur Infostation schicken können, die dort eine Filterung der von der Infostation berechneten Rangliste durchführen.

## Kapitel 5

# Implementierung

Im Rahmen dieser Arbeit und der Diplomarbeit [Hegele 2001] ist eine vollständige Implementierung beider im vorherigen Kapitel beschriebenen Verfahren entstanden. Eine wichtige Randbedingung bei der Entwicklung des Vorabübertragungssystems HorSe (HOaRding-SystEm) war es, eine einfache Integration in die neXus-Plattform zu ermöglichen.

In diesem Kapitel wird zunächst die Architektur des Systems beschrieben. Dabei wird auch auf die Vorkehrungen eingegangen, die getroffen wurden, um die geforderte Integrationsmöglichkeit in neXus so weit wie möglich zu unterstützen. Anschließend wird die Umsetzung der Architektur vorgestellt, bevor am Ende des Kapitels verschiedene Einsatzbeispiele diskutiert werden.

### 5.1 Architektur

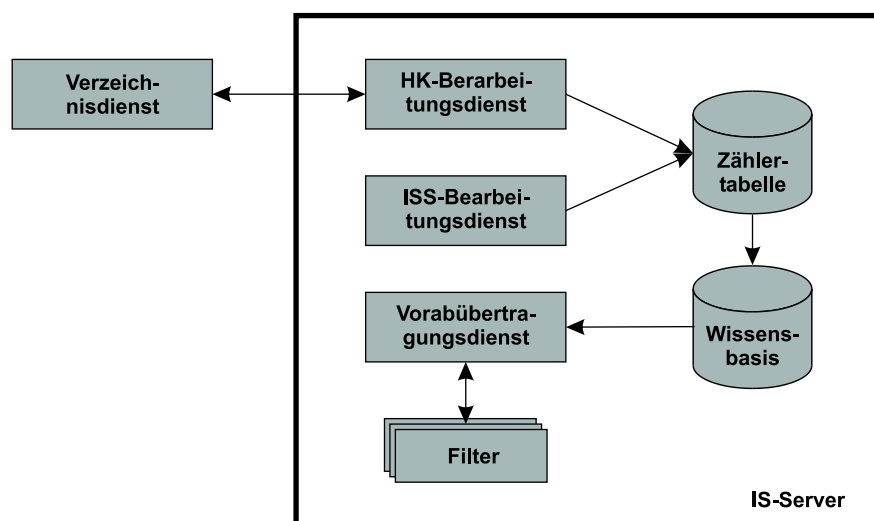
Das System besteht aus Infostation-Servern (ISS) und Hoarding-Klienten (HK). Zum Betrieb jeder Infostation ist genau ein Infostation-Server erforderlich. Um an einer Infostation Informationen vorab zu einem mobilen Endgerät übertragen zu können, muss auf diesem Endgerät der Hoarding-Klient verfügbar sein.

Zusätzlich sind zum Betrieb des Systems die in Abschnitt 3.3.4 erwähnten externen Dienste, Ortsinformationsdienst, Ereignisdienst und Verzeichnisdienst, erforderlich.

Diese Dienste wurden, so weit wie es für den Test des Systems erforderlich war, ebenfalls implementiert. Allerdings wurde davon ausgegangen, dass in einem realen Einsatz von HorSe, z.B. in neXus, diese externen Dienste durch die existierende Infrastruktur bereitgestellt werden. Entsprechend wurde die Implementierung der externen Dienste so einfach wie möglich gehalten.

### 5.1.1 Infostation-Server

Um einen möglichst hohen Parallelitätsgrad zu erreichen, wurden die einzelnen Aufgaben des Infostation-Servers auf getrennte Dienste verteilt. Daher besteht der Infostation-Server aus zwei Diensten zur Bearbeitung von Protokolldateien und einem Vorabübertragungsdienst. Neben diesen Diensten gibt es noch eine Wissensbasis, eine Zählertabelle, in der der Protokolldateizähler, die Zugriffszähler und die Besuchszähler verwaltet werden, sowie die in Abschnitt 4.6 erwähnten Filter (siehe Abbildung 5.1). Die Dienste zur Bearbeitung der Protokolldateien werden im Folgenden kurz Bearbeitungsdienste genannt. Nachfolgend werden die Aufgaben der einzelnen Dienste sowie ihre Schnittstellen beschrieben.



**Abbildung 5.1:** Komponenten der Infostation-Server.

## HK-Bearbeitungsdienst

Dieser Dienst ist für die Bearbeitung von Protokolldateien zuständig, die direkt von Hoarding-Klienten empfangen wurden. Er nutzt den externen Verzeichnisdienst, um festzustellen, an welche benachbarten Infostationen die einzelnen Einträge in der Protokolldatei weitergeleitet werden müssen. Die für jede Infostation jeweils relevanten Einträge werden in einer reduzierten Protokolldatei zusammengefasst, die anschließend an den ISS-Bearbeitungsdienst der entsprechenden Infostation gesendet wird. Schließlich sorgt der HK-Bearbeitungsdienst noch für die Aktualisierung der in der Zählertabelle enthaltenen Zähler.

Der HK-Bearbeitungsdienst bietet nur die folgende Methode nach außen an:

- `uploadLogfile (Logfile, Channels)`: Sie muss durch den mobilen Klienten aufgerufen werden, um die Bearbeitung einer Protokolldatei anzustoßen. Dabei ist `Logfile` die zu bearbeitende Protokolldatei und `Channels` eine Beschreibung der während der Erstellung der Protokolldatei abonnierten Informationskanäle einschließlich den jeweiligen Gewichten  $\bar{g}$  (siehe Abschnitt 4.6.1).

## ISS-Bearbeitungsdienst

Der ISS-Bearbeitungsdienst ist für die Bearbeitung von Protokolldateien zuständig, die von anderen Infostationen empfangen wurden. Im Gegensatz zum HK-Bearbeitungsdienst ist hier keine Weiterleitung von Einträgen der empfangenen Protokolldatei notwendig. Folglich ist der Bearbeitungsdienst ausschließlich für die Aktualisierung der Zähler in der Zählertabelle zuständig.

Der ISS-Bearbeitungsdienst bietet ebenfalls eine Methode `uploadLogfile (Logfile, Channels)` an. Der einzige Unterschied ist, dass sie durch den HK-Bearbeitungsdienst eines anderen Infostation-Servers aufgerufen wird und nicht durch einen mobilen Klienten.

## Vorabübertragungsdienst

Dieser Dienst trifft die Hoarding-Entscheidung und teilt dem mobilen Klienten das Ergebnis mit. Die Schnittstelle zu diesem Dienst umfasst folgende Methoden:

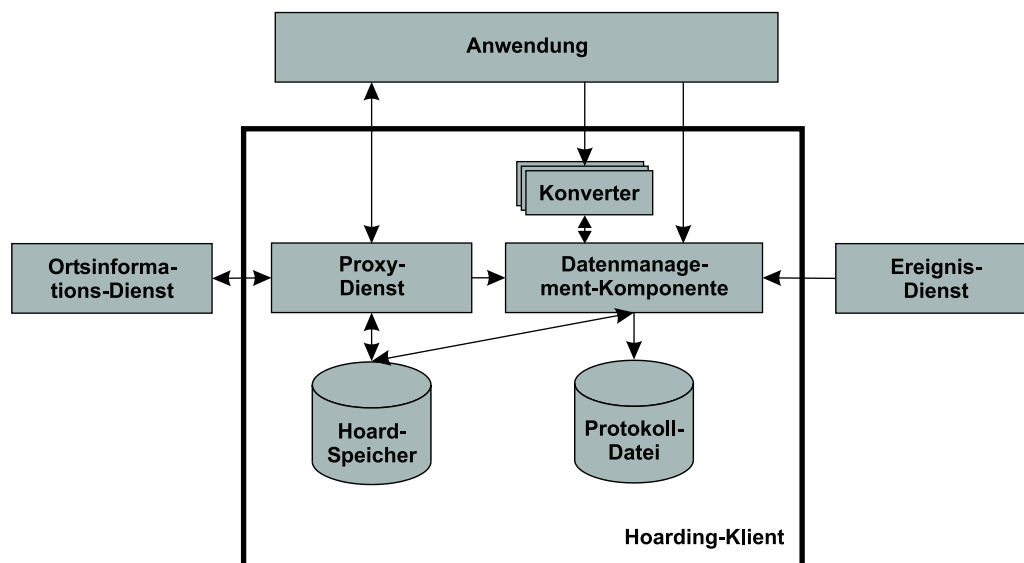
- `ObjectIds determineHoardObjects(MaxObjs, Channels, Filters, Maps)`: Mit dieser Methode kann ein mobiler Klient erfragen, welche Informationsobjekte vorab übertragen werden sollten. Der obligatorische Parameter `MaxObjs` gibt an, wieviele Informationsobjekte der Klient maximal speichern kann. Alle weiteren Parameter sind optional. `Channels` ist wiederum die Beschreibung der abonnierten Informationskanäle einschließlich den Gewichten. `Filters` enthält eine Liste der Filter, die auf die berechnete Rangliste angewendet werden sollen. `Maps` schließlich enthält die externen Besuchswahrscheinlichkeitskarten, die bei der Hoarding-Entscheidung zu berücksichtigen sind. Als Ergebnis wird eine Liste der Informationsobjekte zurückgeliefert, die auf dem mobilen Endgerät zukünftig am wahrscheinlichsten benötigt werden.
- `Channels getChannels()`: Diese Methode gibt eine Liste der durch die Infostationsinfrastruktur unterstützten Informationskanäle zurück. Sie wird zur Initialisierung der Hoarding-Klienten benötigt.
- `Filters getFilters()`: Durch den Aufruf dieser Methode kann ein mobiler Klient erfahren, welche Filter an der besuchten Infostation verfügbar sind.
- `HoardingArea getHoardingArea()`: Diese Methode liefert eine Beschreibung der Geometrie des Hoarding-Gebiets der Infostation sowie der darin enthaltenen Zonen. Diese Beschreibung wird für die Spezifikation externer Besuchswahrscheinlichkeitskarten benötigt.

Als einfache Erweiterung der Architektur eines Infostation-Servers könnte ein lokaler Cache angelegt werden, der häufig benötigte Informationsobjekte vorhält. Würden zusätzlich die Filter noch dahingehend erweitert, dass sie anstatt einer reinen Filterung eventuell auch eine Anpassung an die Fähigkeiten des Endgeräts vornehmen, wie beispielsweise eine Reduzierung der Auflösung eines Bildes, hätte man damit die Funktionalität von so genannten Transcoding-Proxies [Han et al. 1998, Mohan et al. 1999] in das Hoarding-System integriert.

### 5.1.2 Hoarding-Klient

Der Hoarding-Klient besteht aus einer Datenmanagement-Komponente und einem Proxy-Dienst. Des Weiteren verfügt jeder Hoarding-Klient über einen Hoard-Speicher,

in dem die vorab übertragenen Informationsobjekte gespeichert werden, und verwaltet die Protokolldatei zur Aufzeichnung der Informationszugriffe. Zusätzlich kann jeder Hoarding-Klient einen oder mehrere so genannte Konverter bereitstellen, die die in Abschnitt 4.4.6 erwähnten anwendungsspezifischen Schnittstellen zur Spezifikation externer Besuchswahrscheinlichkeitskarten realisieren. Abbildung 5.2 gibt einen Überblick über den Aufbau des Hoarding-Klienten.



**Abbildung 5.2:** Komponenten der Hoarding-Klienten.

### Datenmanagement-Komponente

Die Datenmanagement-Komponente ist zuständig für die Aufzeichnung der Informationszugriffe in der Protokolldatei sowie für die Koordination der Kommunikation mit den Infostation-Servern.

Der Datenmanagement-Komponente wird bei jedem Informationszugriff vom Proxy-Dienst der Bezeichner des referenzierten Objekts sowie der Ort, an dem der Zugriff auftrat, mitgeteilt. Diese Informationen hält die Datenmanagement-Komponente in der Protokolldatei fest.

Sobald das drahtlose LAN einer Infostation betreten wird, wird dies der Datenmanagement-Komponente durch den Ereignisdienst mitgeteilt. Dabei übermittelt der

Ereignisdienst auch die Adresse des Infostation-Servers. Daraufhin initiiert die Datenmanagement-Komponente die Protokollübertragung und schließlich die Informationsübertragung. Zu Beginn der Informationsübertragungsphase erfragt die Datenmanagement-Komponente die Bezeichner der vorab zu übertragenden Informationsobjekte vom Infostation-Server. Danach überprüft sie, ob eventuell schon einige dieser Informationsobjekte im Hoard-Speicher vorhanden sind und fordert alle nicht vorhandenen Objekte vom Informationssystem an.

Die Schnittstelle der Datenmanagement-Komponente besteht aus folgenden Methoden:

- `registerQuery(Position, ObjectID)`: Über diese Methode teilt der Proxy der Datenmanagement-Komponente für jeden Informationszugriff, die Position `Position`, an der er aufgetreten ist, sowie den Bezeichner `ObjectID` des referenzierten Informationsobjekts mit.
- `addMap(Map)`: Mittels dieser Methode kann eine externe Besuchswahrscheinlichkeitskarte `Map` direkt zur Datenmanagement-Komponente übermittelt werden.
- `addKonverter(Konverter)`: Mit dieser Methode meldet sich ein Konverter bei der Datenmanagement-Komponente an. Nach der Anmeldung wird der Konverter zu Beginn jeder Informationsübertragungsphase über die Geometrie des Hoarding-Gebiets sowie der Zonen informiert und nach externen Karten gefragt.
- `Filters getFilters()`: Diese Methode liefert eine Liste der verfügbaren Filter. Diese Liste beinhaltet sowohl die Standardfilter, die von den Infostation-Servern angeboten werden, als auch die auf dem mobilen Klienten implementierten spezifischen Filter.
- `setFilters(Filters)`: Mittels dieser Methode wird festgelegt, welche Filter bei der nächsten Vorabübertragung angewendet werden sollen.
- `Channels getChannels()`: Mit dieser Methode können die von den Infostation-Servern angebotenen Informationskanäle erfragt werden.



- `setChannels(Channels)`: Mit dieser Methode werden die abonnierten Informationskanäle einschließlich den Gewichtungen festgelegt. Jede Änderung der abonnierten Kanäle wird erst unmittelbar vor der nächsten Informationsübertragungsphase gültig.
- `WLAN_Entered(IS_Address)`: Diese Methode wird vom Ereignisdienst aufgerufen, um der Datenmanagement-Komponente mitzuteilen, dass das drahtlose LAN der Infostation, mit der Adresse `IS_Address` betreten wurde.
- `WLAN_Left(IS_Address)`: Diese Methode wird analog zu `WLAN_Entered` benutzt, um die Datenmanagement-Komponente über das Verlassen eines drahtlosen Netzwerks zu unterrichten.

## Proxy-Dienst

Der Proxy-Dienst nimmt die Informationsanfragen der Anwendungen entgegen und erfragt die Position, an der sie aufgetreten sind, beim externen Ortsinformationsdienst. Die entsprechenden Angaben werden an die Datenmanagement-Komponente weitergegeben. Zusätzlich überprüft der Proxy-Dienst für jede Informationsanfrage, ob das angeforderte Informationsobjekt im lokalen Hoard-Speicher vorhanden ist. Ist es vorhanden, so wird die Anfrage unmittelbar mit diesem Objekt beantwortet. Wenn das Objekt im Hoard-Speicher nicht gefunden wird, wird überprüft, ob eine Netzwerkverbindung zum Informationssystem hergestellt werden kann. Gelingt dies, wird das benötigte Objekt direkt beim Informationssystem angefordert und an die Anwendung zurückgegeben. Andernfalls wird eine Fehlermeldung („Hoard-Miss“) an die Anwendung übergeben.

Wie bei Proxy-Diensten üblich, müssen die Anwendungen so konfiguriert werden, dass sie ihre Anfragen nicht direkt an das Informationssystem sondern an den Proxy-Dienst senden. Da der Proxy-Dienst diese Anfragen direkt bearbeiten kann, sind keine expliziten Schnittstellen-Methoden notwendig, um den Proxy-Dienst zu nutzen.

## Konverter

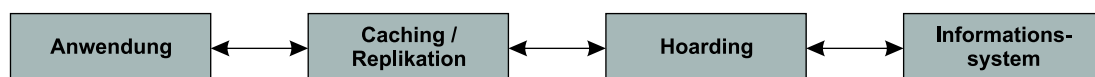
Die Konverter dienen als anwendungsspezifische Schnittstelle zur Spezifikation externer Besuchswahrscheinlichkeitskarten. Jeder Konverter meldet sich bei der Daten-

management-Komponente an. Danach erhält er zu Beginn jeder Informationsübertragungsphase eine Beschreibung des Hoarding-Gebiets und der Zonen. Von der Anwendung erhält er die externen Informationen über die zukünftige Bewegung des Benutzers im anwendungsspezifischen Format, beispielsweise die Beschreibung eines Pfads in Form eines Polygonzuges. Der Konverter rechnet diese anwendungsspezifischen Informationen in Besuchswahrscheinlichkeitskarten um und übermittelt die berechneten Karten an die Datenmanagement-Komponente. Für die gesamte Kommunikation mit der Datenmanagement-Komponente wird nur eine Methode benötigt:

- `Maps createMaps(HoardingArea)`: Diese Methode wird von der Datenmanagement-Komponente zu Beginn der Informationsübertragungsphase aufgerufen. Der Parameter `HoardingArea` beinhaltet die Beschreibung des Hoarding-Gebiets einschließlich der Zonen. Als Ergebnis werden externe Besuchswahrscheinlichkeitskarten geliefert, die die von der Anwendung spezifizierten Informationen über die zukünftige Bewegung des Benutzers repräsentieren.

### Koexistenz mit anderen Datenverwaltungsmechanismen

Der Hoarding-Klient ist so ausgelegt, dass seine Verwendung, bis auf die Umlenkung der Anfragen zum Proxy-Dienst, transparent ist für alle Anwendungen und Systeme, die auf das Informationssystem zugreifen. Damit ist ohne weitere Maßnahmen auch eine Koexistenz des Hoarding-Systems mit weiteren Datenverwaltungsmechanismen, wie zum Beispiel Caching- oder Replikationsmechanismen, möglich. Das Hoarding-System wird einfach den anderen Mechanismen nachgeschaltet (siehe Abbildung 5.3). Da die anderen Mechanismen nicht wissen, dass ihnen ein Hoarding-System nachgeschaltet ist, kann man diese Form der Koexistenz als transparente Koexistenz bezeichnen.



**Abbildung 5.3:** Transparente Koexistenz.

Der Nachteil dieser transparenten Koexistenz ist, dass möglicherweise die gleichen Informationsobjekte sowohl vom Hoarding-System als auch dem vorgeschalteten Datenverwaltungsmechanismus vorgehalten werden, wodurch Speicherplatz vergeudet

wird. Um dies zu vermeiden, dürfte nur ein gemeinsamer Datenbestand verwaltet werden. Dazu müsste aber auch sichergestellt werden, dass die weiteren Datenverwaltungsmechanismen nicht mit dem Hoarding-System um Speicherplatz konkurrieren. Daher wäre für eine solche synergetische Koexistenz eine übergeordnete Strategie zur Verwaltung des gemeinsamen Datenbestands notwendig.

### 5.1.3 Externe Dienste

Das Problem bei der Verwendung externer Dienste ist, dass deren Schnittstellen variieren können. Eine solche Veränderung der Schnittstellen ist zu erwarten, wenn ein Dienst *A* durch einen anderen Dienst *B* ersetzt werden soll. Veränderungen können aber auch durch die Weiterentwicklung eines bestimmten Dienstes auftreten. Insbesondere bei Verwendung der Dienste der neXus-Plattform ist zur Zeit noch mit solchen Veränderungen zu rechnen, da die Entwicklung der Dienste noch nicht abgeschlossen ist.

Um dieses Problem zu lösen, wird in HorSe das bekannte Adapter-Entwurfsmuster (Adapter Design Pattern, [Gamma et al. 1997]) eingesetzt. Dabei geht der Klient, in diesem Fall das Hoarding-System, von einer festen Schnittstelle aus. Um diese Schnittstelle auf die jeweils tatsächliche Schnittstelle des Dienstes abzubilden, werden so genannte Adapterklassen eingeführt. Diese Adapterklassen implementieren die vom Klienten erwartete Schnittstelle und erbringen die geforderte Funktionalität unter Verwendung des verfügbaren Dienstes. Daher hängt die Komplexität der Adapterklassen stark davon ab, wie weit die vom Dienst erwartete Funktionalität mit der erbrachten Funktionalität übereinstimmt. Durch die Verwendung des Adapter-Entwurfsmusters muss bei Änderung einer Dienst-Schnittstelle nur noch die jeweilige Adapterklasse angepasst werden. Das Hoarding-System selbst kann unverändert bleiben.

Nachfolgend werden die Schnittstellen der externen Dienste so beschrieben, wie sie vom Hoarding-System erwartet werden.

#### Ortsinformationsdienst

Der Ortsinformationsdienst wird vom Proxy-Dienst des Hoarding-Klienten benutzt, um die Position zu erfahren, an der eine bestimmte Informationsanfrage aufgetreten ist. Dazu muss der Ortsinformationsdienst nur die Methode `Position getPosition()` anbieten, die die aktuelle Position des mobilen Endgeräts liefert.

### Ereignisdienst

Der Ereignisdienst muss den mobilen Klienten mitteilen, wenn ein drahtloses LAN einer Infostation betreten und verlassen wird. Der Ereignisdienst bietet die folgende Methode an:

- `addClient(Client)`: Mit dieser Methode melden sich die Hoarding-Klienten beim Ereignisdienst an. Nach der Anmeldung werden sie über die beschriebenen Ereignisse informiert.

Um den Hoarding-Klienten den Eintritt von Ereignissen mitzuteilen, ruft der Ereignisdienst die Methoden `WLAN_Entered` beziehungsweise `WLAN_Left` der Datenmanagement-Komponente des Hoarding-Klienten auf.

### Verzeichnisdienst

Der Verzeichnisdienst wird vom HK-Bearbeitungsdienst des Infostation-Servers benutzt, um für eine Position  $p$  alle Infostationen zu erfragen, in deren Hoarding-Gebiet sich  $p$  befindet. Zur Schnittstelle des Verzeichnisdienstes gehören die folgenden Methoden:

- `registerIS(IS_Address, HoardingArea)`: Diese Methode wird von Infostationen benutzt, um sich beim Verzeichnisdienst anzumelden. Erst nach dieser Anmeldung ist die Infostation dem Verzeichnisdienst bekannt. Der Parameter `IS_Address` beinhaltet die Adresse der sich anmeldenden Infostation und `HoardingArea` ist eine Beschreibung des entsprechenden Hoarding-Gebiets.
- `IS_Addresses getIS_Addresses(Position)`: Mit dieser Methode können die Adressen aller Infostationen erfragt werden, deren Hoarding-Gebiet die Position `Position` abdecken.

## 5.2 Umsetzung

Die beschriebene Architektur wurde in Java 1.2 implementiert. Die Implementierung umfasst 91 Klassen. Für die entfernte Kommunikation zwischen den einzelnen Komponenten des Hoarding-Systems wird Java-RMI benutzt. Die Anwendungen kommunizieren über das Hypertext Transferprotokoll (HTTP, [Fielding et al. 1999]) mit dem Hoarding-System. Das Hoarding-System selbst benutzt ebenfalls HTTP, um auf das Informationssystem zuzugreifen.

Für die Kommunikation mit der Außenwelt wird HTTP eingesetzt, weil es dadurch möglich ist, alle gängigen Web-Browser unmittelbar durch das Hoarding-System zu unterstützen. Dies schließt alle Browser-basierten Anwendungen und Systeme, wie beispielsweise ViLis<sup>1</sup> (siehe Abschnitt 5.3.2), mit ein. Um das Hoarding-System nutzen zu können, muss in den Browser-Optionen lediglich der Port des Proxy-Dienstes des Hoarding-Klienten eingetragen werden. Ferner ergibt sich durch die Verwendung von HTTP ein einfach zu realisierender Migrationspfad für die Integration des Hoarding-Systems in die neXus-Plattform (siehe Abschnitt 5.3.3).

### 5.2.1 Infostation-Server

#### Initialisierung

Um für die anderen Komponenten erreichbar zu sein, muss sich jeder Infostation-Server bei einem RMI-Register registrieren. Die Adresse des Infostation-Servers setzt sich aus dem Namen des Rechners, auf dem das Register läuft, dem Port, an das das Register gebunden ist und einem eindeutigen Bezeichner der Infostation zusammen.

Den Namen des Rechners, auf dem das für einen Infostation-Server zuständige Register läuft, sowie den entsprechenden Port, liest der Infostation-Server beim Start aus einer Konfigurationsdatei. Diese Konfigurationsdatei enthält noch eine Reihe weiterer Parameter, die für den Betrieb des Infostation-Servers notwendig sind, wie zum Beispiel die maximale Länge der Zugriffswahrscheinlichkeitstabellen und die Namen von Initialisierungsdateien. Alle diese Parameter können beim Start der Infostation vom Betreiber über die graphische Oberfläche nochmals verändert werden.

---

<sup>1</sup><http://vit.informatik.uni-stuttgart.de>

Zusätzliche Initialisierungsdateien beschreiben unter anderem die zu verwaltenden Informationskanäle, den initialen Inhalt der Zugriffswahrscheinlichkeitstabellen und Besuchswahrscheinlichkeitskarten sowie das Hoarding-Gebiet der Infostation. Wo es die Komplexität der beschriebenen Daten erforderlich machte, wurden diese in einer XML-basierten Syntax beschrieben.

Tag	Bedeutung	Tag	Bedeutung
is_pos_x	x-Koordinate des Mittelpunkts des Hoarding-Gebiets	y	y-Koordinate eines Atoms
is_pos_y	y-Koordinate des Mittelpunkts des Hoarding-Gebiets	area	Zuordnung einer Menge von Atomen zu einer Zone
radius	Radius des Hoarding-Gebiets	from_x	x-Koord. der am weitesten links liegenden Atome eines Bereichs
sq_length	Kantenlänge der Atome	from_y	y-Koord. der am weitesten unten liegenden Atome eines Bereichs
zone	Beschreibung einer Zone	to_x	x-Koord. der am weitesten rechts liegenden Atome eines Bereichs
id	Bezeichner einer Zone	to_y	y-Koord. der am weitesten oben liegenden Atome eines Bereichs
patch	Zuordnung eines Atoms zu einer Zone		
x	x-Koordinate eines Atoms		

**Tabelle 5.1:** XML-Tags zur Beschreibung eines Hoarding-Gebiets und ihre Bedeutung.

Bei der Beschreibung von Hoarding-Gebieten wird vereinfachend davon ausgegangen, dass diese immer eine kreisförmige Fläche haben. Im erweiterten Verfahren ist jedes Hoarding-Gebiet in Atome unterteilt. Diese Atome sind gleich große Quadrate mit einer konfigurierbaren Kantenlänge und wurden für die Implementierung eingeführt, um eine einheitliche Beschreibung der Zonen zu ermöglichen. Jede Zone besteht aus einem oder mehreren dieser Atome. Tabelle 5.1 enthält die XML-Tags, die für die Beschreibung eines Hoarding-Gebiets und seiner Zonen notwendig sind. In Abbildung 5.4 ist ein Beispiel für die Beschreibung eines Hoarding-Gebiets angegeben. Atome, die durch die Beschreibungsdatei keiner Zone zugeordnet werden, werden in einer so genannten Restzone zusammengefasst.

### Benutzungsoberfläche

Das Hauptfenster der Benutzungsoberfläche (siehe Abbildung 5.5(a)) zeigt dem Betreiber eines Infostation-Servers alle Parametereinstellungen sowie die verwalteten In-

<pre> &lt;is_pos_x&gt; 2355.0 &lt;/is_pos_x&gt; &lt;is_pos_y&gt; 5643.0 &lt;/is_pos_y&gt; &lt;radius&gt; 100 &lt;/radius&gt; &lt;sq_length&gt; 20.0 &lt;/sq_length&gt;  &lt;zone&gt;   &lt;id&gt; Kiosk &lt;/id&gt;   &lt;patch&gt;     &lt;x&gt; -3 &lt;/x&gt;     &lt;y&gt; -1 &lt;/y&gt;   &lt;/patch&gt; &lt;/zone&gt; </pre>	<pre> &lt;zone&gt;   &lt;id&gt; Bibliothek &lt;/id&gt;   &lt;patch&gt;     &lt;x&gt; -4 &lt;/x&gt;     &lt;y&gt; 1 &lt;/y&gt;   &lt;/patch&gt;   &lt;area&gt;     &lt;from_x&gt; -5 &lt;/from_x&gt;     &lt;from_y&gt; 2 &lt;/from_y&gt;     &lt;to_x&gt; -3 &lt;/to_x&gt;     &lt;to_y&gt; 2 &lt;/to_y&gt;   &lt;/area&gt; &lt;/zone&gt; </pre>
---	--

**Abbildung 5.4:** Beschreibung eines Hoarding-Gebiets.

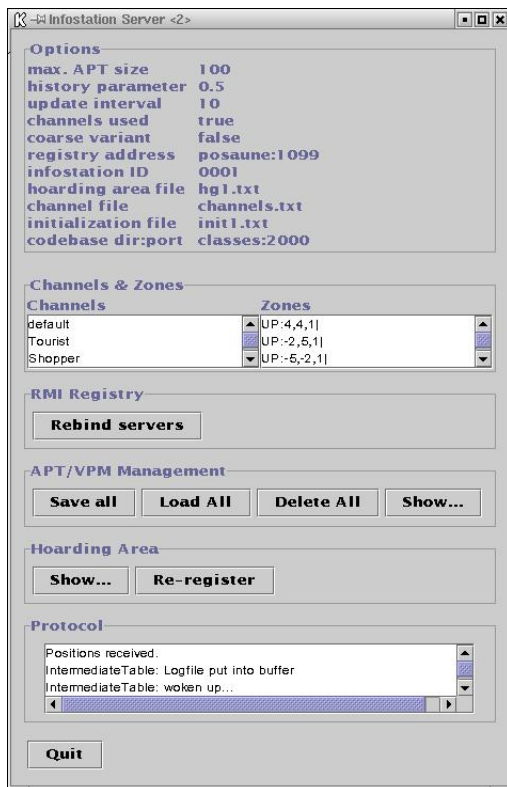
formationskanäle und die Zonen des Hoarding-Gebiets an. Falls die Verbindung zum RMI-Register unterbrochen wurde, kann der Betreiber über den „Rebind“-Knopf versuchen, diese Verbindung wieder herzustellen, ohne den Server neu zu starten. Ferner können Zugriffswahrscheinlichkeitstabellen und Besuchswahrscheinlichkeitskarten jederzeit gespeichert, geladen oder gelöscht werden. Ihr Inhalt kann auf Wunsch in einem eigenen Fenster angezeigt werden. Das Gleiche gilt für das Hoarding-Gebiet und seine Zonen.

## 5.2.2 Hoarding-Klient

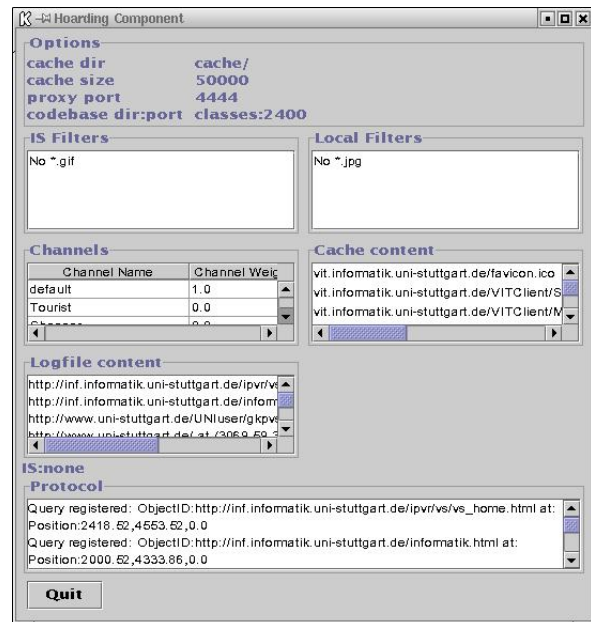
### Initialisierung

Wie beim Start eines Infostation-Servers werden auch beim Start eines Hoarding-Klienten die Voreinstellungen der verschiedenen Parameter aus einer Konfigurationsdatei gelesen. Darin ist zum Beispiel das Verzeichnis angegeben, unter dem der Hoard-Speicher zu finden ist, sowie die maximale Größe des Hoard-Speichers und der Port, an dem der Proxy-Dienst Anfragen der Anwendungen entgegen nehmen soll.

Die in der Datei enthaltenen Voreinstellungen können ebenfalls beim Start eines Hoarding-Klienten über die graphische Benutzeroberfläche geändert werden. Sind bereits beim Start des Hoarding-Klienten Informationsobjekte im Verzeichnis des Hoard-Speichers enthalten, beispielsweise aus früheren Läufen, so kann darauf unmittelbar nach dem Start zugegriffen werden, ohne dass eine Informationsübertragung durchgeführt wird.



(a) Infostation-Server.



(b) Hoarding-Klient.

**Abbildung 5.5:** Benutzungsoberflächen der Hoarding-System-Komponenten.

### Benutzungsoberfläche

Neben den Voreinstellungen zeigt das Hauptfenster der Benutzungsoberfläche des Hoarding-Klienten die von den Infostationen angebotenen Filter, sowie die lokal verfügbaren Filter an (siehe Abbildung 5.5(b)). Der Benutzer kann aus den angezeigten Filtern eine beliebige Anzahl auswählen, die bei der nächsten Informationsübertragung angewendet werden sollen. Zu Testzwecken wurde auf den Infostation-Servern ein Filter für Bilder im gif-Format implementiert und auf der Seite der Hoarding-Klienten ein Filter für Bilder im jpg-Format, der gegebenenfalls zum Infostation-Server übertragen wird. Wie der Transfer des Filter-Codes durchgeführt wird, ist detailliert in [Hegele 2001] beschrieben.

Ferner kann der Benutzer aus einer Liste der unterstützten Informationskanäle, diejenigen auswählen, die er abonnieren möchte und die jeweiligen Gewichte spezifizieren.



Schließlich wird noch eine Reihe von Statusinformationen angezeigt, wie zum Beispiel der Inhalt des Hoard-Speichers und der Protokolldatei sowie Informationen über eine mögliche Verbindung zu einer Infostation.

### 5.2.3 Externe Dienste

Um das Hoarding-System testen zu können, wurden in der Diplomarbeit [Hegele 2001] alle benötigten externen Dienste in einer einfachen Form implementiert. Da diese Dienste nur zu Testzwecken entworfen wurden und in einem realen Einsatz des Hoarding-Systems durch Dienste der jeweiligen Infrastruktur ersetzt werden, lag der Schwerpunkt beim Entwurf auf einer einfachen Realisierungsmöglichkeit.

#### Ortsinformationsdienst

Der implementierte Ortsinformationsdienst basiert auf der Verwendung eines GPS-Sensors sowie der in der Studienarbeit [Hässler 2000] entwickelten Java-Klassen zur Abfrage eines solchen Sensors. Bei jedem Informationszugriff fragt der Proxy-Dienst des Hoarding-Klienten über diese Klassen die aktuelle geographische Position des mobilen Endgeräts beim GPS-Sensor ab.

Da GPS-Sensoren aus technischen Gründen innerhalb von Gebäuden nicht eingesetzt werden können, wurden für die meisten Tests die geographischen Positionen nicht vom GPS-Sensor erfragt, sondern aus einer Datei, in der zuvor typische Positionen gespeichert wurden, ausgelesen.

#### Ereignisdienst

Der implementierte Ereignisdienst nutzt das Simple Network Management Protokoll (SNMP, [Case et al. 1990]), um bei den Zugangspunkten (Access Points) der drahtlosen Funknetze aller Infostationen die IP-Adressen der mobilen Endgeräte zu erfragen, die sich im jeweiligen Empfangsbereich befinden. Sobald ein neues Gerät hinzukommt, wird diesem über den Aufruf der `WLAN_Entered` Methode die Adresse des zugehörigen Infostation-Servers mitgeteilt. Die Adressen der Zugangspunkte sowie der zugehörigen Infostation-Server liest der Ereignisdienst aus einer Konfigurationsdatei.

Eine alternative Realisierung wäre gewesen, direkt an der Netzwerkkarte des mobilen Endgeräts abzufragen, ob und wenn ja mit welchen Zugangspunkt sie assoziiert ist. Die Adresse des Zugangspunkts hätte dann einfach auf die Adresse der Infostation abgebildet werden können. Leider scheiterte dieser Ansatz daran, dass die Hersteller der Netzwerkkarten die Treiberschnittstellen ihrer Karten nicht offen legen, weshalb auf Informationen über den assoziierten Zugangspunkt nicht zugegriffen werden kann.

Eine weitere Möglichkeit wäre, die Infostationen periodisch Nachrichten (Beacons) innerhalb ihres LANs rundsenden zu lassen, die ihre Adresse beinhalten. Sobald ein Endgerät dann in den Empfangsbereich eines drahtlosen LANs kommen würde, würde es die Nachricht der jeweiligen Infostation empfangen. Das Problem bei diesem Ansatz war, dass ohne die Verwendung von Mobile IP [Perkins 1996], das gegenwärtig an der Fakultät Informatik noch nicht installiert ist, alle Zugangspunkte zum gleichen Subnetz gehören müssen. Sonst könnte ein Endgerät nach dem Wechsel des Zugangspunkts und damit des Subnetzes nicht mehr über seine IP-Adresse angesprochen werden. Wenn jedoch alle Zugangspunkte zum gleichen Subnetz gehören, würde ein mobiles Endgerät die Nachrichten aller Infostationen empfangen, sobald irgendein Zugangspunkt erreichbar ist.

### **Verzeichnisdienst**

Der Verzeichnisdienst wurde als zentralisierter Dienst implementiert, der die Adressen der Infostation-Server und die Beschreibungen der zugehörigen Hoarding-Gebiete in einer Hash-Tabelle verwaltet.

Da die implementierten externen Dienste genau die vom Hoarding-System erwarteten Schnittstellen realisieren, war die Implementierung der Adapterklassen sehr einfach. Sie müssen jeweils nur die Anfragen des Hoarding-Systems an den jeweiligen Dienst und umgekehrt die Nachrichten von den Diensten an das Hoarding-System weiterleiten.

## **5.3 Einsatzbeispiele**

In diesem Abschnitt wird beschrieben, wie das entwickelte Hoarding-System in Verbindung mit einigen typischen Systemen eingesetzt werden kann. Bei den ersten bei-

den Systemen, dem Web-Browser und ViLis, konnte der Einsatz auch praktisch durchgeführt werden, während die Integration in neXus nur als Ausblick dienen kann, wie HorSe zukünftig in neXus integriert werden kann, sobald funktionsfähige Prototypen der Plattform zur Verfügung stehen.

### 5.3.1 Web-Browser

Der Verwendung von HorSe zur Unterstützung von drahtlosem Surfen im WWW ist sehr einfach. Es muss lediglich in den Einstellungen des Web-Browsers die Verwendung eines Proxy-Dienstes spezifiziert werden und der Hoarding-Klient als Proxy-Dienst angegeben werden.

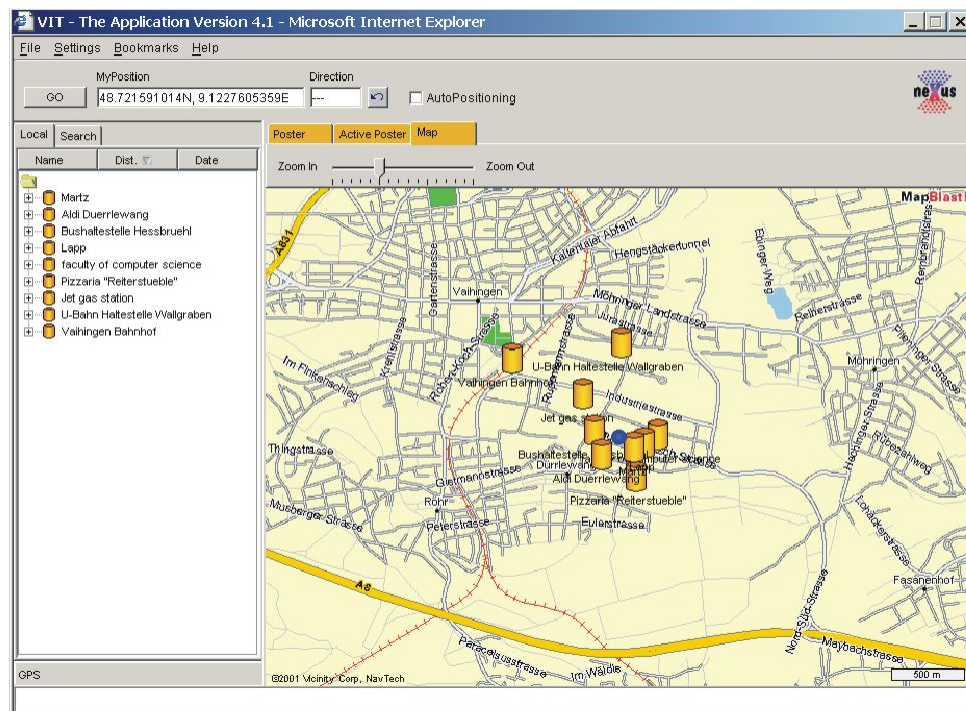
Obwohl der in HorSe integrierte HTTP-Parser nur GET-Anfragen bearbeitet, funktioniert HorSe im Zusammenspiel mit einem Web-Browser sehr gut. Selbst komplexe WWW-Seiten, beispielsweise solche mit Rahmen, werden durch den Proxy-Dienst des Hoarding-Klienten korrekt bearbeitet.

### 5.3.2 ViLis

ViLis ist ein System, das im Rahmen eines Studienprojekts in der Abteilung Verteilte Systeme, entwickelt wurde. Es realisiert die in [Leonhardi et al. 1999] vorgestellte Metapher einer virtuellen Litfaßsäule. Bei einer virtuellen Litfaßsäule handelt es sich um ein Objekt, das nur als Datenstruktur im Rechner existiert, dem aber eine feste geographische Position in der realen Welt zugeordnet ist. Zusätzlich ist jeder virtuellen Litfaßsäule ein Sichtbarkeitsbereich in der realen Welt zugeordnet. Sobald ein Benutzer des ViLis-Systems diesen Sichtbarkeitsbereich betritt, wird ihm die jeweilige Litfaßsäule angezeigt und er kann die mit dieser Litfaßsäule verbundene Information einsehen (siehe Abbildung 5.6).

Der Einsatz von HorSe zur Unterstützung von ViLis gestaltet sich ähnlich einfach wie die Zusammenarbeit mit einem Web-Browser. Das liegt primär daran, dass ViLis in einen Web-Browser integriert ist und sich daher die Anfragen des ViLis-Systems aus Sicht des Hoarding-Klienten nicht von denen eines Web-Browsers unterscheiden.

Ein Vorteil von ViLis ist, dass jede verfügbare Information einer virtuellen Litfaßsäule fest zugeordnet ist und nur innerhalb des Sichtbarkeitsbereichs dieser Litfaßsäule auf



**Abbildung 5.6:** Benutzungsoberfläche des ViLis-Systems.

die Information zugegriffen werden kann. Bei ViLis handelt es sich also um ein inhärent ortsbezogenes System, was besonders hohe Trefferraten erwarten lässt und auch die Angabe initialer Zugriffswahrscheinlichkeitstabellen erleichtert.

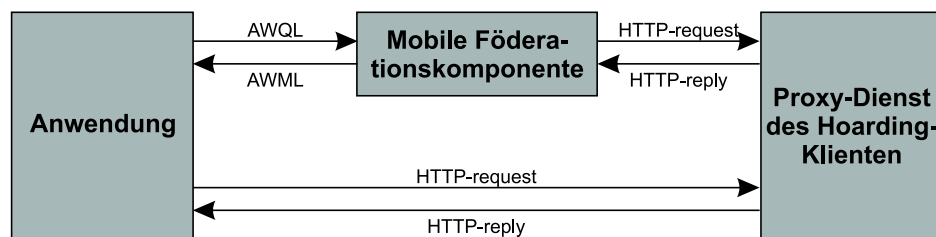
### 5.3.3 neXus

Die zentrale Komponente der neXus-Plattform ist ein Modell der realen Welt, das um virtuelle Objekte erweitert ist [Hohl et al. 1999]. Die von der Plattform unterstützten Anwendungen greifen auf dieses so genannte Augmented World Model über eine spezielle Anfragesprache, die Augmented World Query Language (AWQL) zu. Das Modell selbst ist in einer speziellen Beschreibungssprache, der Augmented World Modeling Language (AWML), repräsentiert. Da in AWQL sehr komplexe Anfragen möglich sind, die sich auch noch auf verschiedene im Augmented World Model verfügbare Informationsquellen beziehen können, ist für die Bearbeitung dieser Anfragen in der neXus-Plattform eine globale Föderationskomponente vorgesehen.

Allerdings ist diese globale Föderationskomponente zu umfangreich, um ihre Funk-

tionalität vollständig auf einem mobilen Endgerät zu replizieren. Daher kann in einem entkoppelten Betrieb nur ein Teil der Funktionalität der Föderationskomponente erbracht werden. Wie diese Teilfunktionalität erbracht wird und welchen Umfang sie haben wird, ist noch Gegenstand der aktuellen Forschung in der neXus-Forscherguppe. Fest steht bislang, dass es eine spezielle, mobile Föderationskomponente geben wird, die diese Teilfunktionalität in einem entkoppelten Betrieb erbringen wird.

AWQL-Anfragen können nicht direkt an HorSe weitergeleitet werden, da der Proxy-Dienst des Hoarding-Klienten HTTP-Anfragen erwartet. Die derzeitige Planung ist jedoch, dass die mobile Föderationskomponente, die AWQL-Anfragen in HTTP-Anfragen umsetzt und auch aus HTTP-Antworten gültige AWML-Beschreibungen generieren kann. Damit könnte HorSe unverändert in neXus-integriert werden. Ferner wird es auch einfache neXus-Anwendungen geben, deren Kommunikation direkt über HTTP abgewickelt wird. Diese Anwendungen können HorSe selbstverständlich auch problemlos nutzen. In Abbildung 5.7 ist die Integration von HorSe in neXus nochmals im Überblick dargestellt.



**Abbildung 5.7:** Integration des Hoarding-Systems in neXus.

Die vom Hoarding-System benötigten externen Dienste werden alle in die neXus-Plattform integriert sein. Die Funktionalität des Verzeichnisdienstes kann direkt durch das Augmented World Model in Verbindung mit der globalen Föderation erbracht werden. Die Infostationen müssen dazu, zusammen mit einer Beschreibung ihrer Hoarding-Gebiete, lediglich als Objekte in das Modell eingebracht werden. Die Infostationen sind bereits schon als Klasse in der Klassenhierarchie zur Modellierung der Augmented World in neXus vorgesehen [Messmer 2001, Hegele 2001]. In [Hegele 2001] ist auch die Anfrage an das Augmented World Model in AWQL angegeben, die alle Infostationen liefert, deren Hoarding-Gebiete eine bestimmte Position abdecken.

## Kapitel 6

# Simulationsumgebung

In diesem Kapitel wird die Simulationsumgebung beschrieben, die für die Evaluierung des Vorabübertragungsverfahrens (siehe Kapitel 7) benutzt wurde [Kubach et al. 2001]. Zunächst wird erläutert, warum es erforderlich war, für diese Evaluierung eine spezielle Simulationsumgebung zu entwickeln und wo die Defizite existierender Modelle, insbesondere bei der Simulation eines mobilen, ortsabhängigen Informationszugriffs, liegen. Anschließend wird das Simulationsmodell sowie die in der Studienarbeit [Hegele 2000] entstandene Implementierung der Simulationsumgebung vorgestellt. Aufgrund ihrer Flexibilität ist die Simulationsumgebung für die Evaluierung beliebiger ortsbezogener Informationssysteme geeignet.

### 6.1 Problemstellung

Zwar existiert eine Reihe von Modellen sowohl zur Modellierung der Mobilität von Benutzern als auch für die Modellierung von Informationszugriffen. Allerdings wird in jedem dieser Modelle jeweils nur einer der beiden Aspekte, das heißt entweder die Mobilität oder die Informationszugriffe, betrachtet. Daher ist es mit diesen Modellen nicht möglich, eine Beziehung zwischen der Mobilität und den Informationszugriffen herzustellen und somit festzustellen, auf welche Informationen an einem bestimmten Ort zugegriffen wird oder an welchen Orten eine bestimmte Information benötigt wird. Aber gerade solche Fragestellungen sind bei der Evaluierung ortsbezogener Systeme, wie dem in dieser Arbeit vorgestellten Verfahren, von besonderer Bedeutung.

Des Weiteren existieren bislang keine Modelle, die die in ortsbezogenen Informationssystemen vorhandenen Beziehungen von Informationen und geographischen Orten nachbilden. Da aber ortsbezogene Systeme gerade diese Beziehungen ausnutzen, ist es für ihre Evaluierung unerlässlich, diese Beziehungen zu modellieren, das heißt festzulegen, welche Informationen an bestimmten Orten überhaupt verfügbar sind und welche Informationen wo bevorzugt werden.

Um existierende Mobilitätsmodelle und Informationszugriffsmodelle in ein übergeordnetes Modell zu integrieren, das darüber hinaus auch noch die Beziehungen zwischen Orten und Informationen berücksichtigt, musste ein neues Simulationsmodell entwickelt werden.

## 6.2 Verwandte Arbeiten

Nachfolgend werden existierende Mobilitätsmodelle und Informationszugriffsmodelle beschrieben. Einige von ihnen werden im entwickelten Simulationsmodell eingesetzt, um bestimmte Teilaspekte des Benutzerverhaltens zu modellieren.

### 6.2.1 Mobilitätsmodelle

Die existierenden Mobilitätsmodelle können in die drei Klassen, Random-Walk-Modelle, Bewegungsmuster-basierte Modelle und Bewegungsbedarfsmodelle, eingeteilt werden.

#### Random-Walk-Modelle

Ein sehr gebräuchliches und einfaches Mobilitätsmodell ist das Random-Walk-Modell. In diesem Modell wird die Geschwindigkeit und die Richtung der zukünftigen Bewegung eines Benutzers zu bestimmten diskreten Zeitpunkten auf zufällige Werte festgelegt, die dann bis zu einer erneuten Festlegung gültig sind. Solche Random-Walk-Modelle werden zum Beispiel in [Zonoozi & Dassanayake 1997] und [de Nitto Personè et al. 1998] eingesetzt.

### **Bewegungsmuster**

Andere Autoren, wie zum Beispiel Liu & Maguire Jr. (1995), behaupten, dass die Bewegung eines Benutzers nicht rein zufällig ist, sondern auch von regelmäßigen Bewegungsmustern getrieben wird. In [Liu et al. 1998] wird beispielsweise eine kreisförmige Bewegung als häufig auftretendes Muster beschrieben. Bei dieser kreisförmigen Bewegung kehren die Benutzer, nachdem sie eine Reihe von Orten besucht haben, stets wieder zu ihrem Ausgangspunkt zurück. Ein Beispiel für eine solche kreisförmige Bewegung sind Berufspendler, die morgens zu ihrem Arbeitsplatz fahren und am Abend wieder nach Hause zurückkehren. Dabei können auch mehrere Kreise verschachtelt sein. Wenn beispielsweise der Berufspendler von seinem Arbeitsplatz aus andere Orte besucht und danach wieder zu seinem Arbeitsplatz zurückkehrt, entsteht dadurch ein neuer Kreis, der in dem Kreis, der die ganztägige Bewegung beschreibt, integriert ist.

Teilweise werden Random-Walk-Modelle und Bewegungsmuster kombiniert. Auch das in [Liu et al. 1998] beschriebene Modell ist ein solches kombiniertes Modell. Es handelt sich um ein hierarchisches Modell, das zwei Ebenen unterscheidet. Die obere Ebene beschreibt die Bewegung zwischen Zellen und die untere Ebene modelliert die Bewegung innerhalb einer Zelle. Die obere Ebene basiert auf Bewegungsmustern, während in der unteren Ebene ein Random-Walk-Modell verwendet wird.

### **Bewegungsbedarf**

In Verkehrsplanungssystemen, wie [INRO Consultants Inc. 2000], wird Mobilität oft dadurch modelliert, dass der Bewegungsbedarf Einzelner oder einer Gruppe von Personen nachgebildet wird. Modelle dieser Art versuchen den Anlass einer Bewegung nachzubilden. Man kann zwei Arten von Bewegungsbedarfsmodellen (travel demand models) unterscheiden [Oppenheim 1995]: die beschreibenden Modelle (descriptive travel demand models) und die verhaltensorientierten Modelle (behavioral travel demand models).

Die beschreibenden Modelle stützen sich auf Start-Ziel-Matrizen, die Informationen darüber enthalten, wieviele Menschen sich zwischen jedem möglichen Paar von Start- und Ziel-Orten bewegen. Für die Erstellung der Matrizen ist in der Regel eine große Menge von Daten, die die Bewegungen der Menschen beschreiben, erforderlich. Diese



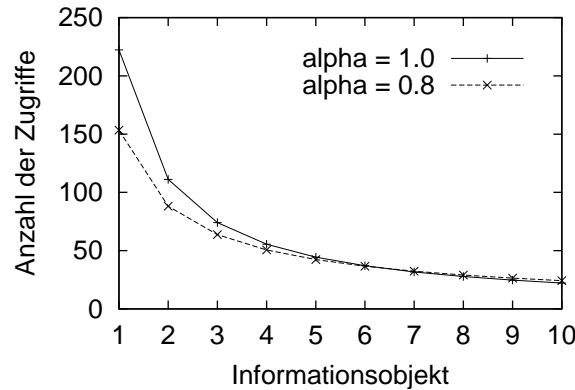
Daten müssen meist in aufwendigen Erhebungen gesammelt werden. Die beschreibenden Modelle können nur eingesetzt werden, um ein bekanntes, beobachtetes Bewegungsverhalten im Rechner nachzubilden. Sie können nicht eingesetzt werden, wenn keine Daten zur Erstellung der Start-Ziel-Matrix vorliegen, weil sie den Bedarf für eine Bewegung nicht plausibel erklären.

In den verhaltensorientierten Modellen [Scourias & Kunz 1999] wird die Mobilität der Benutzer als Ergebnis verschiedener, individueller Entscheidungen der Benutzer betrachtet. Entsprechend versuchen diese Modelle nachzubilden, welche Entscheidungsmöglichkeiten die Benutzer während ihrer Bewegung bezüglich des Bewegungsablaufs haben und wie sie ihre Entscheidungen treffen. Sehr häufig wird davon ausgegangen, dass jede Bewegung beziehungsweise Reise eines Benutzers einem bestimmten Ziel dient und dass der Benutzer versucht dieses Ziel mit dem geringsten Aufwand zu erreichen (Prinzip des geringsten Aufwands).

### **6.2.2 Informationszugriffsmodelle**

Viele Arbeiten zur Modellierung von Informationszugriffen sind im Umfeld des World Wide Web angesiedelt. Ein wichtiger Aspekt dieser Modelle ist die Popularität einzelner Informationsobjekte, das heißt die Verteilung der Informationszugriffe auf die einzelnen Informationsobjekte.

In einer Reihe von Arbeiten [Glassman 1994, Crovella & Bestavros 1996, Nishikawa et al. 1998] wird bei der Simulation von Informationszugriffen davon ausgegangen, dass die Zugriffswahrscheinlichkeiten auf die einzelnen Informationsobjekte Zipf-verteilt sind. Durch die Analyse umfangreicher Zugriffsprotokolle auf sechs verschiedene hochfrequentierte WWW-Server konnte in [Breslau et al. 1999] gezeigt werden, dass die Verteilung der Informationszugriffe auf die einzelnen Informationsobjekte eines WWW-Servers Zipf-ähnlich ist. In der vorliegenden Arbeit wurde ein existierendes ortsbezogenes Informationssystem analysiert und dadurch verifiziert, dass auch in ortsbezogenen Informationssystemen eine Zipf-ähnliche Verteilung beobachtet werden kann (siehe Abschnitt 7.2). Da die Zipf-ähnliche Verteilung die reale Verteilung der Zugriffswahrscheinlichkeiten am besten nachbildet, wurde sie auch für die in Kapitel 7 beschriebenen Simulationen verwendet. Im Folgenden wird die Zipf-ähnliche Verteilung daher kurz näher beschrieben.



**Abbildung 6.1:** Zipf-Verteilung und Zipf-ähnliche Verteilung.

Angewandt auf die Zugriffswahrscheinlichkeiten von Informationsobjekten, wie zum Beispiel WWW-Seiten, besagt das Gesetz von Zipf [Zipf 1949], dass die Zugriffswahrscheinlichkeit auf ein Informationsobjekt umgekehrt proportional zu seinem Beliebtheitsrang ist, das heißt wird auf das beliebteste Objekt mit der Wahrscheinlichkeit  $c$  zugegriffen, so wird auf das  $i$ -beliebteste Objekt mit der Wahrscheinlichkeit  $\frac{c}{i}$  zugegriffen.

Die erwähnten Untersuchungen von Breslau et al. (1999) haben jedoch gezeigt, dass die tatsächlichen Zugriffswahrscheinlichkeiten nur einer so genannten Zipf-ähnlichen Verteilung folgen, bei der auf das  $i$ -beliebteste Objekt mit einer Wahrscheinlichkeit proportional zu  $\frac{1}{i^\alpha}$  zugegriffen wird, wobei für  $\alpha$  meist Werte beobachtet wurden, die ein wenig kleiner als eins sind. Je kleiner der Wert von  $\alpha$  wird, umso geringer ist die Konzentration der Informationszugriffe auf die beliebtesten Informationsobjekte.

**DEFINITION 6.1 (ZIPF-ÄHNLICHE VERTEILUNG)**

Sei eine Menge von Informationsobjekten gegeben. Dann sind die Zugriffswahrscheinlichkeiten auf diese Objekte Zipf-ähnlich verteilt mit dem Parameter  $\alpha$ , wenn auf das  $i$ -beliebteste Informationsobjekt mit der Wahrscheinlichkeit

$$p(i) = \frac{\Omega_\alpha}{i^\alpha}$$

zugegriffen wird.

Sind  $n$  Informationsobjekte in der betrachteten Menge enthalten, so folgt aus der Randbedingung  $\sum_{i=1}^n p(i) = 1$  unmittelbar  $\Omega_\alpha = \frac{1}{\sum_{i=1}^n \frac{1}{i^\alpha}}$ . Für den Spezialfall  $\alpha = 1$  entspricht die

Zipf-ähnliche Verteilung der Zipf-Verteilung gemäß dem Gesetz von Zipf. In Abbildung 6.1 sind die Zugriffszahlen, die man mit einer echten Zipf-Verteilung und einer Zipf-ähnlichen Verteilung mit  $\alpha = 0.8$  bei 1000 Informationszugriffen auf 50 verschiedene Informationsobjekte erhält, gegenübergestellt.

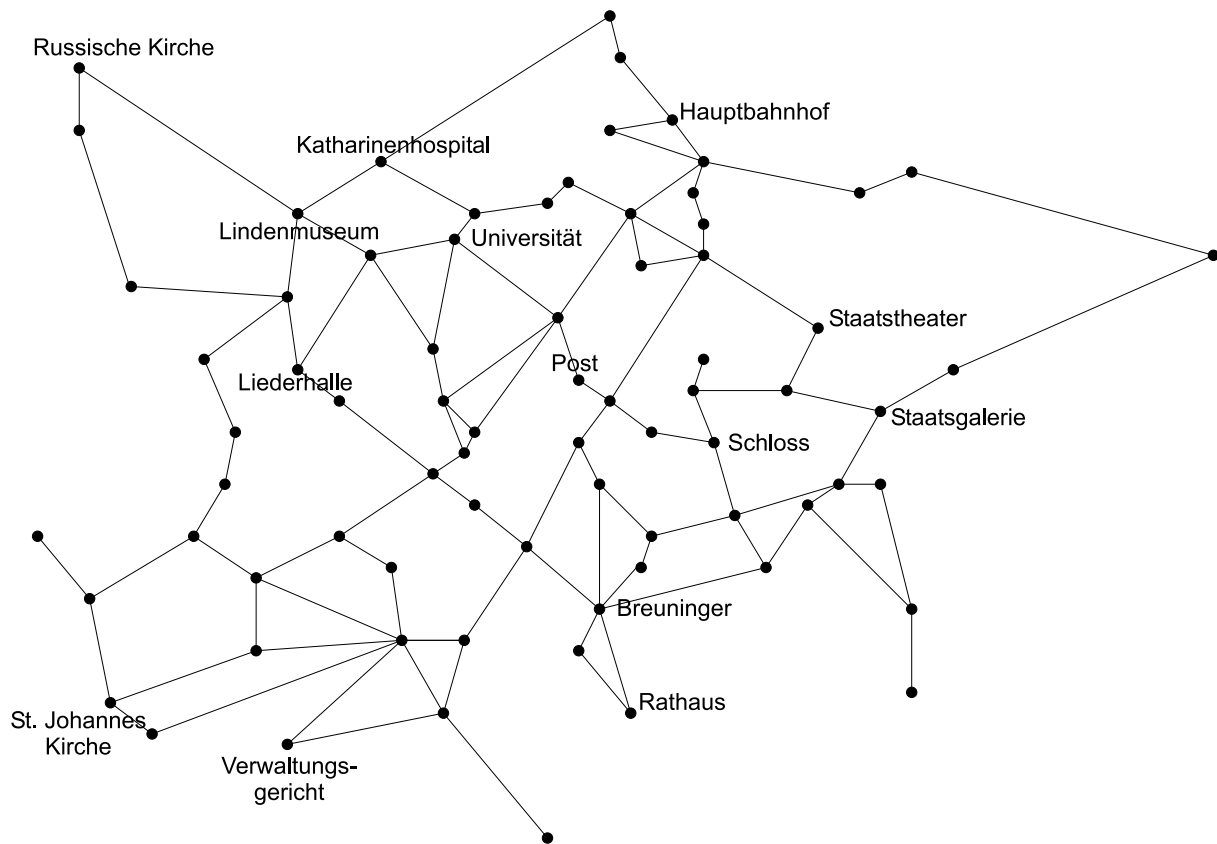
Ein weiterer wichtiger Aspekt von Informationszugriffsmodellen ist die Anzahl der Zugriffe, die pro Zeiteinheit zu erwarten sind. In [Ye et al. 1998] wird von einer Gleichverteilung der Zugriffszahlen pro Zeiteinheit ausgegangen. Die meisten anderen Autoren [Sistla et al. 1998, Rodriguez & Biersack 1998] nehmen hingegen eine Poisson-Verteilung an. Auf die mit dem Vorabübertragungsverfahren erzielten Trefferraten hat die Anzahl der Zugriffe pro Zeiteinheit jedoch keinen Einfluss. Daher wird dieser Aspekt nachfolgend nicht mehr weiter betrachtet.

## 6.3 Simulationsmodell

### 6.3.1 Überblick

Das in dieser Arbeit verwendete Simulationsmodell basiert auf einem Graph  $G(V, E)$ , der die Infrastruktur eines Ausschnitts aus der realen Welt repräsentiert. Daher wird er als Infrastrukturgraph bezeichnet. Die Knoten des Graphen repräsentieren Orte, die die simulierten Benutzer möglicherweise besuchen. Die Kanten modellieren die Verbindungen, wie Straßen, Fußwege, oder Schienen, zwischen diesen Orten. Es wird vorausgesetzt, dass der Infrastrukturgraph zusammenhängend ist, das heißt dass es für jedes beliebige Knotenpaar  $(i, j)$  einen Pfad von  $i$  nach  $j$  gibt. Für die Evaluierung in Kapitel 7 wurde ein Graph mit 115 Knoten und 150 Kanten verwendet, der die Innenstadt von Stuttgart modelliert. In der Darstellung dieses Graphen in Abbildung 6.2 sind zur besseren Übersichtlichkeit einige Knoten und Kanten nicht wiedergegeben.

Für die Modellierung der Benutzerbewegungen wird ein verhaltensorientiertes Bewegungsbedarfsmodell eingesetzt. Entsprechend wird davon ausgegangen, dass die Reise eines Benutzers stets dem Ziel dient, einen oder mehrere Orte, die im Infrastrukturgraphen repräsentiert sind, zu besuchen. Als Voreinstellung wird ferner davon ausgegangen, dass die Benutzer, dem Prinzip des geringsten Aufwands folgend, die als Ziel ausgewählten Orte auf dem kürzesten möglichen Pfad besuchen. Die entwickelte Simulationsumgebung erlaubt es aber auch, diese Voreinstellung zu ändern (siehe Abschnitt 6.4).



**Abbildung 6.2:** Infrastrukturgraph zur Modellierung der Stuttgarter Innenstadt.

Um die Simulationsumgebung so flexibel wie möglich gestalten zu können, werden alle weiteren Teilaspekte des mobilen Informationszugriffs durch drei separate Submodelle beschrieben: das Besuchsmodell, das Informationsmodell und das Zugriffsmodell. Dadurch ist es möglich einzelne Submodelle jederzeit auszutauschen, ohne dass davon die anderen Aspekte betroffen sind. Dies ist insbesondere dann wichtig, wenn die Simulationsumgebung zur Evaluierung weiterer ortsbezogener Systeme eingesetzt werden soll.

Die Grundidee der für die Evaluierung dieser Arbeit eingesetzten Submodelle ist die Unterscheidung zwischen bevorzugten Orten, die von besonderem Interesse für die Benutzer sind, und nicht bevorzugten Orten. Soll beispielsweise der Informationszugriff von Touristen simuliert werden, könnten die Orte, an denen sich Sehenswürdigkeiten befinden, bevorzugte Orte sein. Die Menge der bevorzugten Orte kann frei gewählt werden und ist für alle Submodelle gleich.

### 6.3.2 Besuchsmodell

Das Besuchsmodell wird benutzt, um die Zielknoten auszuwählen, die ein Benutzer während einer Reise besuchen möchte. Ein einfaches Besuchsmodell könnte zum Beispiel aus den Knoten des Infrastrukturgraphen eine feste Anzahl zufällig auswählen, wobei jeder einzelne Knoten mit der gleichen Wahrscheinlichkeit als Ziel gewählt wird.

In der Evaluierung in Kapitel 7 wurde ein Besuchsmodell benutzt, das jedem Knoten des Infrastrukturgraphen eine bestimmte Auswahlwahrscheinlichkeit zuordnet. Dabei kann zwischen bevorzugten und nicht bevorzugten Orten unterschieden werden, weil in der Regel die Benutzer bevorzugte Orte mit einer höheren Wahrscheinlichkeit als Ziel einer Reise auswählen werden. Damit hat das verwendete Besuchsmodell zwei Parameter: die Auswahlwahrscheinlichkeit  $p_p$  eines bevorzugten Ortes und die eines nicht bevorzugten Ortes  $p_n$ .

### 6.3.3 Informationsmodell

Das Informationsmodell ist für die Modellierung der Beziehungen zwischen den Informationsobjekten und den Orten verantwortlich. Es legt fest, auf welche Informationsobjekte von welchen Orten aus zugegriffen werden kann. Wollte man einen ortsunabhängigen Informationszugriff simulieren, müsste das Informationsmodell festlegen, dass von jedem beliebigen Ort aus auf alle Objekte zugegriffen werden kann. In einem inhärent ortsabhängigen System wäre jedem Ort nur eine bestimmte Teilmenge aller Informationsobjekte zugeordnet, wobei alle diese Teilmengen paarweise disjunkt sein müssten.

Das verwendete Informationsmodell ordnet jedem Ort eine feste Anzahl von Informationsobjekten zu, die ausschließlich vom jeweiligen Ort aus referenziert werden. Dabei kann den bevorzugten Orten eine größere Anzahl an Objekten zugeordnet werden als den nicht bevorzugten, weil zu erwarten ist, dass mehr Informationen über die für die Benutzer interessanten Orte angeboten werden.

Zusätzlich zu diesen lokalen Informationsobjekten ist im verwendeten Informationsmodell ein globaler Informationsvorrat vorgesehen, dessen Objekte von allen Orten aus angefragt werden. Die Größe dieses globalen Informationsvorrats wird durch den Parameter  $a_g$  festgelegt. Weitere Parameter des verwendeten Informationsmodells

sind die Anzahl der Objekte  $a_p$ , die bevorzugten Orten zugeordnet sind, sowie die Anzahl  $a_n$  der den nicht bevorzugten Orten zugeordneten Objekte. Die in den Parametern  $a_p$  und  $a_n$  festgelegten Werte können an jedem Ort um einen zufälligen, gleichverteilten Prozentsatz variieren. Die maximale Abweichung wird durch den Parameter  $a_e$  kontrolliert. Durch geeignete Wahl der Parameter ist es mit diesem Informationsmodell möglich, das gesamte Spektrum von inhärent ortsbezogenen bis zu nicht ortsbezogenen Informationssystemen zu simulieren.

### 6.3.4 Zugriffsmodell

Das Zugriffsmodell legt fest, wieviele Informationszugriffe ein Benutzer beziehungsweise seine Anwendung an jedem Ort durchführt und wie sich diese Zugriffe auf die einzelnen an dem jeweiligen Ort gemäß dem Informationsmodell verfügbaren Informationsobjekte verteilen. Beispielsweise könnte für die Anzahl der Zugriffe eine Normalverteilung angenommen werden und für die Verteilung der Zugriffswahrscheinlichkeiten eine Gleichverteilung.

Im Informationsmodell, das in Kapitel 7 verwendet wird, bestimmt ein Parameter  $r_p$  die Anzahl der Informationsobjekte, die an bevorzugten Orten angefragt werden, und ein weiterer Parameter  $r_n$  die Anzahl der an nicht bevorzugten Orten referenzierten Objekte. Auch hier werden zwei verschiedene Parameter für bevorzugte und nicht bevorzugte Orte verwendet, weil davon auszugehen ist, dass an bevorzugten Orten mehr Informationen angefordert werden. Auch die Werte der Parameter  $r_p$  und  $r_n$  können an jedem Ort um einen zufälligen, gleichverteilten Prozentsatz variieren. Hier wird die maximale Abweichung durch den Parameter  $r_e$  kontrolliert.

Ein weiterer Parameter  $p_g$  bestimmt, mit welcher Wahrscheinlichkeit auf Objekte des globalen Informationsvorrats zugegriffen wird. Für die Verteilung der Zugriffe auf die Objekte des globalen Informationsvorrats sowie auf die den einzelnen Orten zugeordneten Objekte wurde eine Zipf-ähnliche Verteilung mit dem Parameter  $\alpha$  angenommen.

## 6.4 Simulationsalgorithmus

Jede Simulation besteht aus einer konfigurierbaren Anzahl  $n$  von unabhängigen Touren durch den Infrastrukturgraphen. Jede Tour einer Simulation wird mit den gleichen

Parametereinstellungen simuliert. Am Ende einer Simulation werden die Ergebnisse der Simulation, in denen alle simulierten Touren berücksichtigt sind, angezeigt. Auf Wunsch kann auch eine Protokolldatei erstellt werden, in der alle Informationszugriffe und Bewegungen der simulierten Benutzer protokolliert werden.

Der erste Schritt des Simulationsalgorithmus (siehe Algorithmus 6.1) ist die Berechnung einer Matrix mit den kürzesten Pfaden des Infrastrukturgraphen. Diese Matrix enthält für jedes Knotenpaar  $(i, j)$  den kürzesten Pfad von  $i$  nach  $j$  sowie dessen Länge. Die in der Matrix gespeicherte Information wird später für die Berechnung der kürzesten Routen benötigt.

```

Berechne Matrix mit kürzesten Pfaden;
for  $i = 1$  to  $n$  do
    zielknoten = besuchsmodell.bestimmteZielknoten();
    kürzesteRoute = bestimmeKürzesteRoute(zielknoten);
5:   for all  $v \in$  kürzesteRoute do
        referenzierteObjekte = zugriffsmodell.bestimmeReferenzierteObjekte(v);
        Registriere Anfragen und Ort in Protokolldatei;
        Aktualisiere Zähler;
    end for
10: end for

```

**Algorithmus 6.1:** Simulationsalgorithmus.

Anschließend werden nacheinander die  $n$  unabhängigen Touren simuliert. Die Simulation jeder einzelnen Tour beginnt mit der Bestimmung der Knoten, die der simulierte Benutzer als Zielknoten seiner Tour auswählt. Für diese Bestimmung wird das Besuchsmodell benutzt.

Danach muss die Route bestimmt werden, auf der die Zielknoten besucht werden. Neben den Zielknoten enthält diese Route im Allgemeinen noch eine Reihe weiterer Knoten, die jeweils auf dem Weg von einem Zielknoten zum nächsten liegen. Wie bereits erwähnt wurde, wird standardmäßig die kürzeste mögliche Route gewählt. Auf Wunsch kann die Reihenfolge, in der die Zielknoten besucht werden, auch zufällig gewählt werden. Des Weiteren wurde die „Nächst gelegener Nachbar“-Heuristik (Nearest Neighbor Heuristics) implementiert, bei der ausgehend von einem zufällig ausgewählten Zielknoten, immer der am nächst gelegene Knoten als nächstes besucht wird. In [Lawler et al. 1985] werden empirische Analysen erwähnt, die zeigen, dass die mit dieser Heuristik gefundenen Routen im durchschnittlichen Fall nur ca. 15% länger sind, als die kürzeste mögliche Route.

Mit Hilfe des Zugriffsmodells wird für jeden Knoten auf der Route bestimmt, welche Informationsobjekte der simulierte Benutzer dort anfordert. Das Informationsmodell wird dabei nicht direkt vom Simulationsalgorithmus benutzt, sondern nur indirekt über das Zugriffsmodell referenziert. Schließlich werden die Informationszugriffe und der jeweils besuchte Knoten gegebenenfalls in der Protokolldatei festgehalten, sowie die Zähler, in denen die Ergebnisse aus den einzelnen Touren zusammengefasst werden, aktualisiert. Unter anderem werden die Anzahl der Anfragen für jedes Informationsobjekt, sowie die Anzahl der Besuche jeden Ortes gezählt.

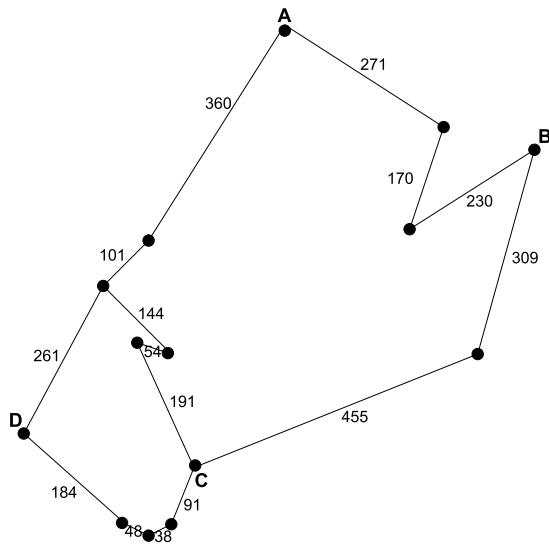
Damit ist die Simulation einer einzelnen Tour abgeschlossen und der nächste Durchgang beginnt, bis schließlich die Anzahl der zu simulierenden Touren erreicht ist.

Das bekanntermaßen NP-vollständige Problem des Handlungsreisenden [Lawler et al. 1985] ist ein Spezialfall der im Simulationsalgorithmus durchgeführten Berechnung einer kürzesten Route, bei dem alle Knoten des Infrastrukturgraphen als Zielknoten ausgewählt wurden. Daher ist auch die im Simulationsalgorithmus durchzuführende Berechnung im Allgemeinen NP-vollständig. Um die Problemgröße möglichst klein zu halten, wird der Infrastrukturgraph  $G$  vor der Berechnung der kürzesten Route auf einen kleineren Graphen  $G'(V', E')$  reduziert.

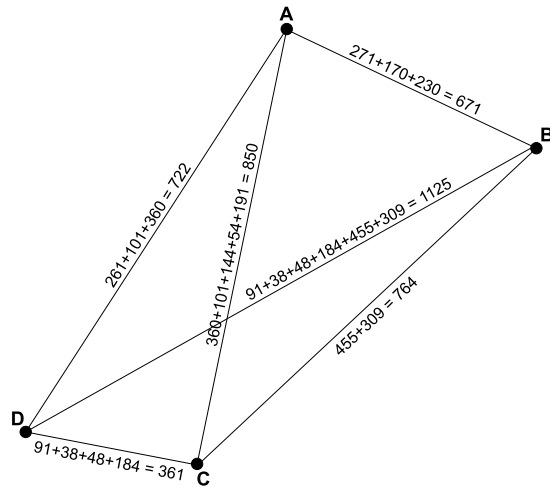
Die Knotenmenge  $V'$  des reduzierten Graphen beinhaltet ausschließlich die Zielknoten der jeweiligen Tour. Jedes Knotenpaar in  $V'$  wird durch eine Kante verbunden. Dabei wird der Kante zwischen zwei Knoten  $i$  und  $j$  die Länge des kürzesten Pfades zwischen diesen beiden Knoten im Originalgraph  $G$  zugewiesen. Abbildung 6.3 zeigt einen einfachen Infrastrukturgraphen (Abbildung 6.3(a)) und die Reduktion für die Zielknoten  $A, B, C, D$  (Abbildung 6.3(b)).

In diesem reduzierten Graphen ist nun das Problem des Handlungsreisenden zu lösen. Es hat sich gezeigt, dass auf aktueller Standard-PC-Hardware in akzeptabler Zeit eine exakte Lösung für bis zu zehn Zielknoten gefunden werden kann [Hegele 2000]. Sollen mehr als zehn Zielknoten besucht werden, empfiehlt es sich, die ebenfalls implementierte „Minimaler Spannbaum“-Heuristik (Minimum Spanning Tree Heuristics, [Lawler et al. 1985]) zu verwenden. Diese Heuristik hat nur noch eine Komplexität von  $\mathcal{O}(n^2)$  und findet garantiert eine Lösung, die maximal doppelt so lang ist, wie die kürzeste mögliche Route. Der Benutzer der Simulationsumgebung kann einen Grenzwert für die Anzahl der Zielknoten spezifizieren, bis zu dem die exakte Berechnung durchgeführt werden soll. Wenn mehr Knoten als Zielknoten ausgewählt wurden, wird automatisch die Heuristik verwendet.

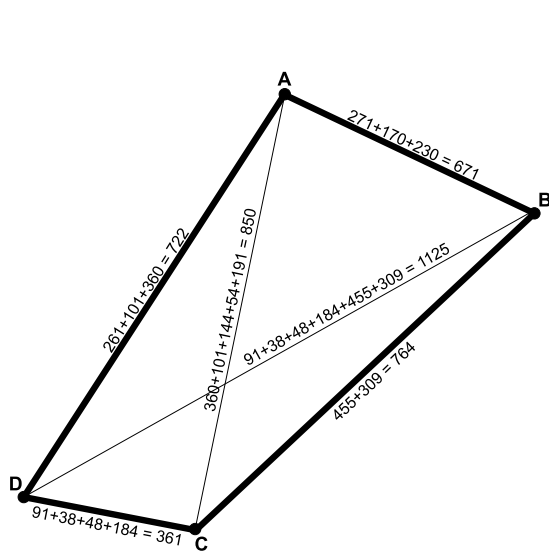




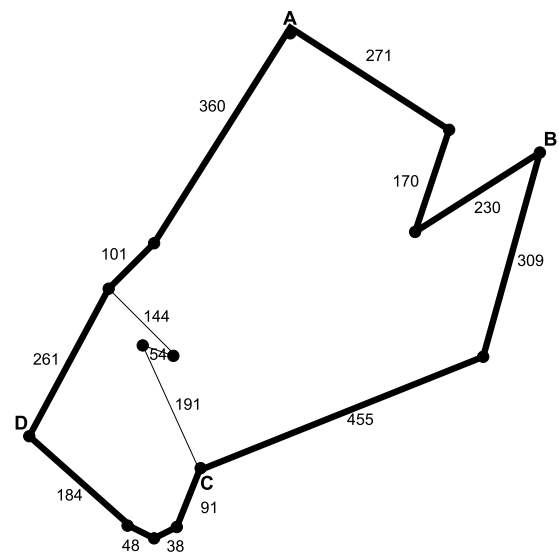
(a) Infrastrukturgraph.



(b) Reduzierter Graph.



(c) Lösung im reduzierten Graph (dicke Linien).



(d) Lösung im Originalgraph.

**Abbildung 6.3:** Berechnung der kürzesten Route für die Zielknoten A, B, C und D.

Unabhängig davon ob eine exakte Lösung des Problems berechnet wurde oder ob die Heuristik eingesetzt wurde, muss die gefundene Route (Abbildung 6.3(c)) in den Originalgraphen übertragen werden. Die Route im Originalgraphen (Abbildung 6.3(d)) wird folgendermaßen bestimmt: Die Zielknoten werden in der Reihenfolge besucht wie sie durch die Lösung des Problems des Handlungsreisenden im reduzierten Graphen gefunden wurde. Als Weg zwischen zwei Zielknoten wird immer der kürzeste mögliche Pfad zwischen diesen beiden Knoten, wie er in der Matrix der kürzesten Pfade gespeichert ist, benutzt.

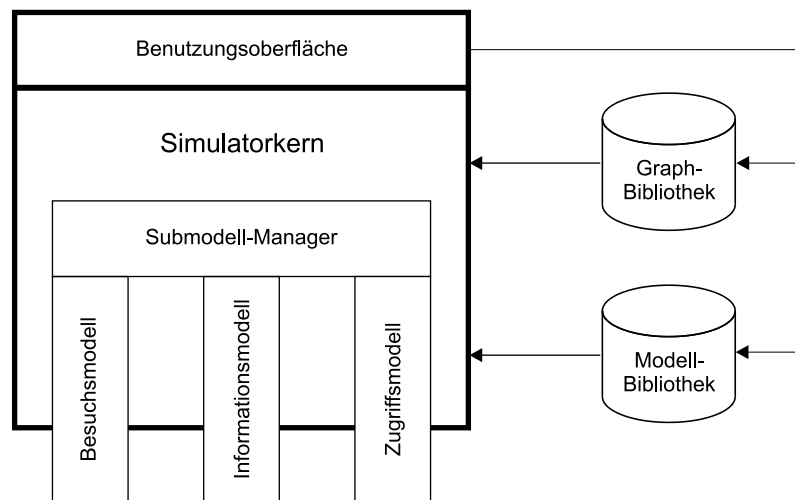
## 6.5 Implementierung

Die Simulationsumgebung wurde in Java 1.2 implementiert. Ein wichtiges Ziel beim Entwurf der Architektur war es, eine möglichst große Flexibilität zu gewährleisten, so dass die Simulationsumgebung auch über den Rahmen dieser Arbeit hinaus eingesetzt werden kann.

### 6.5.1 Architektur

Abbildung 6.4 gibt einen Überblick über die Architektur der Simulationsumgebung. Der eigentliche Simulationsalgorithmus wird durch den Simulatorkern implementiert. Die verwendeten Submodelle sind jeweils in eigene Klassen gekapselt. Dadurch können sie jederzeit durch andere Submodelle ersetzt werden, ohne dass hierzu Änderungen im Simulatorkern oder den anderen Submodellen notwendig wären.

Die Architektur sieht sogar vor, dass der Benutzer des Simulators aus einer Modellbibliothek diejenigen Submodelle über die graphische Oberfläche auswählt, die er für die Simulation benutzen möchte. Allerdings sind bisher neben den für diese Arbeit verwendeten Submodellen nur sehr einfache Submodelle implementiert. Analog kann der Benutzer den zu verwendenden Infrastrukturgraphen aus einer Bibliothek auswählen. Neben dem in Abschnitt 6.3 erwähnten Graphen, sind momentan noch einige eher theoretisch motivierte Graphen, z.B. Gitter, und einige kleinere Graphen, die zu Testzwecken eingesetzt wurden, verfügbar.



**Abbildung 6.4:** Die Architektur der Simulationsumgebung.

Der Modellbibliothek können jederzeit neue Submodelle hinzugefügt werden. Dabei muss lediglich darauf geachtet werden, dass ein neues Submodell, die vom Simulatorkern erwartete Schnittstelle implementiert. Besuchersmodelle müssen die Methode `Vertex[] getDestinations()` implementieren, die eine Menge von Zielknoten als Ergebnis liefert. Ferner wird eine Methode `float getVisitProbability(Vertex)` erwartet, die die Besuchswahrscheinlichkeit eines Knoten liefert.

Jedes Informationsmodell muss die Methoden `InformationObjects[] getAssociatedObjects(Vertex)` und `int getTotalNumberOfObjects()` bereitstellen. Die erste Methode liefert die Informationsobjekte, auf die an einem bestimmten Ort zugegriffen werden kann, während die zweite Methode benutzt wird, um die Anzahl der Informationsobjekte im gesamten Informationsraum zu erfragen.

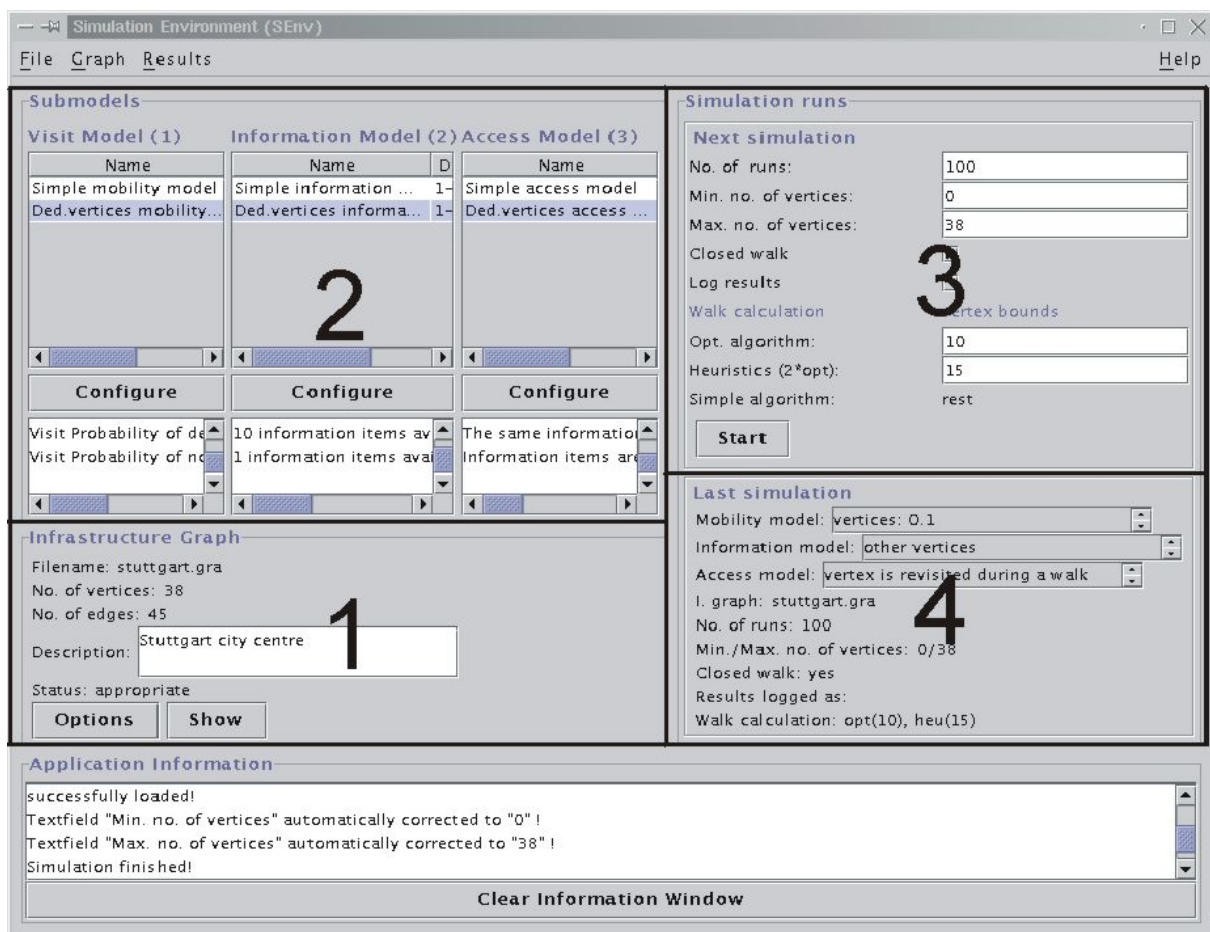
Die Zugriffsmodelle schließlich müssen die Methode `InformationObjects[] getAccessedObjects(Vertex)` anbieten, mit der die an einem bestimmten Ort tatsächlich referenzierten Informationsobjekte bestimmt werden können.

Die Aufgabe des Submodell-Managers ist es, gegebenenfalls neue Submodelle zu instantiieren und diese entsprechend den Einstellungen des Benutzers zu konfigurieren.

### 6.5.2 Benutzungsoberfläche

Nach dem Start der Simulationsumgebung wird das Hauptfenster, wie in Abbildung 6.5 wiedergegeben, angezeigt. Hier kann der Benutzer zunächst den zu verwendenden

Infrastrukturgraphen auswählen (1). Dabei können auch vom Benutzer selbst definierte Graphen, die in einem XML-basierten Format beschrieben werden [Hegele 2000], ausgewählt werden. Nach dem Laden eines Graphen wird zunächst überprüft, ob dieser zusammenhängend ist. Schließlich werden dem Benutzer Informationen über den geladenen Graphen, wie zum Beispiel die Zahl der Knoten, angezeigt. Der Benutzer kann auch eine graphische Anzeige des Graphen in einem getrennten Fenster abrufen.



**Abbildung 6.5:** Hauptfenster der Simulationsumgebung.

Als nächstes wählt der Benutzer die gewünschten Submodelle aus (2) und konfiguriert sie in einem weiteren Fenster. Bevor die Simulation gestartet wird, müssen noch einige allgemeine Einstellungen zum Simulationsablauf vorgenommen werden (3). Zum Beispiel muss die Anzahl der zu simulierenden Touren angegeben werden. Der Be-

reich (4) schließlich zeigt nach erfolgreichem Abschluss einer Simulation Informationen über deren Verlauf an.

Zur Evaluierung der Simulationsergebnisse kann der Benutzer eine Reihe von Tabellen abrufen, die beispielsweise die Anzahl der Zugriffe auf jedes Informationsobjekt anzeigen oder auch die Anzahl der Besuche jedes Knotens im Infrastrukturgraphen. Diese Tabellen können auf Wunsch in einem Format gespeichert werden, das unmittelbar mit dem Werkzeug gnuplot<sup>1</sup> weiterverarbeitet werden kann.

Zusätzlich können die Besuchshäufigkeiten in Karten, vergleichbar den Besuchswahrscheinlichkeitskarten aus Abschnitt 4.4.4, graphisch dargestellt werden. Ferner kann auf Wunsch, wie bereits erwähnt wurde, der gesamte Simulationsverlauf in einer Datei mitprotokolliert werden. Neben der Beschreibung aller Simulationsparameter enthält diese Datei für jeden Knotenbesuch einen Eintrag, der den Bezeichner des Knotens und die dort referenzierten Informationsobjekte angibt.

---

<sup>1</sup><http://www.gnuplot.org/>

## Kapitel 7

# Evaluierung

Zunächst werden in diesem Kapitel die Eigenschaften eines Informationssystems identifiziert, die die Effizienz des vorgestellten Vorabübertragungsverfahrens beeinflussen und Auswirkungen auf den Vorteil des erweiterten Verfahrens gegenüber dem Basisverfahren haben.

Nach einer kurzen Diskussion des erwarteten Verhaltens des Vorabübertragungsverfahrens in Abhängigkeit von diesen Eigenschaften, schließt sich eine mathematische Analyse der zu erwartenden Trefferraten an.

Am Ende des Kapitels werden die umfangreichen Simulationen, die mit der im vorherigen Kapitel vorgestellten Simulationsumgebung und der in Kapitel 5 beschriebenen Implementierung des Vorabübertragungsverfahrens durchgeführt wurden, vorgestellt. Die Wahl der Parameter des Simulationsmodells wurde auf Grundlage einer Analyse eines existierenden ortsbezogenen Informationssystems getroffen. Eine Beschreibung dieser Analyse ist ebenfalls Bestandteil dieses Kapitels.

### 7.1 Eigenschaften eines Informationssystems

Die Effektivität der vorgestellten Verfahren und der Vorteil des erweiterten Verfahrens gegenüber dem Basisverfahren hängen von einer Reihe von Eigenschaften des unterstützten Informationssystems ab. In diesem Abschnitt werden diese Parameter des Informationssystems beschrieben und mit Hilfe der in Abschnitt 3.1 eingeführten Konzentrationsmaße werden Maße definiert, um sie zu quantifizieren.

### 7.1.1 Definitionen

Ein wichtiger Faktor für die Effektivität der Vorabübertragung von Informationen ist die Verteilung der Informationszugriffe auf die einzelnen Informationsobjekte eines Informationsraums. Wenn sich ein großer Anteil der Anfragen auf wenige Informationsobjekte konzentriert, dann können mit wenigen vorab übertragenen Objekten viele Anfragen beantwortet werden. Dabei ist auch die absolute Anzahl der Informationsobjekte, auf die sich die Anfragen konzentrieren, von Bedeutung. Folglich wird das Maß für diese Konzentration, das nachfolgend als Fokussierung bezeichnet wird, mit Hilfe des Herfindahl-Koeffizienten definiert.

#### DEFINITION 7.1 (FOKUSSIERUNG)

Sei  $IR$  ein Informationsraum und  $h : IR \rightarrow \mathbb{N}$  die Häufigkeitsverteilung, die jedem Informationsobjekt in  $IR$  die absolute Häufigkeit zuordnet, mit der es innerhalb eines wohldefinierten Zeitraums referenziert wurde. Dann ist die Fokussierung des Informationsraums  $IR$  der Herfindahl-Koeffizient der Verteilung  $h$ .

Die Konzentration der Zugriffe, die innerhalb einer bestimmten Zone auftreten, auf die einzelnen Informationsobjekte spielt für die Effektivität des erweiterten Verfahrens eine entscheidende Rolle. Als Maß für diese Konzentration der Informationszugriffe innerhalb der einzelnen Zonen dient die mittlere lokale Fokussierung eines Informationssystems. Sie leitet sich aus den lokalen Fokussierungen innerhalb der betrachteten Zonen ab.

#### DEFINITION 7.2 (LOKALE FOKUSSIERUNG)

Sei  $z$  eine Zone und  $h : IR \rightarrow \mathbb{N}$  die Häufigkeitsverteilung, die jedem Informationsobjekt in  $IR$  die absolute Häufigkeit zuordnet, mit der es innerhalb eines wohldefinierten Zeitraums in der Zone  $z$  referenziert wurde. Dann ist die lokale Fokussierung innerhalb der Zone  $z$  der Herfindahl-Koeffizient der Verteilung  $h$ .

#### DEFINITION 7.3 (MITTLERE LOKALE FOKUSSIERUNG)

Sei das Abdeckungsgebiet eines Informationssystems in eine Menge von Zonen  $Z$  unterteilt und sei  $h : Z \rightarrow \mathbb{N}$  eine Funktion, die jeder Zone die Anzahl der dort innerhalb eines wohldefinierten Zeitraums aufgetretenen

Informationsanfragen zuordnet. Sei ferner  $f : Z \rightarrow [0, 1]$  eine Funktion die jeder Zone  $z$  die lokale Fokussierung in  $z$  zuordnet. Dann ist die mittlere lokale Fokussierung  $\bar{F}$  eines Informationssystems definiert durch das gewichtete arithmetische Mittel der lokalen Fokussierungen aller betrachteten Zonen:

$$\bar{F} = \sum_{z \in Z} \frac{h(z)}{\sum_{i \in Z} h(i)} \cdot f(z)$$

Die mittlere lokale Fokussierung hängt natürlich von der gewählten Zoneneinteilung ab. Bei Vergleichen muss daher immer die gleiche Zoneneinteilung zu Grunde gelegt werden.

Da das vorgeschlagene Verfahren die Ortsbezogenheit der Informationen ausnutzt, hat die Abhängigkeit der Zugriffswahrscheinlichkeit auf ein Informationsobjekt vom Aufenthaltsort der Benutzer ebenfalls eine starke Auswirkung auf die Wirksamkeit des Verfahrens. Um auch dafür ein Maß zu haben, wird die Ortsbezogenheit eines Informationsobjekts wie folgt definiert:

**DEFINITION 7.4 (ORTSBEZOGENHEIT)**

Sei das Abdeckungsgebiet eines Informationssystems in eine Menge von Zonen  $Z$  unterteilt. Sei  $IO$  ein Informationsobjekt und  $h : Z \rightarrow \mathbb{N}$  eine Häufigkeitsverteilung, die jeder Zone  $z$  die Häufigkeit zuordnet, mit der das Informationsobjekt  $IO$  in der Zone  $z$  innerhalb eines wohldefinierten Zeitraums referenziert wird. Dann wird der Gini-Koeffizient der Verteilung  $h$  als Ortsbezogenheit des Informationsobjekts  $IO$  bezeichnet.

Die Ortsbezogenheit eines Informationsobjekts hängt ebenfalls von der gewählten Zoneneinteilung ab. Um die Ortsbezogenheit eines Informationsobjekts auch als Maß dafür verwenden zu können, wieviel Nutzen aus eventuell verfügbarem Wissen über die zukünftigen Aufenthaltsorte eines Benutzers gezogen werden kann, muss der Gini-Koeffizient zur Definition benutzt werden. Dass der Herfindahl-Koeffizient ungeeignet ist, kann an den Verteilungen aus Tabelle 3.2 verdeutlicht werden, wenn man annimmt, dass sie nicht die Häufigkeiten angeben, mit denen auf verschiedene Informationsobjekte zugegriffen wird, sondern die Häufigkeiten mit denen in verschiedenen Zonen auf ein bestimmtes Informationsobjekt zugegriffen wird. Bei der ersten Verteilung kann dann kein Nutzen aus der Information gezogen werden, welche der



drei Zonen ein Benutzer besuchen wird, weil in allen drei Zonen gleich häufig auf das Informationsobjekt zugegriffen wird. In der zweiten Verteilung hingegen kann man ableiten, dass bei einem Besuch der Zonen eins bis vier wahrscheinlicher auf das Informationsobjekt zugegriffen werden wird als bei einem Besuch der Zonen sechs bis zehn. Trotzdem ist der Herfindahl-Koeffizient in der ersten Verteilung höher, weil er auch die absolute Konzentration, die für die Hoarding-Entscheidung keine Vorteile bringt, berücksichtigt.

Die Ortsbezogenheit eines einzelnen Informationsobjekts hat nur einen geringen Einfluss auf die Gesamtleistung des Vorabübertragungsverfahrens. Um eine Aussage darüber zu machen, müssen alle Informationsobjekte eines Informationsraums betrachtet werden. Dies führt zur Definition der mittleren Ortsbezogenheit eines Informationsraums:

**DEFINITION 7.5 (MITTLERE ORTSBEZOGENHEIT)**

Sei  $IR$  ein Informationsraum und  $h : IR \rightarrow \mathbb{N}$  die Häufigkeitsverteilung, die jedem Informationsobjekt in  $IR$  die absolute Häufigkeit zuordnet, mit der es innerhalb eines wohldefinierten Zeitraums referenziert wurde. Sei ferner  $o : IR \rightarrow [0, 1]$  eine Funktion, die jedem Informationsobjekt seine Ortsbezogenheit zuordnet. Dann ist die mittlere Ortsbezogenheit  $\bar{O}$  des Informationsraums definiert als gewichtetes arithmetisches Mittel der Ortsbezogenheiten aller Informationsobjekte in diesem Informationsraum:

$$\bar{O} = \sum_{i \in IR} \frac{h(i)}{\sum_{j \in IR} h(j)} \cdot o(i)$$

Damit mit dem erweiterten Verfahren höhere Trefferraten als mit dem Basisverfahren erzielt werden, sollten die Benutzer jeweils nur einen kleinen Teil aller Zonen innerhalb eines Hoarding-Gebiets besuchen. Werden nur wenige Zonen besucht, dann können für jede der besuchten Zonen mehr Informationsobjekte vorab übertragen werden, als wenn sehr viele Zonen besucht werden. Die mit dem erweiterten Verfahren innerhalb eines Hoarding-Gebiets erzielte Trefferrate hängt also auch von der Besuchsrates der Benutzer bezüglich den Zonen ab:

**DEFINITION 7.6 (BESUCHSRATE)**

Sei  $Z$  die Menge aller Zonen eines Hoarding-Gebiets und sei  $B$  die Menge der innerhalb eines wohldefinierten Zeitraums beobachteten Besuche des

Hoarding-Gebiets. Sei weiter  $n : B \rightarrow \mathbb{N}$  eine Funktion, die jedem Besuch des Hoarding-Gebiets die Zahl der dabei besuchten Zonen zuordnet. Dann ist die Besuchsrates  $BR$  definiert als der mittlere Anteil der besuchten Zonen:

$$BR = \frac{1}{|B|} \cdot \sum_{b \in B} \frac{n(b)}{|Z|}$$

### 7.1.2 Erwartetes Verhalten

Wenn die Fokussierung eines Informationsraums hoch ist, dann werden bereits mit dem Basisverfahren hohe Trefferraten erzielt werden, weil es auch die beliebtesten Informationsobjekte identifizieren kann und ein großer Anteil der Anfragen sich auf diese Objekte beziehen wird. In diesem Fall ist der potenzielle Vorteil des erweiterten Verfahrens geringer als bei einem schlechten Abschneiden des Basisverfahrens.

Um darüber zu entscheiden, ob der Einsatz des erweiterten Verfahrens lohnt, reicht es nicht, nur die mittlere Ortsbezogenheit eines Informationsraums zu betrachten. Eine hohe Ortsbezogenheit ist zwar eine notwendige Bedingung für den lohnenswerten Einsatz des erweiterten Verfahrens, wie man aus dem Beispiel in Abbildung 7.1(a) sehen kann, ist sie jedoch keine hinreichende Bedingung. Die Matrix zeigt, wie häufig in jeder Zone auf die einzelnen Informationsobjekte zugegriffen wird. Alle Zugriffe innerhalb einer bestimmten Zone sind in einer Zeile angegeben. Alle Zugriffe auf ein bestimmtes Informationsobjekt sind in einer Spalte angegeben. Damit gibt der Wert in Zeile  $i$  und Spalte  $j$  an, wie häufig in Zone  $i$  auf das Informationsobjekt  $j$  zugegriffen wird. Die mittlere Ortsbezogenheit des Informationsraums ist sehr hoch, weil auf alle Informationsobjekte ausschließlich von einer Zone aus zugegriffen wird. Da es aber immer die gleiche Zone ist, kann kein Nutzen aus dem Wissen gezogen werden, welche Zonen ein Benutzer besucht. Daher bietet das erweiterte Verfahren in diesem Fall keinen Vorteil gegenüber dem Basisverfahren.

Für einen vorteilhaften Einsatz des erweiterten Verfahrens muss zusätzlich noch die mittlere lokale Fokussierung hoch sein. Eine hohe mittlere lokale Fokussierung bedeutet, dass es in jeder Zone klar favorisierte Informationsobjekte gibt. Aus der hohen Ortsbezogenheit folgt dann, dass die Favoriten in den einzelnen Zonen voneinander verschieden sind. Denn sobald ein Informationsobjekt in vielen Zonen gleich beliebt oder unbeliebt ist, ist seine Ortsbezogenheit gering. Besuchen die Benutzer dann auch noch nur einen kleinen Teil der Zonen eines Hoarding-Gebiets, so werden die mit dem

	$i_1$	$i_2$	$i_3$	$i_4$	$i_5$
$z_1$	0	0	0	0	0
$z_2$	0	0	0	0	0
$z_3$	10	10	10	10	10
$z_4$	0	0	0	0	0
$z_5$	0	0	0	0	0

(a) ortsbezogen.

	$i_1$	$i_2$	$i_3$	$i_4$	$i_5$
$z_1$	0	0	10	0	0
$z_2$	0	0	10	0	0
$z_3$	0	0	10	0	0
$z_4$	0	0	10	0	0
$z_5$	0	0	10	0	0

(b) fokussiert.

**Abbildung 7.1:** Ortsbezogene und lokal fokussierte Zugriffsmatrix.

erweiterten Verfahren erzielten Trefferraten hoch sein. Das Beispiel aus Abbildung 7.1(b) zeigt, dass eine hohe mittlere lokale Fokussierung alleine ebenfalls nicht ausreicht, um automatisch einen hohen Nutzen aus dem erweiterten Verfahren ziehen zu können. Dort ist zwar die mittlere lokale Fokussierung sehr hoch, weil aber in allen Zonen das gleiche Informationsobjekt bevorzugt wird, kann kein Nutzen aus dem Wissen gezogen werden, welche Zonen ein Benutzer besuchen wird.

### 7.1.3 Mathematische Analyse

In den nachfolgenden Analysen wird davon ausgegangen, dass die tatsächlich vorliegenden Zugriffs- und Besuchswahrscheinlichkeiten durch die Tabellen und Karten der Infostationen exakt nachgebildet werden. Bei den Analysen des erweiterten Verfahrens wird grundsätzlich davon ausgegangen, dass die jeweils besuchten Zonen dem Vorabübertragungsverfahren von externen Informationsquellen exakt mitgeteilt werden.

#### Fokussierung

Für die Analyse der Auswirkung unterschiedlicher Fokussierungen wird vereinfachend angenommen, dass das gesamte Abdeckungsgebiet des Informationssystems aus einem einzigen Hoarding-Gebiet besteht.

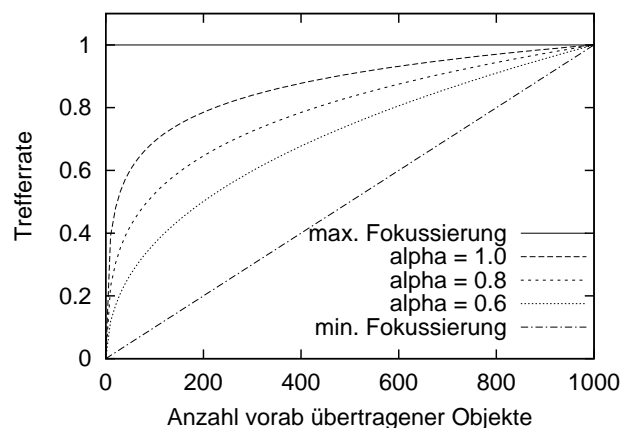
Sei  $p(i)$  die Zugriffswahrscheinlichkeit auf das  $i$ -beliebteste Informationsobjekt innerhalb eines Hoarding-Gebiets. Dann kann die mit dem Basisverfahren zu erwartende

Trefferrate  $ET_B$ , wie bereits in Abschnitt 4.5.2 erläutert wurde, folgendermaßen berechnet werden:

$$ET_B(S) = \sum_{i=1}^S p(i),$$

wobei  $S$  die Anzahl der vorab übertragbaren Informationsobjekte ist.

Der Maximalwert eins der Fokussierung wird erreicht, wenn sich alle Informationszugriffe auf ein bestimmtes Objekt beziehen. Dann ergibt sich eine zu erwartende Trefferrate von 1, sobald mindestens ein Objekt vorab übertragen wird. Die minimale Fokussierung  $\frac{1}{|IR|}$  in einem Informationsraum  $IR$  ergibt sich, wenn auf alle Objekte mit der gleichen Wahrscheinlichkeit zugegriffen wird. Das bedeutet, dass für die Zugriffswahrscheinlichkeit auf jedes Objekt  $i$  gilt:  $p(i) = \frac{1}{|IR|}$ . Daraus ergibt sich für die zu erwartende Trefferrate:  $ET_B(S) = \frac{1}{|IR|} \cdot S$ . In Abbildung 7.2 sind die zu erwartenden Trefferraten für die maximale und minimale Fokussierung in Abhängigkeit von der Anzahl der vorab übertragenen Objekte wiedergegeben. Für die Größe des Informationsraums wurde ein Wert von 1000 Informationsobjekten angenommen.



**Abbildung 7.2:** Zu erwartende Trefferraten für unterschiedlich fokussierte Informationsräume.

Die entsprechenden Kurven für beliebige weitere Zugriffswahrscheinlichkeitsverteilungen und damit auch für beliebige Fokussierungen liegen alle innerhalb des von den beiden Geraden für die minimale und maximale Fokussierung und der Ordinate-Achse aufgespannten Dreiecks. Die weiteren Kurven in Abbildung 7.2 zeigen die

zu erwartenden Trefferraten

$$ET_B(S) = \sum_{i=1}^S p(i) = \sum_{i=1}^S \frac{1}{\sum_{i=1}^{|IR|} \frac{1}{i^\alpha}} \cdot \frac{1}{i^\alpha} = \frac{\sum_{i=1}^S \frac{1}{i^\alpha}}{\sum_{i=1}^{|IR|} \frac{1}{i^\alpha}}$$

für Zipf-ähnliche Verteilungen mit Werten  $\alpha = 1.0$ ,  $\alpha = 0.8$  und  $\alpha = 0.6$  für den Parameter  $\alpha$ . Für die Fokussierungen  $F$  ergeben sich, gemäß

$$F = \sum_{i=1}^{|IR|} p(i)^2 = \frac{\sum_{i=1}^{|IR|} \frac{1}{i^{2\alpha}}}{\left(\sum_{i=1}^{|IR|} \frac{1}{i^\alpha}\right)^2},$$

Werte von 0.02933906569, 0.009440939350 beziehungsweise 0.003054213805.

Sofern im erweiterten Verfahren kein externes Wissen zur Verfügung steht, sind die hier gemachten Aussagen auch für das erweiterte Verfahren gültig. Auf den Nutzen, der aus externem Wissen gezogen werden kann, hat die globale Fokussierung keine Auswirkung. Da der Effekt der lokalen Fokussierung auf diesen Nutzen eng mit der Ortsbezogenheit zusammenhängt, wird die lokale Fokussierung gemeinsam mit der Ortsbezogenheit im nächsten Abschnitt diskutiert.

### Ortsbezogenheit und lokale Fokussierung

Die Auswirkungen, die die Ortsbezogenheit und die lokale Fokussierung der Informationsobjekte auf die Trefferraten haben, lassen sich am einfachsten zeigen, wenn man eine Gleichverteilung der Zugriffswahrscheinlichkeiten auf die Objekte im Informationsraum annimmt.

Sei  $|HG|$  die Gesamtzahl der Hoarding-Gebiete im Abdeckungsbereich des Informationssystems und sei weiter die Ortsbezogenheit bezüglich diesen Hoarding-Gebieten maximal, dann wird auf jedes Informationsobjekt von genau einem Hoarding-Gebiet aus zugegriffen.

Diese maximale Ortsbezogenheit kann auch bei einer niedrigen lokalen Fokussierung auftreten, beispielsweise wenn auf alle Objekte vom gleichen Hoarding-Gebiet aus zugegriffen wird. Dann erhält man für die Zugriffswahrscheinlichkeit innerhalb dieses Hoarding-Gebiets für alle Objekte  $p(i) = \frac{1}{|IR|}$  und für die zu erwartende Trefferrate in

diesem Gebiet:

$$ET_B(S) = \sum_{i=1}^S \frac{1}{|IR|} = \frac{1}{|IR|} \cdot S$$

Da in den übrigen Gebieten, dann auf gar keine Informationen zugegriffen wird, macht die Berechnung einer Trefferrate dort keinen Sinn.

Eine vergleichsweise hohe lokale Fokussierung bei gleichzeitig maximaler Ortsbezogenheit ergibt sich zum Beispiel, wenn die Mengen der innerhalb der einzelnen Hoarding-Gebiete referenzierten Objekte paarweise disjunkt sind und in jedem Hoarding-Gebiet auf die gleiche Anzahl verschiedener Objekte zugegriffen wird. Dann werden in jedem Hoarding-Gebiet  $\frac{|IR|}{|HG|}$  verschiedene Informationsobjekte referenziert. Da die Zugriffswahrscheinlichkeiten auf diese Objekte gleich verteilt sind, ergibt sich für jedes innerhalb eines bestimmten Hoarding-Gebiets referenzierte Objekt eine Zugriffswahrscheinlichkeit  $p(i) = \frac{|HG|}{|IR|}$ . Wenn wiederum  $S$  Informationsobjekte an jeder Infostation vorab übertragen werden können, erhält man mit dem Basisverfahren für die zu erwartende Trefferrate:

$$ET_B(S) = \sum_{i=1}^S \frac{|HG|}{|IR|} = \frac{|HG|}{|IR|} \cdot S$$

Damit ist die zu erwartende Trefferrate in diesem Fall linear von der Anzahl der Hoarding-Gebiete abhängig.

Bei einer minimalen Ortsbezogenheit wird hingegen auf jedes Informationsobjekt im Informationsraum von allen Hoarding-Gebieten aus zugegriffen. Das bedeutet, dass in jedem Hoarding-Gebiet auf alle  $|IR|$  Objekte zugegriffen wird. Damit ergibt sich für die Zugriffswahrscheinlichkeit innerhalb eines Hoarding-Gebiets für alle Objekte  $p(i) = \frac{1}{|IR|}$  und zwangsläufig eine minimale lokale Fokussierung. Für die zu erwartende Trefferrate erhält man:

$$ET_B(S) = \sum_{i=1}^S \frac{1}{|IR|} = \frac{1}{|IR|} \cdot S$$

Die Trefferrate hängt nicht mehr von der Anzahl der Hoarding-Gebiete ab. Die Unterteilung des Abdeckungsgebiets in einzelne Hoarding-Gebiete bringt hier also keinen Vorteil.

Wie bereits erwähnt wurde, kann die Ortsbezogenheit auch benutzt werden, um Aussagen über den potenziellen Vorteil des erweiterten Verfahrens zu treffen. Hierzu muss

die Ortsbezogenheit bezüglich der Zonen eines Hoarding-Gebiets, nicht mehr bezüglich der Hoarding-Gebiete selbst, betrachtet werden. Ferner muss hierzu die Gesamtzahl  $N$  der innerhalb eines Hoarding-Gebiets referenzierten Objekte anstatt der Anzahl  $|IR|$  der Objekte im Informationsraum, sowie die Anzahl  $|Z|$  der Zonen, in die ein Hoarding-Gebiet unterteilt ist, anstatt der Anzahl  $|HG|$  der Hoarding-Gebiete betrachtet werden. Dann können die zuvor getroffenen Aussagen über das Abdeckungsgebiet eines Informationssystems und seine Hoarding-Gebiete analog für ein Hoarding-Gebiet und seine Zonen gemacht werden.

### Besuchsrate

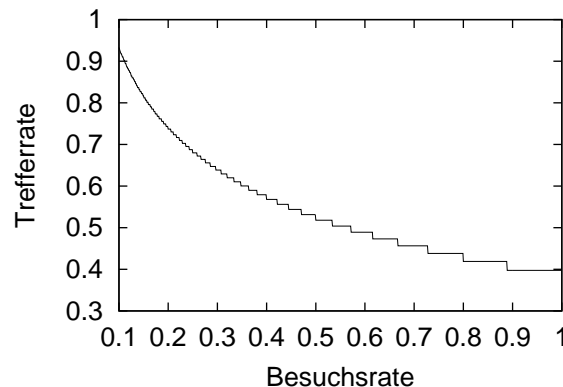
Sei ein Hoarding-Gebiet in  $|Z|$  Zonen unterteilt und sei die Besuchsrate bezüglich dieser Zonen  $BR$ . Dann werden im Mittel  $|B| = |Z| \cdot BR$  Zonen bei einem Besuch des Hoarding-Gebiets besucht.

Seien weiterhin die Mengen der in jeder Zone referenzierten Objekte paarweise disjunkt und sei  $n_o$  die Zahl der unterschiedlichen Objekte in jeder dieser Mengen. Ferner wird angenommen, dass die Zugriffswahrscheinlichkeiten auf die Objekte in jeder dieser Mengen Zipf-ähnlich verteilt sind.

Wenn insgesamt  $S$  Objekte vorab übertragen werden können und der Besuch der  $|B|$  Zonen exakt spezifiziert wird, indem eine externe Karte angegeben wird, in der die Besuchswahrscheinlichkeiten der besuchten Zonen auf eins gesetzt wird und die aller anderen auf null, dann werden für jede der besuchten Zonen im Mittel mindestens  $\lfloor \frac{S}{|B|} \rfloor = \lfloor \frac{S}{|Z| \cdot BR} \rfloor$  Objekte vorab übertragen. Für einige Zonen wird ein Informationsobjekt mehr übertragen, so dass auf jeden Fall die maximale Zahl übertragbarer Informationsobjekte ausgenutzt wird. Für die in jeder Zone mit dem erweiterten Verfahren zu erwartende Trefferrate und damit für die zu erwartende Gesamttrefferrate  $ET_E$  ergibt sich:

$$ET_E = \sum_{i=1}^{\lfloor \frac{S}{|Z| \cdot BR} \rfloor} p(i) = \frac{\sum_{i=1}^{\lfloor \frac{S}{|Z| \cdot BR} \rfloor} \frac{1}{i^\alpha}}{\sum_{i=1}^{n_o} \frac{1}{i^\alpha}}$$

Bei der Berechnung der zu erwartenden Trefferrate wurde vernachlässigt, dass für manche Zonen tatsächlich  $\lfloor \frac{S}{|Z| \cdot BR} \rfloor + 1$  Objekte übertragen werden. Daher liegt der tatsächliche Wert der zu erwartenden Trefferrate geringfügig über dem hier berechneten Wert. Abbildung 7.3 zeigt die zu erwartenden Trefferraten in Abhängigkeit von der Besuchsrate für die Parameter  $S = 80$ ,  $|Z| = 10$ ,  $n_o = 100$ ,  $\alpha = 0.8$ .



**Abbildung 7.3:** Trefferrate in Abhängigkeit von der Besuchsrare.

## 7.2 Analyse einer existierenden Anwendung

Für die Simulationen in Abschnitt 7.3 war es erforderlich, die Parameter des Simulationsmodells (siehe Abschnitt 6.3) festzulegen. Um eine möglichst genaue Vorstellung davon zu erhalten, wie diese Parameter in der Realität aussehen, wurde die nachfolgend beschriebene Analyse eines existierenden mobilen Touristenführers durchgeführt.

Darüber hinaus wird ein erster Eindruck von der Leistungsfähigkeit des Vorabübertragungsverfahrens in einer realen Umgebung vermittelt, indem die Trefferraten bestimmt werden, die mit dem Verfahren in Verbindung mit dieser Anwendung erreicht werden.

### 7.2.1 Beschreibung der Daten

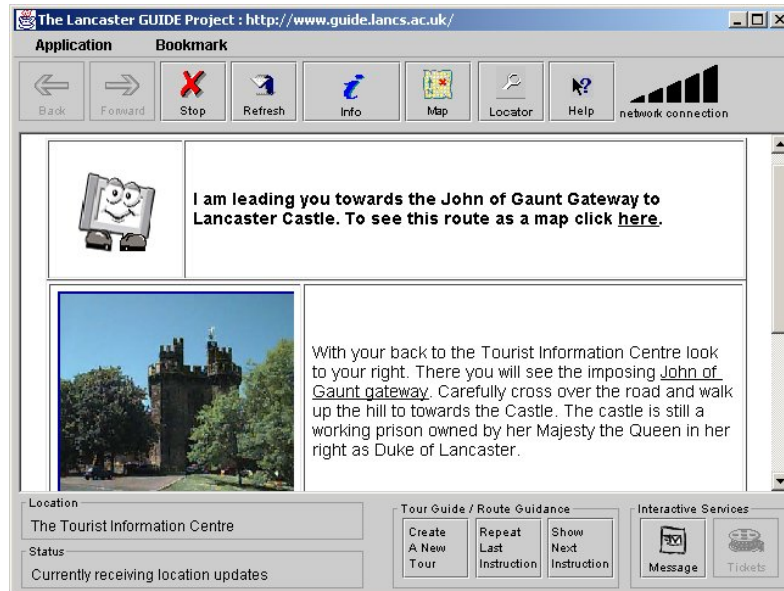
Im Rahmen des GUIDE-Projekts<sup>1</sup> [Cheverst, Davies, Mitchell & Friday 2000] an der Universität Lancaster wurde ein elektronischer Touristenführer, der so genannte „GUIDE“, entwickelt (siehe Abbildung 7.4). Der „GUIDE“ stellt unter anderem Wegbeschreibungen, Öffnungszeiten und Informationen zu historischen Gebäuden bereit. Er deckt nur einen Teil des historischen Stadtzentrums von Lancaster ab. Inzwischen wird der „GUIDE“ durch das Touristenbüro von Lancaster an Touristen verliehen. Als Hardware-Plattform wird ein Fujitsu TeamPad 7600 eingesetzt. Dabei handelt es sich

---

<sup>1</sup><http://www.guide.lancs.ac.uk/>



um einen Windows basierten Stift-PC, dessen Batteriekapazität für ca. zwei Stunden Dauerbetrieb ausreicht.



**Abbildung 7.4:** Die Benutzungsoberfläche der GUIDE-Anwendung.

Die in der nachfolgenden Analyse verwendeten Protokolldaten stammen aus Feldversuchen, die im Rahmen dieses Projekts durchgeführt wurden. Diese Feldversuche sind in der Literatur mehrfach beschrieben worden [Cheverst, Davies, Mitchell, Friday & Efstratiou 2000, Cheverst, Davies, Mitchell & Friday 2000, Davies et al. 1998]. Die Daten wurden freundlicherweise von den für die Implementierung des GUIDE-Systems verantwortlichen Personen für die Evaluierung dieser Arbeit zur Verfügung gestellt.

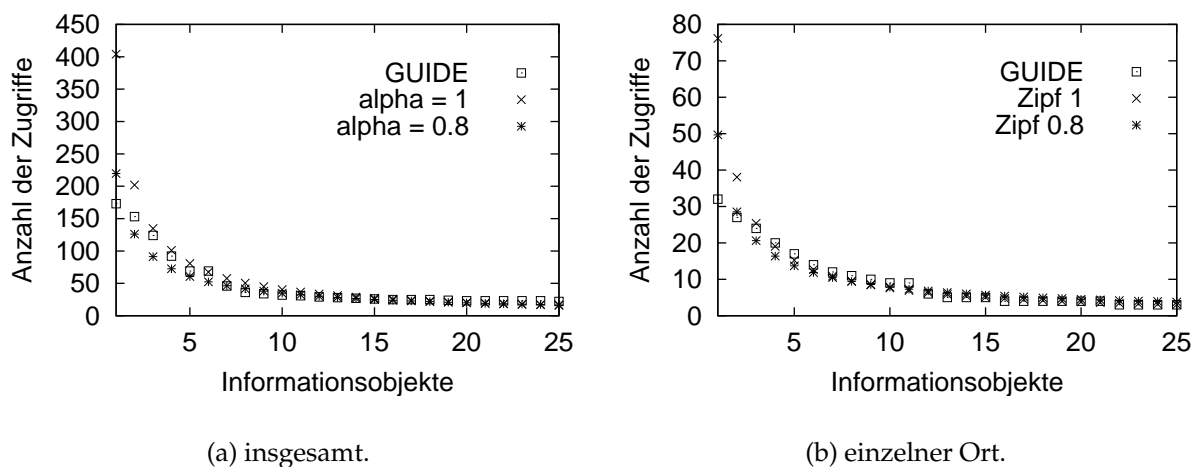
Der vollständige Datensatz umfasst 360 kbyte und besteht aus 24 unabhängigen Protokolldateien. In diesen Dateien sind jeweils alle Informationsanfragen protokolliert, die ein Benutzer während einer Besichtigungstour gestellt hat. Zusätzlich zu den Anfragen wurden auch die Orte, an denen die Anfragen gestellt wurden, festgehalten. Die 24 Dateien stammen von 20 verschiedenen Benutzern, das heißt von vier Benutzern wurden jeweils zwei Touren protokolliert. Insgesamt wurden 2639 Informationsanfragen protokolliert, die sich auf 385 verschiedene Informationsobjekte verteilt haben. Unter diesen Objekten befinden sich auch solche, die keine ortsbezogenen Informationen beinhalten, wie beispielsweise Objekte zur graphischen Benutzerführung, die auf dem Server gespeichert sind. Leider ist die Unterscheidung der verschiedenen Orte

sehr grob, so dass insgesamt nur fünf Orte (Tourist Information Centre (TIC), Shire, Castle, Priory und Judge) unterschieden werden.

## 7.2.2 Ergebnisse

### Verteilung der Informationszugriffe

Anhand der vorliegenden Protokolldateien konnte die Annahme bestätigt werden, dass die Zugriffswahrscheinlichkeiten auf einzelne Informationsobjekte Zipf-ähnlich verteilt sind. Als eine gute Wahl für den Parameter  $\alpha$  stellte sich der bereits in [Breslau et al. 1999] vorgeschlagene Wert von 0.8 heraus. Wie Abbildung 7.5 zeigt, stimmt die für das GUIDE-System beobachtete Verteilung der Zugriffe sowohl insgesamt als auch an einem einzelnen Ort sehr gut mit der theoretischen Verteilung überein, die sich ergibt, wenn Zipf-ähnlich verteilte Zugriffswahrscheinlichkeiten zu Grunde gelegt werden.



**Abbildung 7.5:** Verteilung der Zugriffe auf einzelne Informationsobjekte.

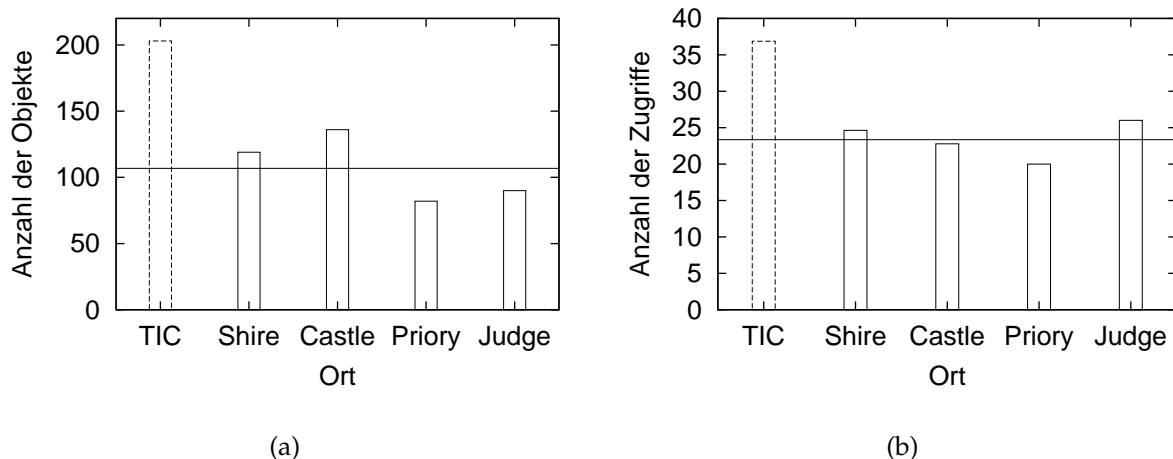
### Anzahl der referenzierten Informationsobjekte

Ein weiterer wichtiger Parameter des Simulationsmodells ist die Anzahl der an einem Ort referenzierten unterschiedlichen Informationsobjekte. Um eine Abschätzung für

diesen Parameter zu erhalten, wurde in den vorliegenden Daten untersucht, auf wieviele verschiedene Informationsobjekte an jedem Ort zugegriffen wird. Darin sind auch die bereits erwähnten globalen Informationsobjekte enthalten.

In den vorliegenden Protokolldaten tauchen nur Informationsobjekte auf, auf die auch mindestens einmal zugegriffen wurde. Sollte dennoch ein Informationsobjekt angeboten werden, auf das bei allen registrierten Informationszugriffen nie zugegriffen wurde, so ist die Zugriffswahrscheinlichkeit auf dieses Objekt so gering, dass die Zugriffe auf das Objekt vernachlässigbar sind.

Bei der Bestimmung der Anzahl der pro Ort referenzierten Informationsobjekte kommt dem Ort „Tourist Information Center“ (TIC) insofern eine Sonderstellung zu, als dass es sich dabei nicht um eine einzelne touristische Attraktion handelt, sondern um ein Informationszentrum, das Informationen über alle Attraktionen anbietet. Entsprechend weicht die Anzahl der dort referenzierten Informationsobjekte im Vergleich zu den übrigen Orten nach oben ab. Für die übrigen Orte ergibt sich eine durchschnittliche Anzahl an referenzierten Informationsobjekten von 106.75. Diese Zahl ist in Abbildung 7.6(a) als Linie eingezeichnet.



**Abbildung 7.6:** Gesamtanzahl der unterschiedlichen, an einem Ort referenzierten Informationsobjekte (a) und durchschnittliche Anzahl der Informationszugriffe eines Benutzers je Ort (b).

### Mittlere Anzahl der Informationszugriffe eines Benutzers je Ort

Wie bereits bei der Betrachtung der an einem Ort referenzierten Informationsobjekte kommt dem Ort „Tourist Information Centre“ auch hier eine Sonderstellung zu. Entsprechend weicht die mittlere Anzahl der Informationszugriffe, die ein Benutzer an diesem Ort durchführt, von der an den übrigen Orten nach oben ab. Für die übrigen Orte ergibt sich ein Durchschnittswert von 23.357, der in Abbildung 7.6(b) wieder als Linie eingezeichnet ist.

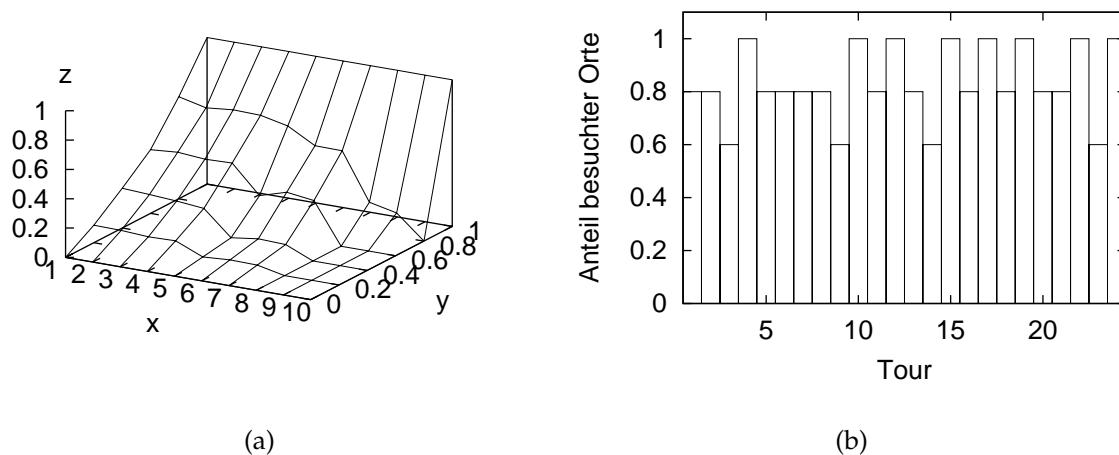
### Eigenschaften des GUIDE-Systems

Die in Abschnitt 7.1 beschriebenen Eigenschaften eines Informationssystems wurden mit Hilfe der dort definierten Metriken für das GUIDE-System quantifiziert. Damit wurden sowohl Anhaltspunkte für die in realen Systemen zu erwartenden Werte als auch Vergleichswerte für die entsprechenden Werte des Simulationsmodells gewonnen. Tabelle 7.1 fasst die Werte für das GUIDE-System zusammen.

Metrik	Wert
mittlere lokale Fokussierung	0.03169702349205264
Fokussierung	0.02475058228935759
mittlere Ortsbezogenheit	0.48851837817354743
Besuchsrate	0.8333333333333335

**Tabelle 7.1:** Werte der Metriken für das GUIDE-System.

Die Ortsbezogenheit der einzelnen Informationsobjekte ist recht unterschiedlich, weil, wie zuvor beschrieben wurde, auch Objekte vorkommen, die keinen ortsspezifischen Inhalt tragen und daher an (fast) allen Orten abgefragt werden. Abbildung 7.7(a) zeigt diese unterschiedlichen Ortsbezogenheiten anhand der Lorenzkurven für zehn exemplarische Informationsobjekte, die entsprechend ihrer Ortsbezogenheit aufsteigend sortiert sind. Auf der x-Achse sind die verschiedenen Objekte abgetragen. Die y-Achse gibt den Anteil der Orte an und auf der z-Achse sind die Trefferraten abgetragen.



**Abbildung 7.7:** Lorenzkurven zur Veranschaulichung der Ortsbezogenheit einzelner Informationsobjekte (a) und Anteil der besuchten Orte je Tour (b).

Die Besuchsrates ist sehr hoch. Wie Abbildung 7.7(b) zeigt, besuchen sehr viele Benutzer alle Orte. Keiner der Benutzer besucht weniger als drei der fünf möglichen Orte.

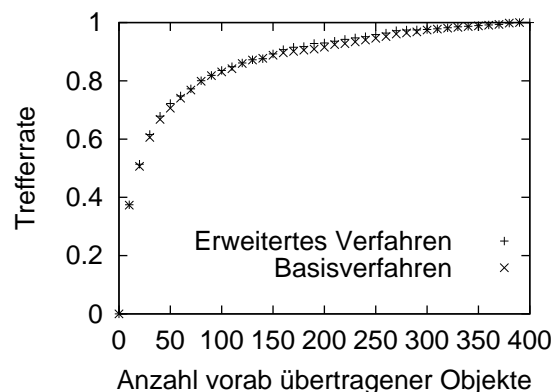
### 7.2.3 Vorabübertragung von Informationen im GUIDE-Szenario

Zwei Eigenschaften des GUIDE-Systems würden sich auf den Einsatz des in dieser Arbeit vorgestellten Verfahrens nachteilig auswirken. Das erste Problem ist, dass im GUIDE die Ortsinformation nur sehr grob ist. Insgesamt werden überhaupt nur fünf Orte unterschieden. Folglich wird fast jeder Ort von jedem Touristen besucht. Da das erweiterte Vorabübertragungsverfahren aber gerade Wissen über nicht besuchte Orte ausnutzt, um eine Übertragung der zugehörigen Informationen zu vermeiden, fällt der Vorteil dieses Verfahrens gegenüber dem Basisverfahren im GUIDE-Szenario sehr gering aus.

Ein weiteres Problem beim Einsatz der Vorabübertragungsverfahren sind die nicht ortsbezogenen Informationsobjekte, da sie fast immer vorab übertragen werden müssen. Hier könnte es Sinn machen, diese grundsätzlich auf den mobilen Endgeräten vorzuhalten, wodurch eine Vorabübertragung dann nicht mehr notwendig wäre.

Abbildung 7.8 zeigt die Trefferraten, die man mit beiden Vorabübertragungsverfahren im GUIDE-System erzielen würde. Um diese zu bestimmen, wurde davon ausgegangen, dass jeder Benutzer zu Beginn seiner Tour Zugang zu einer Infostation hat und

genügend Zeit im Empfangsbereich der Infostation verbringt, um den für die Vorabübertragung verfügbaren Speicherplatz vollständig auszunutzen. Für die Evaluierung des erweiterten Verfahrens wurde angenommen, dass jeder Ort in einer eigenen Zone liegt.



**Abbildung 7.8:** Durchschnittlich erzielte Trefferraten.

In Vorbereitung der Tests beider Verfahren wurden die Infostationen mit allen 24 verfügbaren Protokolldateien initialisiert. Das den Infostationen zur Verfügung stehende Wissen basierte also auf der Beobachtung von 24 Touren. Anschließend wurde für jede der 24 Touren getrennt bestimmt, welche Trefferraten mit beiden Verfahren bei einer gegebenen Anzahl an vorabübertragenen Informationsobjekten erzielt worden wäre. In Abbildung 7.8 ist jeweils der Durchschnitt der ermittelten Trefferraten angegeben.

Aufgrund des kleinen Informationsraums und der damit verbundenen hohen Fokussierung der Informationszugriffe werden trotz der hohen Besuchsrates bereits mit wenig vorabübertragenen Informationsobjekten hohe Trefferraten erzielt.

### 7.2.4 Bewertung

Mit der Analyse des GUIDE-Systems konnte lediglich ein erster Eindruck vermittelt werden, welche Trefferraten mit dem Vorabübertragungsverfahren erzielt werden können. Da es sich aber nur um eine konkrete Ausprägung eines mobilen Informationssystems handelt, konnte mit den GUIDE-Daten nicht festgestellt werden, wie sich die einzelnen Parameter eines Informationssystems auf die Trefferrate auswirken.

Ein weiterer Nachteil der GUIDE-Protokolldaten ist ihr vergleichsweise geringer Umfang von nur 24 Touren, der die Repräsentativität der Daten in Frage stellt. Ferner ist das GUIDE-System eher als ein kleines Testsystem zu betrachten, was sich zum einen in der geringen Größe des Informationsraums und zum anderen durch das vergleichsweise kleine Abdeckungsgebiet mit nur fünf touristischen Attraktionen zeigt.

Das in dieser Arbeit vorgestellte Verfahren zielt jedoch auf sehr große Informationsräume und Abdeckungsgebiete. Erst in solchen großen Informationssystemen ist es erforderlich bei der Auswahl vorab zu übertragender Informationsobjekte möglichst selektiv vorzugehen. Im erweiterten Verfahren wird auch von mehr unterscheidbaren Orten ausgegangen als dies im GUIDE mit seiner rein zellbasierten Lokalisierung der Fall ist. Insbesondere kann das erweiterte Verfahren seine Stärken erst bei einer geringeren Besuchsrates zeigen, die sich bei einem größeren Abdeckungsgebiet und einer genaueren Unterscheidung der Orte automatisch ergeben dürfte.

## 7.3 Simulationen

Die Bewertung der GUIDE-Analyse im vorherigen Abschnitt hat gezeigt, dass die Beurteilung der Vorabübertragungsverfahren allein im GUIDE-Szenario für eine vollständige Evaluierung der Verfahren nicht ausreichend ist. Daher wurden ergänzend die in diesem Abschnitt beschriebenen Simulationen durchgeführt. Dazu wurde die Simulationsumgebung und das Simulationsmodell, wie sie in Kapitel 6 vorgestellt wurden, verwendet. Die Wahl der Standardwerte für die Modellparameter wurde so getroffen, dass ein Informationssystem nachgebildet wird, wie es für den typischen Einsatz des Vorabübertragungsverfahrens erwartet wird. Bei dieser Wahl wurden die Erkenntnisse aus der Analyse des GUIDE-Systems berücksichtigt. In den Simulationen wurde dann jeweils einer der Modellparameter variiert, um dessen Auswirkung auf die erzielten Trefferraten zu analysieren.

### 7.3.1 Parametrisierung des Simulationsmodells

Für die Parametrisierung des Besuchsmodells wurden zehn Knoten des verwendeten Infrastrukturgraphen, die von besonderem Interesse für Touristen sind, als bevorzugte Knoten ausgewählt. Für die Besuchswahrscheinlichkeit der bevorzugten Knoten

wurde  $p_p = 0.3$  als Standardwert angenommen, im Falle nicht bevorzugter Knoten  $p_n = 0.0$ . Dies bedeutet nur, dass nicht bevorzugte Knoten nie als explizites Ziel einer Tour ausgewählt werden, aber dennoch besucht werden können.

Entsprechend den Erkenntnissen aus der GUIDE-Analyse wurde der Parameter für die Anzahl der einem bevorzugten Ort zugeordneten lokalen Informationsobjekte standardmäßig auf den Wert  $a_p = 100$  gesetzt. Da in den 106.75 durchschnittlich pro Ort referenzierten Objekten im GUIDE die globalen Informationsobjekte bereits beinhaltet sind, die in dem hier verwendeten Modell zu den lokalen Objekten noch hinzukommen, ist der Wert  $a_p = 100$  für das Vorabübertragungsverfahren pessimistisch, das heißt eher zu groß, gewählt. Für nicht bevorzugte Orte wurde ein Wert von  $a_n = 5$  angenommen. Diese Werte können zufällig um jeweils  $a_e = 25\%$  nach oben oder unten variieren. Die Größe des globalen Informationsvorrats betrug  $a_g = 1500$  Objekte.

Da in den Simulationen das mögliche Potenzial der Vorabübertragungsverfahren analysiert werden soll, wurde zunächst von einem inhärent ortsabhängigen System ausgegangen, das heißt die Zugriffswahrscheinlichkeit auf globale Informationen wurde mit  $p_g = 0.0$  angenommen. Später wurde dieser Parameter variiert und somit unterschiedlich stark ortsbezogene Systeme simuliert.

Die Anzahl der durchschnittlich pro Benutzer an einem bevorzugten Ort abgerufenen Informationsobjekte wurde auf den Wert  $r_p = 25$  Objekte gesetzt, für nicht bevorzugte Orte wurde  $r_n = 1$  Objekt gewählt. Diese Werte konnten wiederum um  $r_e = 25\%$  variieren, was im Falle nicht bevorzugter Orte keine Auswirkungen hat, da immer auf ganze Informationsobjekte gerundet werden muss.

Für den Parameter  $\alpha$  der Zipf-ähnlichen Verteilung der Zugriffswahrscheinlichkeiten wurde entsprechend den Erkenntnissen aus [Breslau et al. 1999] und der GUIDE-Analyse ein Wert von  $\alpha = 0.8$  gewählt. Die Länge des Aktualisierungsintervalls  $\Delta$  wurde auf 100 Protokolldateien gesetzt und der Historienparameter  $\gamma$  auf 0.5.

Bei der Evaluierung des erweiterten Verfahrens wurde im Allgemeinen davon ausgegangen, dass der genaue Pfad des Benutzers extern spezifiziert wird und dass diesem externen Wissen voll vertraut wird ( $\beta_i = 1$ ). Zusätzlich wurde angenommen, dass sich die Benutzer auch jeweils so verhalten, wie dies extern spezifiziert wurde und dass jeder Ort in einer eigenen Zone liegt.

Die Anzahl der vorab übertragbaren Informationsobjekte wurde mit 80 angenommen. Bei einer durchschnittlichen Objektgröße von 50 kbyte entspricht dies vier MBytes,



einem Wert bei dem davon ausgegangen werden kann, dass er auf heute handelsüblichen PDAs für die Speicherung vorab übertragener Informationen bereit gestellt werden kann.

Für die Simulationen wurde eine Infostation betrachtet, deren Hoarding-Gebiet den gesamten Infrastrukturgraphen abdeckt. In Vorbereitung einer jeden Simulation wurde die Infostation mit 1000 Protokolldateien initialisiert. Die Simulationen selbst bestanden aus der Simulation von jeweils 100 Touren unabhängiger Benutzer. Die Hoarding-Entscheidung basierte jeweils auf dem Wissen der Infostationen, wie es nach der Initialisierung vorlag. Für jede Tour wurde die erzielte Trefferrate bestimmt. Die Kurven zeigen jeweils die durchschnittliche Trefferrate, die im Verlauf der 100 Touren erzielt wurde.

Mit den Standardparametern ergeben sich für die einzelnen Metriken zur Beschreibung eines Informationssystems die Werte, wie sie in Tabelle 7.2 angegeben sind. Die mittlere lokale Fokussierung ist etwas höher als im GUIDE-System, weil an jedem Ort ausschließlich auf die ihm zugeordneten Informationsobjekte zugegriffen wird. Da jedoch der gesamte Informationsraum deutlich größer ist als im GUIDE ist die globale Fokussierung geringer. Für die Bestimmung der mittleren Ortsbezogenheit wurden nur die bevorzugten Orte betrachtet. Da der Zugriff auf globale Informationen zunächst ausgeschlossen wurde, ergibt sich, wie erwartet, für die Ortsbezogenheit der Maximalwert 0.9. Die Besuchsrates ist niedriger als im GUIDE-System, weshalb nun ein größerer Vorteil des erweiterten Verfahrens gegenüber dem Basisverfahren zu erwarten ist.

Metrik	Wert
mittlere lokale Fokussierung	0.05863611273856345
Fokussierung	0.002861545600416785
mittlere Ortsbezogenheit	0.9
Besuchsrates	0.11186956521739132

**Tabelle 7.2:** Werte der Metriken für das simulierte Informationssystem gemäß den Standardparameterwerten.

### 7.3.2 Ergebnisse

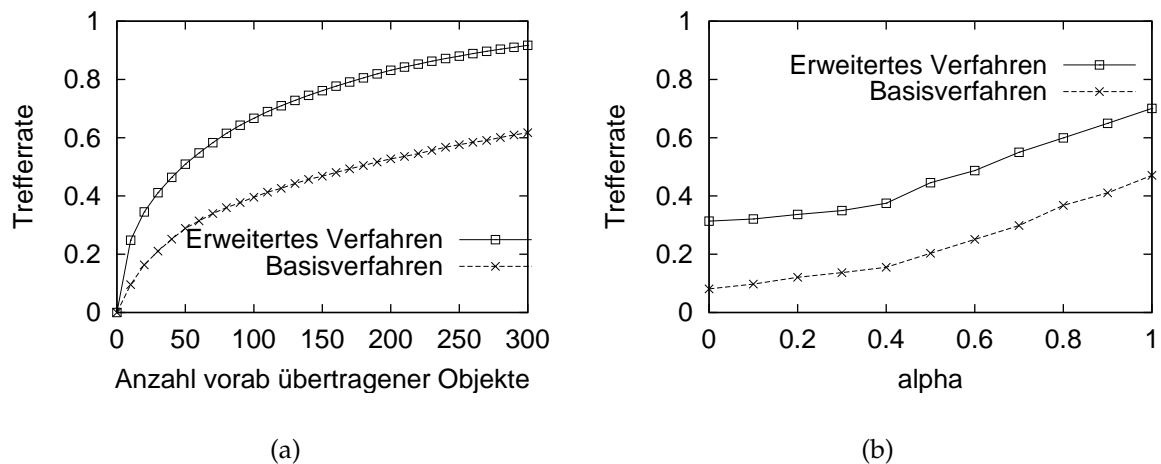
#### Anzahl vorab übertragener Informationsobjekte

Mit einer ersten Simulation wurde untersucht, welchen Einfluss die Zahl der vorab übertragenen Informationsobjekte auf die Trefferraten hat (siehe Abbildung 7.9(a)). Bereits mit dem Standardwert von 80 Objekten wird mit dem erweiterten Verfahren eine Trefferrate von 61.5% erreicht, während mit dem Basisverfahren 36% erreicht werden. Der Grund für den nicht linearen Anstieg der Kurven ist, dass die populärsten Informationsobjekte als erstes für eine Vorabübertragung ausgewählt werden, also auch bereits wenn nur wenige Objekte übertragen werden. Die Objekte, die mit einer steigenden Zahl vorab übertragener Objekte zusätzlich übertragen werden, sind immer weniger beliebt und entsprechend seltener wird darauf zugegriffen, so dass durch ihre Vorabübertragung die Trefferrate weniger stark zunimmt.

Zum Vergleich wurde auch die Trefferrate bestimmt, die man mit einem Caching-Verfahren erzielen würde. Aus Gründen, die bereits in Abschnitt 2.1.2 erläutert wurden, ist Caching als Datenverwaltungsmechanismus in ortsbezogenen Systemen weniger geeignet und entsprechend niedrig ist mit 22.9% die erzielte Trefferrate. Begünstigend wurde dabei angenommen, dass unbegrenzt viel Speicherplatz für das Caching zur Verfügung steht, so dass die ermittelte Trefferrate die maximal erzielbare Trefferrate unabhängig von der Ersetzungsstrategie darstellt.

#### Fokussierung

Um die Auswirkung unterschiedlicher Fokussierungen des Informationssystems auf die Trefferraten zu analysieren, wurde der Parameter  $\alpha$  der Zipf-ähnlichen Verteilung variiert (siehe Abbildung 7.9(b)). Ist  $\alpha = 1$ , konzentrieren sich sehr viele Zugriffe auf wenige sehr beliebte Informationsobjekte und es werden sehr hohe Trefferraten erzielt. Bei  $\alpha = 0$  hingegen sind die Trefferraten sehr niedrig, da die Zugriffswahrscheinlichkeiten dann gleich verteilt sind und beide Vorabübertragungsverfahren keinen Nutzen aus der Beobachtung der Zugriffswahrscheinlichkeiten ziehen können. Niedrige Werte von  $\alpha$  konnten bislang in der Praxis jedoch nicht beobachtet werden. In [Breslau et al. 1999] liegen die niedrigsten beobachteten Werte von  $\alpha$  bei 0.64.



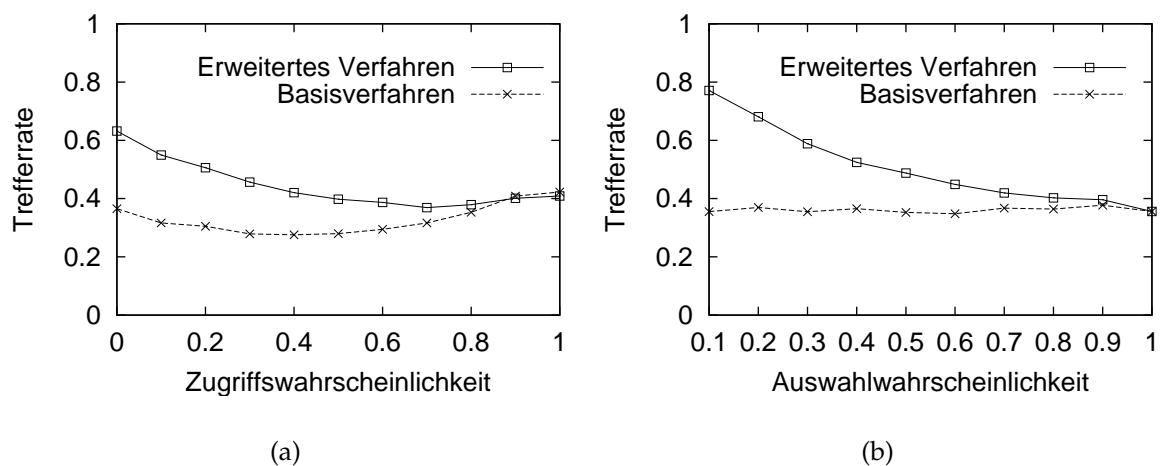
**Abbildung 7.9:** Trefferraten in Abhängigkeit von der Anzahl der übertragenen Informationsobjekte (a) und dem Parameter  $\alpha$  der Zipf-ähnlichen Zugriffswahrscheinlichkeitsverteilung (b).

### Ortsbezogenheit

Der Einfluss der Ortsbezogenheit des Informationssystems auf die Trefferraten wurde durch Variation der Zugriffswahrscheinlichkeit auf den globalen Informationsvorrat evaluiert (siehe Abbildung 7.10(a)). Solange die Ortsbezogenheit hoch ist, das heißt die Zugriffswahrscheinlichkeit auf den globalen Informationsvorrat ist gering, kann das erweiterte Verfahren sehr viel Nutzen aus der Kenntnis des Bewegungspfades ziehen und einen entsprechend großen Vorteil gegenüber dem Basisverfahren erzielen. Der Grund dafür ist, dass dann die Zuordnung der Objekte zu den Orten sehr deutlich ist und somit Objekte, die den nicht besuchten Orten zugeordnet sind, von der Vorabübertragung ausgeschlossen werden können. Mit einer zunehmenden Zugriffswahrscheinlichkeit auf den globalen Informationsvorrat geht diese Zuordnung mehr und mehr verloren, bis schließlich an allen Orten auf die gleichen Objekte zugegriffen wird. Dann liegt ein nicht ortsbezogenes Informationssystem vor und das erweiterte Verfahren ist dem Basisverfahren nicht mehr überlegen.

Wenn die Zugriffswahrscheinlichkeit auf den globalen Informationsvorrat null ist, beziehen sich alle Informationsanfragen auf eines der 1525 Informationsobjekte die einem Ort zugeordnet sind. Sobald jedoch die Zugriffswahrscheinlichkeit auf den globalen Informationsvorrat erhöht wird, werden zusätzlich immer mehr der 1500 Objek-

te des globalen Informationsvorrats referenziert. Dadurch sinkt die Fokussierung des Informationssystems und mit ihr auch die Trefferraten. Mit einer weiteren Erhöhung der Zugriffswahrscheinlichkeit auf globale Informationen beginnt jedoch eine zunehmende Konzentration auf ausschließlich die Objekte des globalen Informationsvorrats. Dadurch nimmt dann auch wieder die Fokussierung zu, was sich positiv auf die Trefferraten auswirkt. Mit größeren globalen Informationsvorräten ist die sich schließlich ergebende Fokussierung und daher auch der Anstieg der Trefferraten mit steigenden Zugriffswahrscheinlichkeiten schwächer. Auch der Wendepunkt von fallenden zu steigenden Trefferraten wird dann erst später erreicht.



**Abbildung 7.10:** Trefferraten in Abhängigkeit von der Zugriffswahrscheinlichkeit auf den globalen Informationsvorrat (a) und der Auswahlwahrscheinlichkeit bevorzugter Orte (b).

### Besuchsrate

Um unterschiedliche Besuchsraten zu erhalten, wurde die Wahrscheinlichkeit, mit der bevorzugte Orte als Ziel einer Tour ausgewählt werden, schrittweise von 0,1 auf 1,0 erhöht (siehe Abbildung 7.10(b)). Wenn die Auswahlwahrscheinlichkeit klein ist, werden nur wenige Orte besucht und das erweiterte Verfahren kann auf Grund der Kenntnis des Bewegungspaths sehr viele Informationen, die zu den vielen unbesuchten Orten gehören, von der Vorabübertragung ausschließen. Entsprechend groß ist der Vorteil des erweiterten Verfahrens gegenüber dem Basisverfahren.

Mit steigender Auswahlwahrscheinlichkeit gibt es jedoch immer weniger unbesuchte Orte, womit im erweiterten Verfahren weniger Objekte von der Vorabübertragung ausgeschlossen werden können. Daher sinkt die mit dem erweiterten Verfahren erzielte Trefferrate bis sie schließlich nur noch so hoch ist wie die Trefferrate des Basisverfahrens. Die Trefferrate des Basisverfahrens ist unabhängig von der Auswahlwahrscheinlichkeit, weil hier immer für alle Orte Informationen übertragen werden.

### **Größe des Informationsraums**

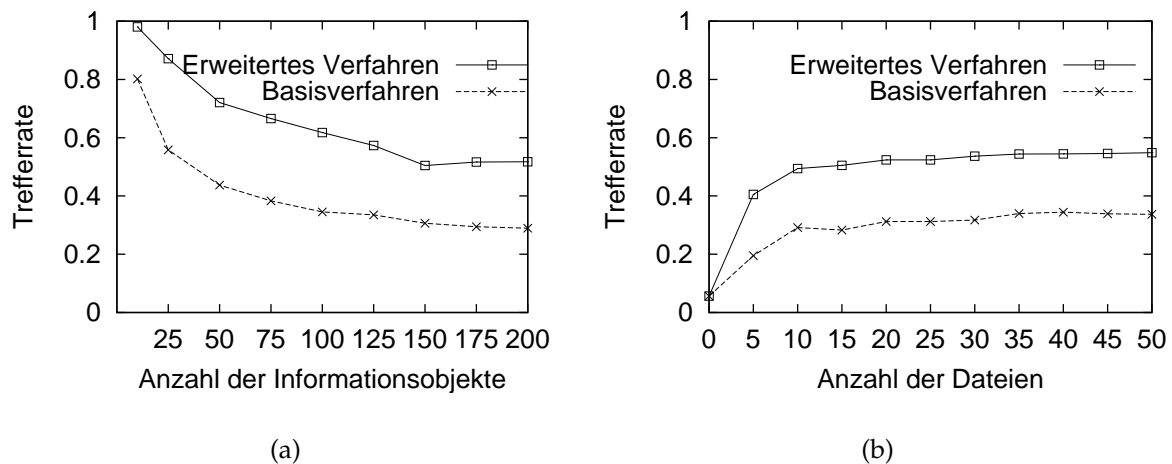
Natürlich beeinflusst auch die Größe des Informationsraums die erzielten Trefferraten. Bei größeren Informationsräumen und gleicher Anzahl vorab übertragener Informationsobjekte ist ein kleinerer Anteil der Objekte im Informationsraum lokal verfügbar. Folglich werden sich mehr Anfragen auf lokal nicht vorhandene Informationsobjekte beziehen und damit die Trefferrate sinken. In der Simulation, deren Ergebnisse in Abbildung 7.11(a) dargestellt sind, wurde die Größe des Informationsraums dadurch verändert, dass den bevorzugten Orten unterschiedlich viele Informationsobjekte zugewiesen wurden.

Die Trefferraten sinken nicht linear mit der Zunahme der Größe des Informationsraums, da aufgrund der angenommen Zipf-ähnlichen Verteilung der Zugriffswahrscheinlichkeiten, auf zusätzliche Informationsobjekte in bereits sehr großen Informationsräumen nur selten zugegriffen wird.

### **Lerngeschwindigkeit**

Wenn keine initialen Zugriffswahrscheinlichkeitstabellen und Besuchswahrscheinlichkeitskarten durch den Betreiber einer Infostation angegeben werden können, muss ein neu gestarteter Infostation-Server das Benutzerverhalten zunächst kennenlernen, bevor zuverlässige Hoarding-Entscheidungen getroffen werden können. Um festzustellen, wieviele Protokolldateien analysiert werden müssen, bis akzeptable Hoarding-Entscheidungen getroffen werden können, wurde eine weitere Simulation durchgeführt.

In dieser Simulation wurde die Infostation mit einer zunehmenden Zahl  $n$  von Protokolldateien initialisiert. Wobei das Aktualisierungsintervall  $\Delta$  abweichend von der



**Abbildung 7.11:** Trefferraten in Abhängigkeit von der Anzahl der Objekte, die bevorzugten Orten zugeordnet sind (a), und der Anzahl der initial registrierten Protokolldateien (b).

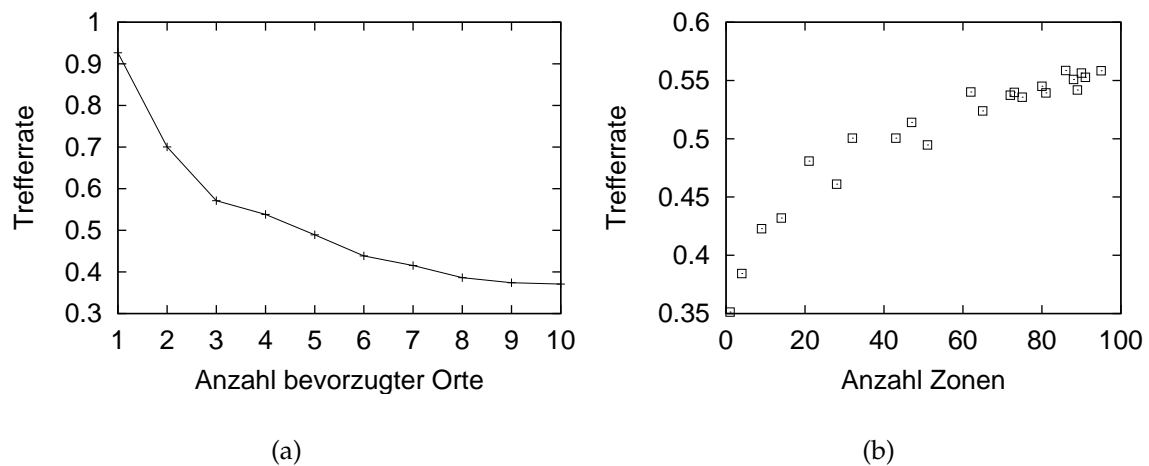
Standardlänge jeweils so gewählt wurde, dass genau nach Registrierung der  $n$  Protokolldateien eine Aktualisierung der Tabellen und Karten durchgeführt wurde. Wie Abbildung 7.11(b) zeigt, kann bereits nach Registrierung von nur zehn Protokolldateien, mit annähernd den maximalen Trefferraten gerechnet werden. Auch wenn noch gar keine Protokolldatei übertragen wurde, ist die erzielte Trefferrate größer als null. Dies liegt daran, dass in diesem Fall zufällig ausgewählte Informationsobjekte übertragen werden, auf die gelegentlich zugegriffen wird.

### Größe des Hoarding-Gebiets

Der Einfluss der Größe eines Hoarding-Gebiets auf die Trefferrate ist in Abbildung 7.12(a) dargestellt. Die Kurve wurde anhand von Simulationen ermittelt, bei denen die Größe des Hoarding-Gebiets jeweils so gewählt wurde, dass eine bestimmte Anzahl bevorzugter Orte innerhalb des Hoarding-Gebiets liegt.

Abweichend vom Standardwert wurde die Auswahlwahrscheinlichkeit bevorzugter Orte auf 1.0 gesetzt, um sicherzustellen, dass auch in jeder Tour das gesamte Hoarding-Gebiet bereist wird und somit zu gewährleisten, dass auch in sehr kleinen Hoarding-Gebieten noch genügend Knoten besucht werden, um statistisch fundierte Aussagen zu ermöglichen. Da stets alle bevorzugten Orte besucht werden, ergeben sich mit dem

erweiterten Verfahren die gleichen Trefferraten wie mit dem Basisverfahren. Folglich wurde diese Simulationsreihe nur für das Basisverfahren durchgeführt.



**Abbildung 7.12:** Trefferraten in Abhängigkeit von der Größe des Hoarding-Gebiets (a) und der Anzahl der Zonen (b).

### Größe der Zonen

Um den Einfluss der Zonengröße auf das erweiterte Verfahren zu ermitteln, wurde eine Reihe von Simulationen mit Zonen unterschiedlicher Größe durchgeführt. In jeder Simulation wurde das Hoarding-Gebiet in ein Raster rechteckiger Zonen gleicher Größe unterteilt. Dieses Raster wurde von Simulation zu Simulation verfeinert, so dass die Zahl der Zonen zunahm und ihre Größe entsprechend abnahm.

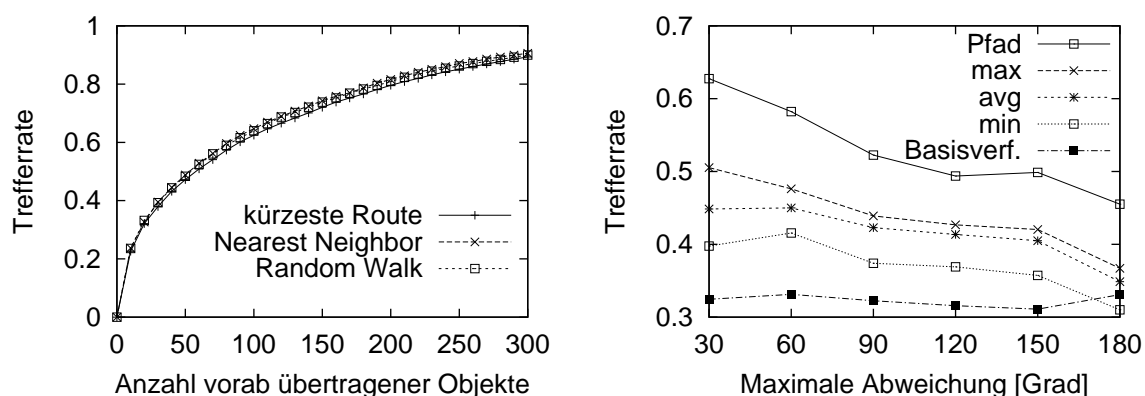
Durch die Aufteilung des Hoarding-Gebiets in ein Raster entstanden teilweise auch Zonen, in denen sich keine Knoten des Infrastrukturgraphen befanden. Da solche Zonen für das Vorabübertragungsverfahren nutzlos sind, wurden sie nicht weiter betrachtet. Daher sind in Abbildung 7.12(b) bei der Anzahl der Zonen nur solche Zonen berücksichtigt, in denen sich auch tatsächlich Knoten befanden.

Selbstverständlich nimmt die Trefferrate mit einer Erhöhung der Zonenzahl tendenziell zu (siehe auch Abschnitt 4.5.2), dennoch kam es in den Simulationen vor, dass eine Erhöhung der Zonenzahl zu einer Reduzierung der mittleren Trefferrate führte.

Der Grund dafür ist, dass die Zonen dann nach der Erhöhung der Zonenzahl ungünstiger lagen als zuvor und so zum Beispiel zwei bevorzugte Orte, die zuvor noch in getrennten Zonen lagen, von der gleichen Zone abgedeckt wurden.

## Pfadberechnung

Man könnte vermuten, dass die Annahme, dass alle Besucher ihre Ziele stets auf der kürzesten möglichen Route besuchen, die erzielten Trefferraten begünstigen, weil stets die gleichen Wege benutzt werden. Tatsächlich ist jedoch nur entscheidend, welche Knoten besucht werden, nicht aber in welcher Reihenfolge dies geschieht. Um dies zu überprüfen, wurde zusätzlich ein „Random Walk“-Modell und eine „Nearest Neighbor“-Strategie zur Bestimmung der Besuchsreihenfolge verwendet. Die erzielten Simulationsergebnisse (siehe Abbildung 7.13(a)) bestätigen, dass die Besuchsreihenfolge keinen Einfluss auf die Trefferraten hat. Die Abbildung zeigt exemplarisch die Ergebnisse für das erweiterte Verfahren.



**Abbildung 7.13:** Einfluss der Besuchsreihenfolge auf die Trefferraten (a) und Nutzen von unscharfem Wissen (b).

## Unscharfes Wissen

Bislang wurde das erweiterte Verfahren nur im Zusammenhang mit exakt spezifizierten Pfaden benutzt. Ist jedoch nur unschärferes externes Wissen verfügbar, ist zu erwarten, dass die erzielten Trefferraten zwischen denen des Basisverfahrens und denen,



die mit exaktem Pfadwissen erzielt werden, liegen. Um dies zu verifizieren, wurde eine weitere Simulation durchgeführt.

Für diese Simulation wurde das Besuchsmodell geändert, so dass sich die simulierten Benutzer vorzugsweise in südlicher Richtung bewegen. Dazu wurden die Auswahlwahrscheinlichkeiten der bevorzugten Orte geändert. Ein Startknoten im Norden des Infrastrukturgraphen (Hauptbahnhof) bekam die Auswahlwahrscheinlichkeit eins zugewiesen. Die weiteren Auswahlwahrscheinlichkeiten wurden so gewählt, dass alle bevorzugten Orte, die innerhalb eines vom Startknoten aus nach Süden gerichteten Kegels liegen, eine Auswahlwahrscheinlichkeit größer null erhielten, wobei die Wahrscheinlichkeiten nach außen hin abnahmen. Allen anderen Orten wurde die Wahrscheinlichkeit null zugewiesen. Der Öffnungswinkel des Kegels war ein Simulationsparameter und wurde auf Werte zwischen  $30^\circ$  und  $180^\circ$  gesetzt.

Die Spezifikation von unscharfem externem Wissen wurde unter Verwendung relativer und absoluter Besuchswahrscheinlichkeitskarten simuliert. Die relativen Karten spezifizierten wiederum für einen vom Startknoten aus nach Süden gerichteten Kegel unterschiedliche relative Besuchswahrscheinlichkeiten, die ebenfalls zum Kegelrand hin abnahmen. Die absoluten Besuchswahrscheinlichkeitskarten spezifizierten eine Besuchswahrscheinlichkeit von null für alle außerhalb des Kegels liegenden Orte. Auch hier wurde der Öffnungswinkel des Kegels variiert.

Für jedes sich aus der Variation des Öffnungswinkels ergebende Besuchsmodell wurde eine eigene Simulationsreihe durchgeführt. Zunächst wurde die Trefferrate für das erweiterte Verfahren mit exakt spezifiziertem Pfad und das Basisverfahren ermittelt. Danach wurde das unscharfe externe Wissen evaluiert. Hierzu wurden jeweils verschiedene Karten mit den Öffnungswinkeln  $15^\circ$ ,  $30^\circ$ ,  $60^\circ$ ,  $90^\circ$ ,  $120^\circ$ ,  $150^\circ$  und  $180^\circ$  spezifiziert. Für jedes Besuchsmodell, das sich aus einem bestimmten Öffnungswinkel ergibt, wurden also jeweils eine ganze Reihe von Karten mit verschiedenen Öffnungswinkel spezifiziert und die entsprechenden Trefferraten ermittelt.

Abbildung 7.13(b) zeigt die Trefferraten in Abhängigkeit des Öffnungswinkels, der für das Besuchsmodell verwendet wurde. Die Kurve *max* gibt die maximale Trefferrate an, die mit unscharfem Wissen erzielt wurde, wenn also der Öffnungswinkel für die externen Karten optimal gewählt wurde. Entsprechend gibt *min* die schlechteste Trefferrate an, wenn also der spezifizierte Öffnungswinkel sehr ungünstig gewählt wurde, zum Beispiel Spezifikation eines Öffnungswinkels von  $180^\circ$ , wenn die Benutzer nur

maximal um  $30^\circ$  von der südlichen Richtung abweichen. Die Kurve avg gibt schließlich den Mittelwert der Trefferraten an, die mit den unterschiedlichen spezifizierten Öffnungswinkeln erreicht wurden. Wie erwartet liegen die Trefferraten, die mit unscharfem Wissen erzielt werden, zwischen den Werten des Basisverfahrens und den mit exaktem Pfadwissen erzielten Werten.

Der teilweise eckige Verlauf der ermittelten Kurven ist dadurch zu erklären, dass die Orte, die im Infrastrukturgraphen repräsentiert sind, nicht gleichmäßig in der Ebene verteilt sind, so dass eine Erhöhung des Öffnungswinkels um den gleichen Wert zu einer unterschiedlichen Zunahme an im Mittel besuchten Orten führen kann.

## Kapitel 8

# Zusammenfassung und Ausblick

In dieser Arbeit wurden zwei Verfahren, ein Basisverfahren und ein erweitertes Verfahren, zur Vorabübertragung ortsbezogener Informationen vorgestellt. Sie ermöglichen es, bei einem mobilen Informationszugriff weitgehend oder vollständig auf teure und unzuverlässige Kommunikationsverbindungen über drahtlose Weitverkehrsnetze zu verzichten. Da für die verwendete Infostationsinfrastruktur lediglich standardisierte, gut verfügbare und günstige Komponenten benötigt werden, kann ein potenzieller Informationsanbieter einfach und kostengünstig eine solche Infrastruktur installieren und damit einen Zugriff auf seine Informationen unabhängig von Mobilfunknetzbetreibern ermöglichen.

Im Basisverfahren weiß jede Infostation nur, welche Informationen in ihrer Umgebung vorzugsweise referenziert werden. Daher ist die einzige Information über die Bewegungen eines Benutzers, die in diesem Verfahren ausgenutzt werden kann, die Tatsache, dass sich der Benutzer zum Zeitpunkt der Vorabübertragung an einer bestimmten Infostation befindet. Entsprechend wird er, so weit wie möglich, mit allen Informationen versorgt, die in der Umgebung dieser Infostation benötigt werden. Eventuell vorhandenes Wissen über die zukünftigen Aufenthaltsorte des Benutzers kann im Basisverfahren nicht genutzt werden.

Die wichtigste Innovation des erweiterten Verfahrens ist es, die lokalen Zugriffswahrscheinlichkeiten in beliebig kleinen Gebieten und deren Besuchswahrscheinlichkeiten getrennt voneinander zu beobachten. Durch diese Entkoppelung von Zugriffs- und

Besuchswahrscheinlichkeiten ist es möglich, die beobachteten Besuchswahrscheinlichkeiten durch externe Angaben zu ersetzen. So kann in den Anwendungen oder beim Benutzer verfügbares Wissen über dessen zukünftige Aufenthaltsorte ausgenutzt werden, um die Vorhersage der benötigten Informationen zu präzisieren. Die Schnittstelle für diese externen Angaben ist sehr generisch gehalten, so dass sehr unterschiedliche Angaben, beispielsweise ein exakter Bewegungspfad oder auch nur eine bevorzugte Bewegungsrichtung, möglich sind. Dennoch kann das Verfahren auch gänzlich ohne externe Angaben auskommen, dann ist es jedoch dem Basisverfahren nicht mehr überlegen. Das erweiterte Verfahren kann also ein breites Spektrum von externem Wissen nutzen, ohne die Verfügbarkeit von solchem Wissen vorauszusetzen. Damit ist es sehr viel flexibler als vergleichbare Ansätze.

Ein Vorteil beider Verfahren ist, dass keine feste Zuordnung von Informationen zu Orten vorausgesetzt werden muss, wie dies in verwandten Ansätzen der Fall ist, sondern dass diese Zuordnung durch Beobachtung des Benutzerverhaltens selbständig abgeleitet werden kann. Die Verfahren adaptieren sich sogar dynamisch an mögliche Änderungen in dieser Zuordnung.

Neben den beiden Verfahren wurden Erweiterungen vorgestellt, die es erlauben, das Profil eines Benutzers und individuelle Eigenschaften des Benutzers oder seines Endgeräts bei der Entscheidung, welche Informationsobjekte vorab übertragen werden sollen, zu berücksichtigen.

Beide Verfahren wurden vollständig implementiert und sowohl anhand von theoretischen Betrachtungen als auch anhand von umfassenden Simulationen evaluiert. Um die für die Effizienz der Verfahren relevanten Eigenschaften eines Informationssystems quantifizieren zu können, wurden eigens spezielle Metriken entwickelt. Für die Simulationen musste ein Simulationsmodell aufgestellt werden, das die Beziehungen zwischen Orten und Informationen widerspiegelt, und eine entsprechende Simulationsumgebung implementiert werden. Die Simulationsergebnisse zeigen, dass die Ausnutzung externer Informationen über den zukünftigen Aufenthaltsort eines Benutzers im Allgemeinen sehr vorteilhaft ist.

Auch wenn Mobilfunknetze zukünftig höhere Datenraten bieten werden, werden die Vorteile der drahtlosen lokalen Netze gegenüber den Mobilfunknetzen erhalten bleiben. Sie werden auch weiterhin stets höhere Datenraten bei geringeren Infrastrukturkosten bieten. Damit wird auch der Einsatz der in dieser Arbeit vorgestellten Verfahren

---

in Verbindung mit einer Infostationsinfrastruktur stets von Vorteil sein, insbesondere zur Unterstützung neuer Anwendungen, die immer umfangreichere Daten, wie zum Beispiel Video oder 3D-Modelle, benötigen werden.

Interessante Aspekte für zukünftige Forschungsarbeiten könnten sich aus einer stärkeren Spezialisierung der Verfahren auf bestimmte Systeme, wie zum Beispiel die neXus-Plattform, ergeben. In neXus könnte man beispielsweise versuchen, neben den Ortsinformationen weitere Kontextinformationen, die möglicherweise aus dem neXus-Modell gewonnen werden können, für eine Vorabübertragung auszunutzen. Möchte man vollständige Teilmodelle aus dem neXus-Modell vorab übertragen, stellen sich weitere interessante Fragen, zum Beispiel ob Informationen über mobile Objekte vorab übertragen werden sollten und falls ja, wie diese konsistent gehalten werden können. Man kann sich auch vorstellen, das Modell in unterschiedlichen Detaillierungsgraden vorzuhalten. Dadurch würde sich die Auswahlmöglichkeit bieten, entweder einen kleinen geographischen Ausschnitt aus dem Modell mit hohem Detaillierungsgrad vorab zu übertragen oder einen größeren geographischen Ausschnitt in einem geringeren Detaillierungsgrad. Auch Abhängigkeiten zwischen einzelnen Objekten des Modells müssten beachtet werden. Sonst könnte es vorkommen, dass ein Objekt vorab übertragen wird, das von anderen nicht übertragenen Objekten, zum Beispiel auf Grund einer räumlichen Inklusion, abhängig ist und dessen isolierte Übertragung daher keinen Sinn macht.

Die in der Arbeit bereits erwähnten übergeordneten Strategien zur Kooperation mit anderen Datenverwaltungsmechanismen stellen sicherlich eine weitere interessante Aufgabe dar (siehe Abschnitt 5.1.2). Auch die Kooperation zwischen verschiedenen Hoarding-Klienten, die sich zufällig treffen und über ein Ad-Hoc-Netzwerk Informationen austauschen können, ist denkbar.

Die Erprobung der entwickelten Verfahren im praktischen Einsatz, beispielsweise einem größeren Stadtinformationssystem, wäre ebenfalls sehr interessant. Insbesondere könnte die Akzeptanz bei den Benutzern überprüft werden, sowie deren Vertrauen in den Informationsanbieter beziehungsweise die eventuell eingesetzten Anonymisierungstechniken.

Das notwendige Vertrauen des Benutzers entweder gegenüber dem Informationsanbieter oder gegenüber den verwendeten Anonymisierungstechniken zeigt, dass neben den noch offenen technischen Fragen, die in naher Zukunft sicherlich bald gelöst wer-

den, auch Fragen nach den gesellschaftlichen Folgen eines allgegenwärtigen, mobilen Informationszugriffs beantwortet werden müssen. Schon heute können die Mobilfunknetzbetreiber sehr genau beobachten, an welchen Orten sich ihre Kunden aufhalten, Betreiber von WWW-Servern können beobachten, welche Informationen von welchen Computern aus abgerufen werden. Durch die Möglichkeit eines mobilen Informationszugriffs werden diese beiden Aspekte zusammengeführt, so dass dann beobachtet werden kann, wo und wann auf welche Informationen zugegriffen wird. Damit drängt sich zum Beispiel die Frage auf, ob mit der, ohne Zweifel, faszinierenden Möglichkeit zu jeder Zeit und überall auf Informationen (information anytime-anywhere) zugreifen zu können, auch ein weiterer Schritt hin zum gläsernen Menschen der modernen Informationsgesellschaft unternommen wird.

## Anhang A

# Glossar und Abkürzungsverzeichnis

### **Abdeckungsgebiet einer Infostation**

Geographisches Gebiet innerhalb dessen das →WLAN einer →Infostation verfügbar ist.

### **Abdeckungsgebiet eines Informationssystems**

Geographisches Gebiet innerhalb dessen auf das Informationssystem zugegriffen werden kann.

### **Absolute Besuchswahrscheinlichkeitskarte**

Funktion, die → Zonen eines →Hoarding-Gebiets absolute Besuchswahrscheinlichkeiten zuordnet.

### **AP – Access Point**

→Zugangspunkt.

### **Besuchsmodell**

Teil des in dieser Arbeit verwendeten Simulationsmodells, der die Besuchswahrscheinlichkeiten von Orten festlegt.

### **Bewegungsbedarfsmodell**

Modell zur Beschreibung der Mobilität von Benutzern auf Grundlage des Bedarfs für Bewegungen.

### **Bewegungsmuster**

Wiederholt, mit einer gewissen Regelmäßigkeit auftretende Bewegung.

**BWK – Besuchswahrscheinlichkeitskarte**

Oberbegriff für →absolute Besuchswahrscheinlichkeitskarten und →relative Besuchswahrscheinlichkeitskarten.

**EDGE – Enhanced Data Rates for Global Evolution**

Protokoll zur paketorientierten Datenübertragung in →GSM-Netzen.

**Entkoppelter Betrieb**

Betriebsart, in der ein mobiles Endgerät ohne Netzanbindung betrieben wird.

**Externes Wissen**

Wissen über die zukünftigen Aufenthaltsorte eines Benutzers, das dem erweiterten Vorabübertragungsverfahren von externen Informationsquellen bereitgestellt wird.

**GPRS – General Packet Radio System**

Protokoll zur paketorientierten Datenübertragung in →GSM-Netzen.

**GPS – Global Positioning System**

Satelliten-gestütztes System zur Bestimmung geographischer Positionen.

**GSM – Global System for Mobile Communications**

Weit verbreiteter Mobilfunkstandard der zweiten Generation (digitale Mobilkommunikation).

**Hoarding**

Vorabübertragung von Informationen, in der Regel mit dem Ziel, einen →entkoppelten Betrieb zu ermöglichen.

**Hoarding-Entscheidung**

Entscheidung, welche Informationen auf das mobile Endgerät eines Benutzers übertragen werden sollen, um seine zukünftigen mobilen Informationszugriffe so gut wie möglich zu unterstützen.

**Hoarding-Gebiet**

Geographisches Gebiet, für das eine →Infostation Informationen bereitstellt.

**Hoarding-Verfahren**

Verfahren zur Vorabübertragung von Informationen, in der Regel mit dem Ziel, einen →entkoppelten Betrieb zu ermöglichen.

**HSCSD – High Speed Circuit Switched Data**

Protokoll zur leitungsvermittelten Datenübertragung in →GSM-Netzen.



---

## **HTTP – Hypertext Transfer Protocol**

Im World Wide Web verwendetes Protokoll zur Übertragung von Dokumenten.

## **Informationsmodell**

Teil des in dieser Arbeit verwendeten Simulationsmodells, der die Beziehungen zwischen Informationen und Orten festlegt.

## **Informationsobjekt**

Einzelne, diskrete Informationseinheit, die von einem Informationssystem bereitgestellt wird.

## **Informationsraum**

Gesamtheit der von einem Informationssystem bereitgestellten →Informationsobjekte.

## **Infostation**

Aus einem →WLAN und einem Software-Prozess bestehende Einheit, die eine breitbandige Datenübertragung in einem Gebiet ermöglicht, in dem ansonsten nur eine schmalbandige oder gar keine Datenübertragung möglich ist.

## **Infostationsinfrastruktur**

Eine Menge von →Infostationen, die über ein geographisches Gebiet verteilt sind und miteinander kommunizieren können.

## **Infrastrukturgraph**

Graph zur Repräsentation der Orte und Verbindungswege in einem geographischen Gebiet.

## **LBS – Location-Based Service**

→ortsbezogener Dienst.

## **Mobile IP**

Erweiterung des Internet-Protokolls zur Unterstützung mobiler Endgeräte.

## **Ortsabhängiger Informationszugriff**

Informationszugriff, bei dem die Zugriffswahrscheinlichkeiten auf einzelne →Informationsobjekte eines →Informationsraums vom Aufenthaltsort des Benutzers abhängen.

## **Ortsbezogener Dienst**

Dienst, der es seinen Benutzern ermöglicht, in Abhängigkeit von ihrem Aufenthaltsort auf ortsspezifische Informationen zuzugreifen.

**Ortsbezogenes Informationssystem**

Informationssystem, dessen →Informationsraum aus →ortsbezogenen Informationsobjekten besteht.

**Ortsbezogenes Informationsobjekt**

Informationsobjekt, bei dem die Wahrscheinlichkeit, mit der es angefordert wird, vom Aufenthaltsort des Benutzers abhängt.

**PDA – Personal Digital Assistant**

Mobiler Kleinrechner, der meist zur Termin- und Adressverwaltung eingesetzt wird.

**Prefetching-Verfahren**

Verfahren zur Vorabübertragung von Informationen, in der Regel ausschließlich mit dem Ziel, die Antwortzeit eines Systems zu reduzieren.

**Random-Walk-Model**

Modell zur Beschreibung der Mobilität von Benutzern, bei dem ein zufallsgesteuerter Bewegungsablauf angenommen wird.

**Relative Besuchswahrscheinlichkeitskarte**

Funktion, die →Zonen eines →Hoarding-Gebiets relative Besuchswahrscheinlichkeiten zuordnet.

**Schwach gekoppelter Betrieb**

Betriebsart, in der ein mobiles Endgerät mit schmalbandiger Netzanbindung betrieben wird.

**SNMP – Simple Network Management Protocol**

Protokoll zur Überwachung und Verwaltung von Rechnernetzen.

**Travel Demand Model**

→Bewegungsbedarfsmodell.

**Trefferrate**

Maß zur Beurteilung von Vorabübertragungsverfahren, das das Verhältnis der Informationsanfragen, welche mit vorab übertragenen Informationen beantwortet werden können, zu der Gesamtzahl der Informationsanfragen angibt.

**UMTS – Universal Mobile Telecommunications System**

Mobilfunkstandard der dritten Generation.

---

### **WAP – Wireless Application Protocol**

Das Wireless Application Protocol definiert einen Standard für die Bereitstellung von text- und grafikbasierten Informationen und Diensten für mobile Endgeräte.

### **WLAN – Wireless Local Area Network**

Drahtloses lokales Netz.

### **Zipf-ähnliche Verteilung**

Wahrscheinlichkeitsverteilung, bei der das  $i$ . wahrscheinlichste Ereignis mit der Wahrscheinlichkeit  $p(i) = \frac{\Omega_i}{i^\alpha}$  eintritt.

### **Zone**

Geographisch begrenzter Teil eines →Hoarding-Gebiets. Zwei verschiedene Zonen überlappen sich nicht gegenseitig.

### **Zugangspunkt**

Übergangspunkt zwischen einem drahtlosen Netzwerk und dem Festnetz.

### **Zugriffsmodell**

Teil des in dieser Arbeit verwendeten Simulationsmodells, der das Zugriffsverhalten der Benutzer beschreibt.

### **ZWT – Zugriffswahrscheinlichkeitstabelle**

Tabelle, die die Zugriffswahrscheinlichkeiten auf einzelne →Informationsobjekte angibt.

## Anhang B

# Berechnung des Gini-Koeffizienten

Wie in Abschnitt 3.1 beschrieben, ist der Gini-Koeffizient definiert als das Verhältnis der Fläche zwischen der Hauptdiagonalen und der Lorenzkurve zur Fläche zwischen der Hauptdiagonalen und der Abszissen-Achse. Da hier nur das Verhältnis von Flächen betrachtet wird, kann in der nachfolgenden Rechnung auf Längen- und Flächeneinheiten verzichtet werden. Die verwendeten Bezeichner sind ebenfalls in Abschnitt 3.1 beschrieben.

Zunächst kann der Inhalt  $A_L$  der Fläche unter der Lorenzkurve wie folgt berechnet werden:

$$\begin{aligned} A_L &= \sum_{j=1}^N \frac{v_j + v_{j-1}}{2} \cdot (u_j - u_{j-1}) \\ &= \frac{1}{2} \cdot \sum_{j=1}^N \frac{\sum_{i=1}^j x_i + \sum_{i=1}^{j-1} x_i}{n} \cdot \left( \frac{j}{N} - \frac{j-1}{N} \right) \\ &= \frac{1}{N \cdot n} \cdot \sum_{j=1}^N \left( \sum_{i=1}^j x_i \right) - \frac{1}{2N} \\ &= \frac{1}{N \cdot n} \cdot \sum_{i=1}^N ((N - i + 1) \cdot x_i) - \frac{1}{2N} \\ &= \frac{2N + 1}{2N} - \frac{\sum_{i=1}^N i \cdot x_i}{N \cdot n} \end{aligned}$$

---

Damit kann schließlich der Gini-Koeffizient  $GK$  berechnet werden:

$$\begin{aligned} GK &= \frac{\text{Fläche zwischen Hauptdiagonale und Lorenzkurve}}{\text{Fläche zwischen Hauptdiagonale und Abszissen-Achse}} \\ &= 2 \cdot \text{Fläche zwischen Hauptdiagonale und Lorenzkurve} \\ &= 2 \cdot \left( \frac{1}{2} - A_L \right) \\ &= 1 - 2 \cdot \left( \frac{2N+1}{2N} - \frac{\sum_{i=1}^N i \cdot x_i}{N \cdot n} \right) \\ &= \frac{2 \cdot \sum_{i=1}^N i \cdot x_i}{N \cdot n} - \frac{N+1}{N} \end{aligned}$$

# Literaturverzeichnis

- Abowd, G.; Atkeson, C. G.; Hong, J.; Long, S.; Kooper, R.; Pinkerton, M. (1997). Cyberguide: A Mobile Context-Aware Tour Guide, *ACM Journal on Wireless Networks* 3(5): 421–433.
- Acharya, S.; Alonso, R.; Franklin, M.; Zdonik, S. (1995). Broadcast Disks: Data Management for Asymmetric Communication Environments, *Proceedings of the 1995 ACM SIGMOD International Conference on Management of Data*, San Jose, CA, USA, pp. 199–210.
- Alonso, R.; Barbará, D.; Cova, L. (1990). Using Stashing to Increase Node Autonomy in Distributed File Systems, *Proceedings of the Ninth Symposium on Reliable Distributed Systems*, Huntsville, AL, USA, pp. 12–21.
- Badrinath, B. R.; Imieliński, T.; Frenkiel, R.; Goodman, D. (1996). NIMBLE: Many-Time, Many-Where Communication Support for Information Systems in Highly Mobile and Wireless Environments, <http://www.cs.rutgers.edu/~badri/dataman/nimble/>.
- Barbará, D.; Imieliński, T. (1994). Sleepers and Workaholics: Caching Strategies in Mobile Environments, *Proceedings of the 1994 ACM SIGMOD International Conference on Management of Data*, Minneapolis, MN, USA, pp. 1–12.
- Berners-Lee, T.; Masinter, L.; McCahill, M. (1994). Uniform Resource Locators (URL), RFC 1738.
- Bohley, P. (1996). *Statistik - einführendes Lehrbuch für Wirtschafts- und Sozialwissenschaftler*, Oldenbourg, München.

- 
- Breslau, L.; Cao, P.; Li, F.; Philips, G.; Shenker, S. (1999). Web Caching and Zipf-like Distributions: Evidence and Implications, *Proceedings of the IEEE INFOCOM '99*, New York City, NY, USA, pp. 126–134.
- Case, J. D.; Fedor, M.; Schoffstall, M.; Davin, C. (1990). A Simple Network Management Protocol (SNMP), RFC 1157.
- Chan, B.; Si, A.; Leong, H. (1998). Cache Management for Mobile Databases: Design and Evaluation., *Proceedings of the IEEE International Conference on Data Engineering (ICDE '98)*, Orlando, FL, USA, pp. 54–63.
- Chang, H.; Tait, C.; Cohen, N.; Shapiro, M.; Mastrianni, S.; Floyd, R.; Housel, B.; Lindquist, D. (1997). Web Browsing in a Wireless Environment: Disconnected and Asynchronous Operation in ARTour Web Express, *Proceedings of the Third Annual ACM/IEEE International Conference on Mobile Computing and Networking (MobiCom '97)*, Budapest, Ungarn, pp. 260–269.
- Cheverst, K.; Davies, K.; Mitchell, K.; Friday, A. (2000). Experiences of Developing and Deploying a Context-Aware Tourist Guide: The GUIDE project, *Proceedings of the Sixth Annual ACM SIGMOBILE International Conference on Mobile Computing and Networking (MobiCom 2000)*, Boston, MA, USA, pp. 20–31.
- Cheverst, K.; Davies, N.; Mitchell, K.; Friday, A.; Efstratiou, C. (2000). Developing a Context-Aware Electronic Tourist Guide: Some Issues and Experiences, *Proceedings of ACM Conference on Human Factors in Computing Systems 2000 (CHI 2000)*, Den Haag, Niederlande, pp. 17–24.
- Chim, J.; Green, M.; Lau, R.; Leong, H.; Si, A. (1998). On Caching and Prefetching of Virtual Objects in Distributed Virtual Environments, *Proceedings of the Sixth ACM International Conference on Multimedia (ACM MULTIMEDIA '98)*, Bristol, Großbritannien, pp. 171–180. [http://www.acm.org/sigmm/MM98/electronic\\_proceedings/chim/index.html](http://www.acm.org/sigmm/MM98/electronic_proceedings/chim/index.html).
- Crovella, M.; Bestavros, A. (1996). Self-Similarity in World Wide Web Traffic: Evidence and Possible Causes, *Proceedings of the 1996 ACM SIGMETRICS Conference on Measurement & Modeling of Computer Systems*, Philadelphia, PA, USA, pp. 160–169.
- Davidson, S.; Garcia-Molina, H.; Skeen, D. (1985). Consistency in Partitioned Networks: A Survey, *ACM Computer Surveys* **17**(3): 341–370.

- Davies, N.; Cheverst, K.; Mitchell, K.; Friday, A. (1999). Caches in the Air: Disseminating Information in the Guide System, *Proceedings of the Second IEEE Workshop on Mobile Computing Systems and Applications (WMCSA '99)*, New Orleans, LA, USA, pp. 11–19.
- Davies, N.; Mitchell, K.; Cheverst, K.; Blair, G. (1998). Developing a Context Sensitive Tourist Guide, *Proceedings of the First Workshop on Human Computer Interaction with Mobile Devices*, Universität Glasgow, Großbritannien, pp. 64–68.
- de Nitto Personè, V.; Grassi, V.; Morlupi, A. (1998). Modeling and Evaluation of Prefetching Policies for Context-Aware Information Services, *Proceedings of the Fourth Annual ACM/IEEE International Conference on Mobile Computing and Networking (MobiCom '98)*, Dallas, Texas, USA, pp. 55–64.
- Demers, A. J.; Petersen, K.; Spreitzer, M. J.; Terry, D. B.; Theimer, M. M.; Welch, B. B. (1994). The Bayou Architecture: Support for Data Sharing among Mobile Users, *Proceedings of the First Workshop on Mobile Computing Systems and Applications (WMCSA '94)*, Santa Cruz, CA, USA, pp. 2–7.
- Eberspächer, J.; Vögel, H.-J. (1998). *GSM Switching, Services, and Protocol*, Teubner, Stuttgart.
- Effelsberg, W.; Härder, T. (1984). Principles of Database Buffer Management, *ACM Transactions on Database Systems* 9(4): 560–595.
- Ericsson GmbH (2000). Von GSM zu UMTS - Informationen über neue Mobilfunktechniken, Broschüre, [http://www.ericsson.de/broschueren/von\\_gsm\\_zu\\_umts.pdf](http://www.ericsson.de/broschueren/von_gsm_zu_umts.pdf).
- Federrath, H.; Pfitzmann, A. (1998). Neue Anonymitätstechniken - Eine vergleichende Übersicht, *Datenschutz und Datensicherheit* 22(11): 628–632. Vieweg, Wiesbaden.
- Fielding, R.; Gettys, J.; Mogul, J. C.; Frystyk, H.; Masinter, L.; Leach, P.; Berners-Lee, T. (1999). Hypertext Transfer Protocol - HTTP/1.1, RFC 2616.
- Gamma, E.; Helm, R.; Johnson, R.; Vlissides, J. (1997). *Design Patterns – Elements of Reusable Object-Oriented Software*, Addison-Wesley, Reading, MA, USA.
- Gergeleit, M. (2001). *Das drahtlose Ethernet*, Addison-Wesley, München.



- 
- Glassman, S. (1994). A Caching Relay for the World Wide Web, *Proceedings of the First International Conference on the World-Wide Web (WWW '94)*, Genf, Schweiz. <http://www.cern.ch/WWW94/>.
- Goodman, D.; Borràs, J.; Mandayam, N.; Yates, R. (1997). INFOSTATIONS: A New System Model for Data and Messaging Services, *Proceedings of the 47th Annual IEEE Vehicular Technology Conference (VTC '97)*, Phoenix, AZ, USA, pp. 969–973.
- Gwertzman, J.; Seltzer, M. (1995). The Case for Geographical Push-Caching, *Proceedings of the Fifth Workshop on Hot Topics in Operating Systems (HOTOS-V)*, Rosario, WA, USA, pp. 51–55.
- Han, R.; Bhagwat, P.; LaMaire, R.; Mummert, T.; Perret, V.; Rubas, J. (1998). Dynamic Adaptation in an Image Transcoding Proxy For Mobile Web Browsing, *IEEE Personal Communications Magazine* 5(6): 8–17.
- Hegele, M. (2000). *Entwicklung einer Simulationsumgebung für kontextbezogenen Informationszugriff in einer mobilen Umgebung*, Studienarbeit Nr. 1790, Universität Stuttgart.
- Hegele, M. (2001). *Entwurf und Implementierung einer Architektur zur Vorabübertragung ortsabhängiger Informationen in einer mobilen Umgebung*, Diplomarbeit Nr. 1890, Universität Stuttgart.
- Heidemann, J.; Thomas, W.; Guy, R.; Popek, G. (1992). Primarily Disconnected Operation: Experiences with Ficus, *Proceedings of the Second Workshop on Management of Replicated Data*, Los Angeles, CA, USA, pp. 2–5.
- Heise-Newsticker (2001). AOL-Studie prophezeit mobilem Internet großes Potenzial, <http://www.heise.de/newsticker/data/klp-18.04.01-003/>.
- Hennessy, J. L.; Patterson, D. A. (1990). *Computer Architecture. A Quantitative Approach*, Morgan Kaufmann Publishers, Inc., San Mateo, CA, USA.
- Hofmann-Wellenhof, B.; Lichtenegger, H.; Collins, J. (2001). *Global Positioning System. Theory and Practice*, Springer Verlag, Wien.
- Hohl, F.; Kubach, U.; Leonhardi, A.; Rothermel, K.; Schwehm, M. (1999). Next Century Challenges: Nexus - An Open Global Infrastructure for Spatial-Aware Applica-

- tions, *Proceedings of the Fifth Annual ACM/IEEE International Conference on Mobile Computing and Networking (MobiCom '99)*, Seattle, WA, USA, pp. 249–255.
- Housel, B.; Lindquist, D. (1996). WebExpress: A System for Optimizing Web Browsing in a Wireless Environment, *Proceedings of the Second Annual ACM/IEEE International Conference on Mobile Computing and Networking (MobiCom '96)*, Rye, New York, USA, pp. 108–116.
- Howard, J. H. (1988). An Overview of the Andrew File System, *Proceedings of the USE-NIX Winter 1988 Technical Conference*, Dallas, TX, USA, pp. 23–36.
- Hässler, A. (2000). Entwicklung einer GPS-Bibliothek in Java, Studienarbeit Nr. 1774, Universität Stuttgart.
- Huston, L.; Honeyman, P. (1993). Disconnected Operation for AFS, *Proceedings of the USENIX Mobile and Location-Independent Computing Symposium*, Cambridge, MA, USA, pp. 1–10.
- Imieliński, T.; Viswanathan, S.; Badrinath, B. R. (1994). Energy Efficient Indexing on Air, *Proceedings of the 1994 ACM SIGMOD International Conference on Management of Data*, Minneapolis, MN, USA, pp. 25–36.
- INRO Consultants Inc. (2000). EMME/2 - Transportation Planning System, <http://www.inro.ca/>.
- Jain, R.; Lin, Y.-B.; Lo, C.; Mohan, S. (1994). A Caching Strategy to Reduce Network Impacts of PCS, *IEEE Journal on Selected Areas in Communications* **12**(8): 1434–1444.
- Joseph, A. D.; Tauber, J. A.; Kaashoek, M. F. (1997). Mobile Computing with the Rover Toolkit, *IEEE Transactions on Computers* **46**(3): 337–352.
- Kistler, J.; Satyanarayanan, M. (1992). Disconnected Operation in the Coda File System, *ACM Transactions on Computer Systems* **10**(1): 3–25.
- Kubach, U.; Hegele, M.; Rothermel, K. (2001). A Simulation Framework for Mobile, Location-Dependent Information Access, *Proceedings of the Sixth IEEE Symposium on Computers and Communications (ISCC '01)*, Hammamet, Tunisia, pp. 506–511.
- Kubach, U.; Rothermel, K. (2000). An Adaptive, Location-Aware Hoarding Mechanism, *Proceedings of the Fifth IEEE Symposium on Computers and Communications (ISCC 2000)*, Antibes, France, pp. 615–620.

- 
- Kubach, U.; Rothermel, K. (2001a). Ein Hoarding-Verfahren für ortsbezogene Informationen, *Informatik - Forschung und Entwicklung, Springer* . Noch nicht erschienen.
- Kubach, U.; Rothermel, K. (2001b). Exploiting Location Information for Infostation-Based Hoarding, *Proceedings of the Seventh Annual ACM SIGMOBILE International Conference on Mobile Computing and Networking (MobiCom '01)*, Rome, Italy, pp. 15–27.
- Kubach, U.; Rothermel, K. (2001c). A Map-Based Hoarding Mechanism for Location-Dependent Information, *Proceedings of the Second International Conference on Mobile Data Management (MDM '01)*, Hong Kong, China, pp. 145–157.
- Kuenning, G.; Popek, G. (1997). Automated Hoarding for Mobile Computers, *Proceedings of the 16th ACM Symposium on Operating Systems Principles (SOSP '97)*, St. Malo, France, pp. 264–275.
- Lawler, E.; Lenstra, J.; Rinnooy Kan, A.; Shmoys, D. (1985). *The Traveling Salesman Problem*, Wiley-Interscience, Chichester, Großbritannien.
- Leite, F.; Pereira, J. (2001). Location-Based Services and Emergency Communications in IMT-2000, *ITU News* . No. 07/2001: 4-13.
- Leonhardi, A.; Kubach, U.; Rothermel, K. (1999). Virtual Information Towers - A Metaphor for Intuitive, Location-Aware Information Access in a Mobile Environment, *Proceedings of the Third International Symposium on Wearable Computers (ISWC '99)*, San Francisco, CA, USA, pp. 15–20.
- Leonhardi, A.; Rothermel, K. (2001). Architecture of a Large-Scale Location Service, *Technical Report TR 2001-01*, Fakultät Informatik, Universität Stuttgart.
- Li, K.; Hudak, P. (1989). Memory Coherence in Shared Virtual Memory Systems, *ACM Transactions on Computer Systems (TOCS)* 7(4): 321–359.
- Liu, G.; Marlevi, A.; Maguire Jr., G. (1996). A Mobile Virtual-Distributed System Architecture for Supporting Wireless Mobile Computing and Communications, *ACM Journal on Wireless Networks* 2(1): 77–86.
- Liu, G.; Maguire Jr., G. (1995). Efficient Mobility Management Support for Wireless Data Services, *Proceedings of the 45th Annual IEEE Vehicular Technology Conference (VTC '95)*, Chicago, IL, USA, pp. 902–906.

- Liu, T.; Bahl, P.; Chlamtac, I. (1998). Mobility Modeling, Location Tracking, and Trajectory Prediction in Wireless ATM Networks, *IEEE Journal on Selected Areas in Communications* **16**(6): 922–936.
- Messmer, J. (2001). *Modellierung der Augmented World in neXus*, Diplomarbeit Nr. 1870, Universität Stuttgart.
- Miller, B.; Bisdikian, C. (2000). *Bluetooth Revealed: The Insider's Guide to an Open Specification for Global Wireless Communications*, Prentice Hall, Englewood Cliffs, NJ, USA.
- Mohan, R.; Smith, J. R.; Li, C.-S. (1999). Adapting Multimedia Internet Content for Universal Access, *IEEE Transactions on Multimedia* **1**(1): 104–114.
- Nishikawa, N.; Hosokawa, T.; Mori, Y.; Yoshida, K.; Tsuji, H. (1998). Memory-Based Architecture for Distributed WWW Caching Proxy, *Proceedings of the Seventh International World Wide Web Conference (WWW7)*, Brisbane, Australia, pp. 205–214.
- Oppenheim, N. (1995). *Urban Travel Demand Modeling: From Individual Choices to General Equilibrium*, Wiley-Interscience, New York, NY, USA.
- Pascoe, J. (1997). The Stick-e Note Architecture: Extending the Interface Beyond the User, *Proceedings of the 1997 International Conference on Intelligent User Interfaces (IUI '97)*, Orlando, FL, USA, pp. 261–264.
- Patterson, R.; Gibson, G.; Satyanarayanan, M. (1993). A Status Report on Research in Transparent Informed Prefetching, *ACM Operating Systems Review* **27**(2): 21–34.
- Perkins, C. (1996). IP Mobility Support, RFC2002.
- Pricewaterhouse-Coopers (2001). Technology Forecast 2001 - 2003, <http://www.pwcglobal.com/tech-forecast>.
- Reichart, M. (2000). *Workflow Management in Mobile Computing Environments*, Diplomarbeit Nr. 1831, Universität Stuttgart.
- Ren, Q.; Dunham, M. (2000). Using Semantic Caching to Manage Location Dependent Data in Mobile Computing, *Proceedings of the Sixth Annual ACM SIGMOBILE International Conference on Mobile Computing and Networking (MobiCom 2000)*, Boston, MA, USA, pp. 210–221.

- 
- Rodriguez, P.; Biersack, E. W. (1998). Bringing the Web to the Network Edge: Large Caches and Satellite Distribution, *Proceedings of the ACM/IEEE Workshop on Satellite-Based Information Services (WOSBIS'98)*, Dallas, TX, USA.
- Satyanarayanan, M.; Kistler, J.; Kumar, P.; Okasaki, M.; Siegel, E.; Steere, D. (1990). Co-da: A Highly Available File System for a Distributed Workstation Environment, *IEEE Transactions on Computers* **39**(4): 447–459.
- Satyanarayanan, M.; Kistler, J.; Mummert, L.; Ebling, M. R.; Kumar, P.; Lu, Q. (1993). Experience with Disconnected Operation in a Mobile Computing Environment, *Proceedings of the USENIX Mobile and Location-Independent Computing Symposium*, Cambridge, MA, USA, pp. 11–28.
- Scourias, J.; Kunz, T. (1999). An Activity-Based Mobility Model and Location Management Simulation Framework, *Proceedings of the Second ACM International Workshop on Modeling, Analysis and Simulation of Wireless and Mobile Systems (MSWiM '99)*, Seattle, WA, USA, pp. 61–68.
- Sistla, A.; Wolfson, O.; Huang, Y. (1998). Minimization of Communication Cost Through Caching in Mobile Environments, *IEEE Transactions on Parallel and Distributed Systems* **9**(4): 378–390.
- Smith, A. J. (1985). Disk Cache-Miss Ratio Analysis and Design Considerations, *ACM Transactions on Computer Systems* **3**(3): 161–203.
- Tait, C. D.; Lei, H.; Acharya, S.; Chang, H. (1995). Intelligent File Hoarding for Mobile Computers, *Proceedings of the First Annual ACM/IEEE International Conference on Mobile Computing and Networking (MobiCom'95)*, Berkeley, CA, USA, pp. 119–125.
- Tan, K.-L.; Ooi, B. C. (2000). *Data Dissemination in Wireless Computing Environments*, Kluwer Academic Publishers, Norwell, MA, USA.
- Walke, B. (2000a). *Mobilfunknetze und ihre Protokolle - Band 1*, Teubner, Stuttgart.
- Walke, B. (2000b). *Mobilfunknetze und ihre Protokolle - Band 2*, Teubner, Stuttgart.
- Witt, M. (ed.) (2000). *GPRS - Start in die mobile Zukunft*, mitp, Bonn.
- Ye, T.; Jacobsen, H.-A.; Katz, R. (1998). Mobile Awareness in a Wide Area Wireless Network of Info-Stations, *Proceedings of the Fourth Annual ACM/IEEE International*

*Conference on Mobile Computing and Networking (MobiCom '98)*, Dallas, TX, USA, pp. 109–120.

Zipf, G. (1949). *Human Behavior and the Principle of Least Effort*, Addison-Wesley, Cambridge, MA, USA.

Zonoozi, M.; Dassanayake, P. (1997). User Mobility Modeling and Characterization of Mobility Patterns, *IEEE Journal on Selected Areas in Communications* **15**(7): 1239–1252.