

Fokussierte Informationsadaption am Beispiel des Ortsbezuges

Astrid Lubinski

Universität Rostock
FB Informatik

Lubinski@informatik.uni-rostock.de

Abstract: In vielen mobilen Anwendungen, in denen Personen mit einem autonom arbeitenden mobilen Gerät unterwegs sind, werden spezielle Anforderungen an die Informationsauswahl gestellt. Einerseits sollen die Informationen auf das Nutzerumfeld eingestellt sein und sich gemäß seiner Mobilität anpassen und andererseits sollen die Informationen auch ohne ständige Netzverbindung verfügbar sein. Mobiler Replikation jedoch stehen immer noch begrenzende mobile Ressourcen entgegen.

In diesem Paper wird eine fokussierte und entsprechend dem Abstand von Fokus unterschiedlich reduzierte Auswahl und Replikation von Informationen in Datenbanksystemen ausgeführt, die einerseits die gewünschte Adaption an den mobilen Kontext realisiert und darüber hinaus den mobilen Ressourcen besser gerecht zu werden vermag als eine generelle Replikation aller angefragten Informationen.

1 Einleitung

Mobiler Zugriff auf Datenbanksysteme unterscheidet sich von stationären Arbeitsumgebungen, unabhängig davon, ob lediglich der Nutzer mobil ist oder ob er zur Unterstützung seiner Mobilität auch mobile Hardware nutzt. Die Umgebungscharakteristika, die die mobile Arbeit beeinflussen und i.d.R. beschränken, haben wir im *mobilen Kontext* zusammengefasst. Die Kontexte:

1. *Personen*, 2. *Ressourcen* und 3. *Informationen* charakterisieren das Umfeld des mobil arbeitenden Nutzers. Zentraler Kontext ist der *Ort*, an dem sich ein Nutzer aufhält und der sich in der *Zeit*-Dimension ändert (siehe auch [HeLu98]).

Diesen mobilen Kontext gilt es, möglichst transparent in die Nutzerarbeit zu integrieren, um dessen mobile Tätigkeit nicht zu behindern, sondern im Gegenteil zu unterstützen. Wechselnde Kontexte — und diese sind bei mobilen Nutzern intendiert — erfordern eine dynamische und möglichst vorausschauende Adaption von Daten und Applikationen an den mobilen Kontext. Aufgrund des starken Ortsbezugs ist das direkte Umfeld des Nutzers von vordringlichem Interesse für ihn. Mit größerer Entfernung vom Fokus schwindet die Relevanz der Informationen.

Es soll hier ein Modell eines fokussierten Zugriffes entwickelt werden, das detaillierte Informationen für das direkte Umfeld des Nutzers liefert und alle weiteren Informationen abhängig von ihrer Entfernung vom Fokus stärker abstrahiert repliziert. Dieses „Fisheye-View“-ähnliche Modell soll nicht begrenzt auf Bilddaten sondern allgemeingültig für in Datenbanken verwaltete Daten entworfen werden. Vorteil des Modells ist neben der Detailtreue der Information im direkten Fokus des Nutzers, in mobiler Umgebung z.B. zum Aufenthaltsort eines Touristen,

einer Arbeitsaufgabe oder speziellen Interessen, dass die weiteren Informationen ressourcenschonend reduziert als eine Art „Vorabinformation“ repliziert werden.

In diesem Artikel soll das Modell am Beispiel des Ortskontextes als Fokusgröße dargestellt werden. Zunächst werden dazu kurz die zugrunde liegende Orts- und Reduktionshierarchien der Informationen beschrieben. Die Grundidee der Arbeit besteht darin, abhängig vom Fokus eine Hierarchiestufe in der Ortshierarchie auszuwählen und diese mit einer geeigneten Reduktionsstufe der Informationen zu verbinden. Die Informationen werden mit dem entsprechenden Reduktionsverfahren abstrahiert. Ein Ortsbezug der Informationen und die Kenntnis darüber, welche Attribute der Datenbank diesen beinhalten, wird ebenso vorausgesetzt wie die technischen Möglichkeiten zur Ortsbestimmung. Probleme der Replikation und Replikatkonsistenz werden hier nicht betrachtet (siehe dazu z.B. [Pito96][LuHe00],[Goll02]).

2 Ortsfokussierte Datenreduktion

2.1 Ortsrelevanz

Viele mobile Tätigkeiten sind ortsrelevant, d.h., ihre Ausprägungen hängen von dem Ort ab, wo sie erfolgen. Die dazugehörigen Informationen sind ebenfalls ortsrelevant und lassen sich ebenso wie die Orte hierarchisch darstellen.

Ein Ort I_1 , z.B. eine Stadt, läßt sich gliedern in mehrere überschneidungsfreie Stadtteile und diese wiederum in Straßen und Hausnummern. Benachbarte Orte sind auf jeder Hierarchiestufe ermittelbar. I_5 und I_6 ergeben z.B. zusammen das Areal von I_2 : $I_5 \cup I_6 = I_2$; $I_2 \cup I_3 \cup I_4 = I_1$ u.s.w. (vgl. Abb. 1).

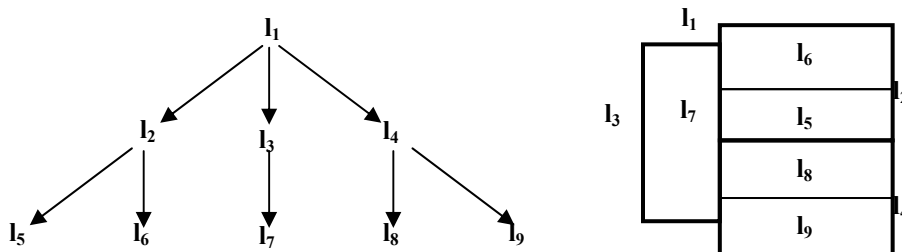


Abb.1: Beispiel-Ortshierarchie als Baum und als Fläche (vereinfacht)

2.2 Reduktionsmethoden und Reduktionsgraphen

Ebenso wie die Vereinigung der Ortsbereiche den auf einer jeweils höheren Ebene liegenden Ortsbereich ergibt, ist eine dazugehörige Information i die Aggregation der zu den unterliegenden Ortsbereichen I_i zugehörigen Informationen i_i . Die Einzelinformationen zu den Sehenswürdigkeiten der Stadtteile $1..n$ ergeben zusammengefasst die Informationen zu den Sehenswürdigkeiten der Stadt insgesamt.

Ein solches Zusammenfügen von Informationen hat ein Ergebnis, das so groß ist, wie die Summe aller Größen der Informationen i_i , die mit den Blättern des Ortshierarchiebaumes korrelieren (z.B. i_5 mit I_5). Eine solche Antwortmenge

($\mathbf{d} = \mathbf{i}_5 \cup \mathbf{i}_6 \cup \mathbf{i}_7 \cup \mathbf{i}_8 \cup \mathbf{i}_9$) kann sehr groß sein, insbesondere, wenn sie multimediale Informationen enthält. Sowohl die auf Nutzeranforderung replizierten als auch die prädiktiv geladenen Daten stehen so in Diskrepanz zu den verfügbaren Ressourcen und sind inhaltlich zu breit angelegt, wenn sie auf die ganze Stadt bezogen sind, bzw. zu einschränkend betrachtet, wenn die niedrigste Hierarchiestufe, die Straße, in der sich der Tourist aufhält, gewählt wurde. Eine sinnvolle stufenweise Reduktion der Datenmenge kann Abhilfe schaffen.

Abb.2 zeigt allgemein die Methoden zur Datenreduktion, die mit Datenbanksystemmitteln umsetzbar sind (siehe detaillierter in [HeLu98],[DRR97]).

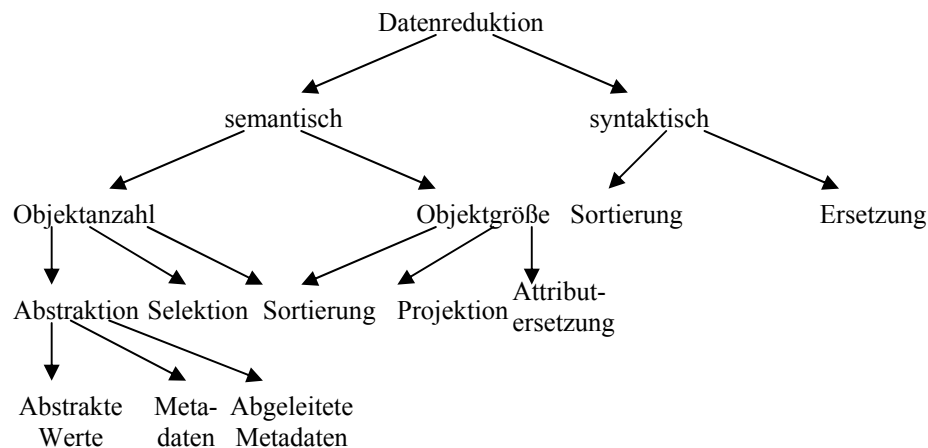


Abb.2: Datenreduktionsmethoden

Nicht-inhaltsbasierte Methoden ersetzen die Objekte ohne Beachtung ihrer Semantik und ohne eine angemessene, Semantik-beachtende Anpassung vorzunehmen. Es ist vorstellbar, an dieser Stelle medienspezifische Reduktionstechniken wie z.B. Text Retrieval-Methoden zu integrieren.

Reduktionsmaßnahmen können in verschiedener Weise wirken:

- die Daten werden selbst schon in mehrfachen Reduktionsstufen gespeichert
- die Anfrage wird geeignet verändert oder
- das Anfrageergebnis wird reduziert.

Die Anfrage wird verändert, indem Projektionen oder Selektionen hinzugefügt bzw. die zugriffenen Objekte ausgetauscht oder auch sortiert werden. Sortierungen sind Voraussetzung für eine progressive nutzergesteuerte Reduktion. Ersetzungen von Daten erfolgen zumeist aufgrund des Medientyps. Neben der Reduktion der Daten kommt eine Größenverringern einhergehend mit einer Funktionsreduktion der mobilen Datenbankmanagementsysteme in Betracht, wie z.B. in [LeHö02] behandelt, die für Informationsadaptionen jedoch nicht weiter relevant ist.

Abb.3 zeigt in der Mitte eine Beispielhierarchie reduzierter Informationen. Auf das initiale Anfrageergebnis \mathbf{d}_0 lassen sich jeweils einige der Reduktionsmethoden, die hier beispielhaft durchnummeriert als \mathbf{r}_1 bis \mathbf{r}_6 dargestellt sind, sinnvoll anwenden. Durch die Anwendung der Reduktionsmethode \mathbf{r}_j auf Informationen \mathbf{d}_i verringert sich sowohl die Größe der resultierenden Information \mathbf{d}_{i+1} als auch deren (subjektiver) Informationsgehalt. Es lassen sich vom syntaktischen als auch semantischen Standpunkt aus gültige bzw. sinnvolle Tupel $(\mathbf{d}_i, \mathbf{r}_j)$ definieren. Die Gültigkeit ergibt sich durch die Anwendbarkeit. Es ist nicht

gültig bzw. sinnvoll, eine reduzierende Projektion bei Aussparung der Schlüsselattribute oder Bild-reduzierende Methoden auf Fließtext anzuwenden.

Die Reduktion kann auf \mathbf{d}_{i+1} fortgesetzt werden, theoretisch bis zum maximalen Reduktionsergebnis \mathbf{d}_n , wo keinerlei Information mehr vorliegt, die dann auch niemandem nützt. Von dieser „Nicht-Information“ abgesehen ist in Abb.3 \mathbf{d}_5 die am stärksten reduzierte Information, während die Reduktion mit \mathbf{r}_1 zu \mathbf{d}_2 kaum Einsparungen erbringt. Es lassen sich „Reduktionsketten“ (z.B. $\mathbf{d}_0 \rightarrow \mathbf{d}_2 \rightarrow \mathbf{d}_5 \rightarrow \mathbf{d}_n$) bilden, deren Teilergebnisse mittels der Reduktionsmethoden \mathbf{r}_1 und zweimaliger Anwendung von \mathbf{r}_3 erzeugt wurden. Der Übergang $\mathbf{d}_2 \rightarrow \mathbf{d}_5$ ist durch das Entfernen von z.B. Videoclips zu Sehenswürdigkeiten sehr groß im Gegensatz zum Übergang $\mathbf{d}_3 \rightarrow \mathbf{d}_4$, wo vielleicht eine Selektionsbedingung hinzugefügt wurde.

2.3 Fokusgesteuerter Datenzuschnitt

Da, wie bereits festgestellt wurde, keine gleichmäßige Reduktion des Gesamtergebnisses angestrebt wird, sondern fokusabhängig die Reduktionsstärke variiert werden soll, werden die Teilergebnisse \mathbf{i}_5 bis \mathbf{i}_9 jeweils separat reduziert. Vereinfacht wird dies, indem \mathbf{n} ineinander geschachtelte Fokusbereiche ähnlich weit vom Fokus entfernter Ortsbereiche gebildet werden. Hierzu sind Ähnlichkeits- und Abstandsmaße einsetzbar, wobei an dieser Stelle eine vereinfachte Form beschrieben wird. Die Informationen der Fokusbereiche werden nun graduell abgestuft detailliert:

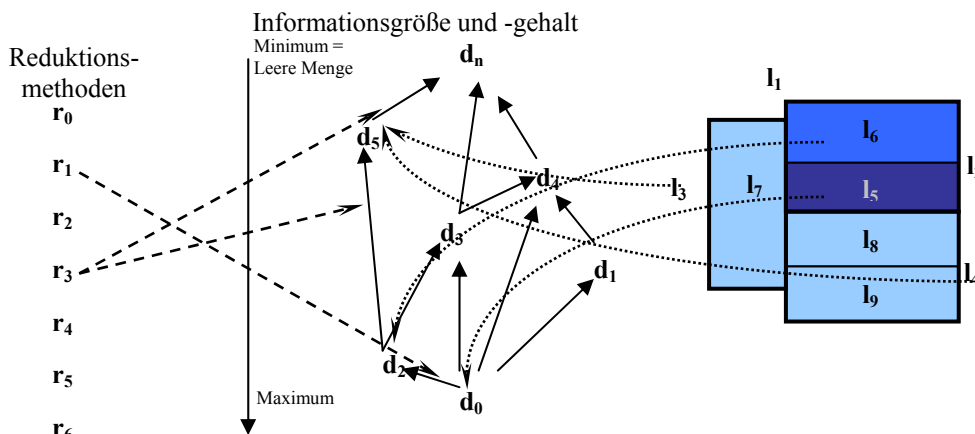


Abb.3: Informations-Reduktions-Hierarchie, Reduktionsmethoden und fokussierte graduelle Informationsreduktion

$\mathbf{d} = \mathbf{i}_1 \cup \dots \cup \mathbf{i}_n$, wobei \mathbf{i}_j weniger stark reduziert ist als \mathbf{i}_{j+1} (entsprechend den \mathbf{d}_i in der Reduktionshierarchie). Für das in Abb.3 dargestellte Beispiel heißt dies, dass sich das Gesamtergebnis aus der Vereinigung folgender Informationen zusammensetzt:

- den unreduzierten, detailgetreuen Informationen \mathbf{i}_5 (entsprechend \mathbf{d}_0) des Ortsbereiches \mathbf{I}_5 , in dem sich der Nutzer aufhält,
- den Informationen \mathbf{i}_6 des benachbarten Ortsbereiches \mathbf{I}_6 aus demselben Ortsbereich der Hierarchie \mathbf{I}_2 , reduziert entsprechend \mathbf{d}_2 und
- den Informationen \mathbf{i}_7 bis \mathbf{i}_9 der weiteren Ortsbereiche \mathbf{I}_3 und \mathbf{I}_4 , reduziert entsprechend \mathbf{d}_5 .

$$\mathbf{d} = \mathbf{d}_0(\mathbf{i}_5) + \mathbf{d}_2(\mathbf{i}_6) + \mathbf{d}_5(\mathbf{i}_3, \mathbf{i}_4)$$

Es gibt im Beispiel keinen Bereich, für den Informationen reduziert nach d_n geliefert werden, d.h. keine leere Information.

3 Zusammenfassung

Die Ortsabhängigkeiten mobiler Anwendungen lassen sich zugunsten der Anwender zunutze machen, indem eine ortsabhängige Fokussierung der Daten vorgenommen wird und die Datenreduktion graduelle Abstufungen erhält. Je weiter etwas entfernt ist, desto mehr werden die Informationen dazu aufgrund des vermutlich geringeren Interesses des Nutzers reduziert. Dies bringt zudem notwendige Ressourcensparnisse. Den einfachen Fall angenommen, dass die 3 Fokusbereiche jeweils die gleiche Datenmenge enthalten und die Reduktion die Information der Größe nach jeweils halbieren würde, betrüge diese beispielhaft dargestellte Reduktion insgesamt 5/12 der Datenmenge.

Günstig auf das Verfahren wirkt sich aus, wenn eine nicht zu große Anzahl Knoten auf jeder (Orts-)Hierarchiestufe ist, um abhängig von der Reduktionsmethodenzuordnung eine bessere Informationszuordnung zu gewährleisten.

Eine Verbesserung der Adaption bietet neben Abstandsfunktionen für lokale Abstände ebenso wie für Abstände zwischen Informationen die Bewertung der Bewegungsrichtung des Nutzers als Indikator seiner zukünftigen Interessen bzw. der prädiktiven Beachtung der Fokusverschiebungen.

Mittels empirischer Untersuchungen ist zudem festzustellen, ob es für einen Anwender zu nachvollziehbaren und brauchbaren Ergebnissen führt, die Informationen der Fokusbereiche entsprechend der Teilergebnisse aus unterschiedlichen Reduktionsketten zu entnehmen. Die Vermutung geht dahin, dass jeweils dieselbe Reduktionskette zu benutzen ist. Hier werden sich Überlegungen anschließen, welche Reduktionsmethoden jeweils miteinander kompatibel sind.

Literaturverzeichnis

- [DRR97] Barbará, D.; DuMouchel, W.; Faloutsos, C.; u.a.: The New Jersey Data Reduction Report. Bulletin of the Technical Committee on Data Engineering, 4(20), S.3-45, 1997.
- [Goll02] Gollmick, C: Using Replication Proxy Servers for Scalable Mobile Database Access. In Proceedings of EDBT Ph.D. Workshop, Prag, 2002.
- [HeLu98] Heuer, A.; Lubinski, A.: Data Reduction - an Adaptation Technique for Mobile Environments. Proc. of the IMC'98, Rostock, 1998.
- [LuHe00] Lubinski, A.; Heuer, A.: Configured Replication for Mobile Applications. Proceedings of the Baltic DB&IS'2000, Vilnius, 2000.
- [LeHö02] Leich, T.; Höpfner, H.: Konzeption eines Anfragesystems für leichtgewichtige, erweiterbare DBMS, In Proc. Workshop Mobile DB und Informationssysteme - Datenbanktechnologie überall und jederzeit, Magdeburg, 2002.
- [Pito96] Pitoura, E.: A Replication Schema to Support Weak Connectivity in Mobile Systems. Int. Conference on DEXA, Wien, 1996.