

# Die Selektivitätsabschätzung im Rahmen der Anfrageoptimierung im HE<sub>A</sub>D-System

Guntram Flach<sup>†</sup>

Universität Rostock  
Fachbereich Informatik  
18051 Rostock

e-mail: gflach@informatik.uni-rostock.de

## 1. Einführung

### 1.1. Motivation

Zur effizienten, unter zeitkritischen Gesichtspunkten durchzuführenden Anfrageverarbeitung ist es notwendig, unter Berücksichtigung geeigneter Zielfunktionen eine Bewertung gegebener äquivalenter Anfragegraphen durch ein Kostenmodell [FILM94, Flac93] vorzunehmen.

Da im HE<sub>A</sub>D-System eine statische Anfrageoptimierung eingesetzt wird, ist die Größe der Zwischenergebnisdatenmengen zur Laufzeit nicht bekannt, so daß eine Abschätzung auf der Basis von Datenbankstatistiken erfolgen muß. Um eine Kostenbewertung vornehmen zu können, müssen daher entsprechende **Statistiken** vorliegen, insbesondere hinsichtlich der Kardinalität (Tupelanzahl) der einzelnen Relationen, Fragmente und Replikate, der Größe einzelner Attribute und Tupel sowie bezüglich der Häufigkeitsverteilung von Attributwerten. Aus diesen Angaben wird bei der Optimierung dann für die einzelnen Ausführungspläne die Größe von Zwischenergebnissen ermittelt, um wiederum die Kosten für nachfolgend auszuführende Operationen abzuschätzen.

Es ist ein **Selektivitätsfaktor**  $SF^\ddagger$  ( $0 \leq SF \leq 1$ ) zu ermitteln, der angibt, welcher Anteil der Relation sich für eine bestimmte Bedingung qualifiziert. Es sei  $card(R)$  die Kardinalität von Relation  $R$  und  $SF(\alpha)$  der Selektivitätsfaktor für Prädikat  $\alpha$  (auf  $R$ ) mit  $SF(\alpha) = card(\sigma_\alpha(R)) = SF(\alpha) * card(R)$ .

### 1.2. Anfrageoptimierung im System HE<sub>A</sub>D

Das Projekt HE<sub>A</sub>D (**H**eterogeneous **E**xtensible and **D**istributed Database Management System) [FILM94] ist der Prototyp eines heterogenen und verteilten Datenbankverwaltungssystems. Ziel der Entwicklungsarbeiten sind Untersuchungen zur parallelen Anfragebearbeitung in heterogen verteilten Datenbanksystemen. Nachfolgend sind wesentliche Charakteristiken aufgeführt:

- Parallelisierung von komplexen Anfragen.
- Neben der horizontalen Parallelität wird die Parallelisierung durch Pipelining und Partitionierung untersucht.

---

<sup>†</sup> Die Entwicklungsarbeit wurde unterstützt durch die DFG, Kennzeichen: Me 1346/1-2

<sup>‡</sup> Wir sprechen von einer "hohen Selektivität", falls der Selektivitätsfaktor einen kleinen Wert annimmt (z.B.  $< 0.01$ ), also nur ein kleiner Teil der Relation sich qualifiziert

- Steuerung des Grades der Multiprogrammierung und Platzierung freier Operatoren durch Methoden der Lastverteilung und -balancierung.
- Integration der aktuellen *Netzbelastung* sowie der aktuellen *Prozessorauslastung* in das Kostenmodell sowie Berücksichtigung von Hauptspeicherrestriktionen bei der Join-Folgeoptimierung.

Optimierungsziel im Rahmen der Anfragebearbeitung ist die Antwortzeitverkürzung.

## 2. Selektivitätsabschätzung

### 2.1. Gleichverteilung der Attributwerte

Im HEAD-System werden zur Abschätzung der Zwischenergebnisdatenmengen nach Ausführung einer relationalen Operation Berechnungsvorschriften in Anlehnung an [CePe85] eingesetzt. Die Grundlage für die Gültigkeit dieser Berechnungsvorschriften ist die Annahme, daß die Werte aller Attribute einer Gleichverteilung unterliegen.

Berechnungsvorschriften oder Näherungsformeln für die Algebraoperation *Selektion* der Form: *Attr rel const*, mit  $rel \in [ <, \leq, >, \geq ]$  existieren in [CePe85] nicht. Für eine derartige Abschätzung sind Kenntnisse über die Wertebereiche der einzelnen Attribute notwendig.

Nachfolgend wird ein Berechnungsansatz<sup>†</sup> aufgezeigt, bei dem die Metadaten um die jeweiligen Grenzwerte der Wertebereiche erweitert werden:

Prädikat	Selektivitätsfaktor SF
$Attr = value$	$\frac{1}{val(A[R])}$
$Attr < value$	$\frac{(value - min)}{(max - min)}$
$Attr > value$	$\frac{(max - value)}{(max - min)}$
$Attr_1 = Attr_2$	$\frac{1}{(max(card(Attr_1), card(Attr_2)))}$
$Attr \text{ BETWEEN } value_1 \text{ AND } value_2$	$\frac{(value_2 - value_1)}{(max - min)}$
$Attr \text{ IN } (list \text{ of } values)$	$\frac{(number \text{ of } list \text{ items})}{SF(Attr = value)}$
$Attr \text{ IN } subquery$	$\frac{card(result \text{ of } subquery)}{(product \text{ of } the \text{ cardinalities})}$
$(\alpha) \text{ OR } (\beta)$	$SF(\alpha) + SF(\beta) - SF(\alpha) * SF(\beta)$
$(\alpha) \text{ AND } (\beta)$	$SF(\alpha) * SF(\beta)$
$\text{NOT } (\alpha)$	$1 - SF(\alpha)$

**Tab. 1:** Selektivitätsfaktoren

Problematisch bei der skizzierten Vorgehensweise ist, daß die erwähnten Annahmen (Gleichverteilung von Attributwerten, Unabhängigkeit verschiedener Attribute) oft nicht gegeben sind. Folglich können die darauf basierenden Schätzungen starken Fehlern unterworfen sein, so daß es ggf. zur Auswahl ungünstiger Pläne kommt.

<sup>†</sup> gilt unter der Annahme der Unabhängigkeit der Werte verschiedener Attribute

## 2.2. Histogrammverfahren

In der weiteren Entwicklungsarbeit wird im HEAD-System zur Verbesserung der Abschätzungsergebnisse eine Statistik-Komponente eingesetzt, mit deren Hilfe unter Verwendung von Histogrammen Aussagen über die Verteilung der Attributwerte einer Relation möglich sind. Die generelle Vorgehensweise wird nachfolgend beschrieben:

- Zerlegung des Attributwertbereiches in endlich viele aneinander grenzende Intervalle  $\Delta_1, \dots, \Delta_k$
- Ermittlung, wieviele Attributwerte innerhalb der Intervalle ( $i \leq j \leq k$ ) liegen.
- Ableitung der Verteilung der Attributwerte aus dem entstandenen Stufenbild (Histogramm)

Verschiedene Variationen hinsichtlich der Intervallaufteilung, gleiche Intervallbreite, gleiche Attributwertanzahl (Höhe), variable Intervallbreite, können zur Selektivitätsabschätzung eingesetzt werden [Lync88, Ioan93, PiCo84].

### 2.2.1. Histogrammintervall mit gleicher Attributwertanzahl

Es hat sich herausgestellt [PiCo84], daß der maximale Abschätzungsfehler durch die Höhe jedes Histogrammbereiches und nicht durch die Breite bestimmt wird, d.h., daß jedes Histogramm-Bucket die gleiche Attributwertanzahl umfassen sollte. Es ist demnach zunächst notwendig, die Relation bzgl. des jeweiligen Attributes zu sortieren, um gleich hohe Histogramme zu generieren. Der Selektivitäts-Abschätzungsfehler kann durch eine Erhöhung der Histogrammbucketanzahl weiter reduziert werden.

Wir betrachten ein Attribut  $A$  der Relation  $R$  ( $T$  als Tupelanzahl der Relation  $R$ ). Die Intervallschritte  $STEP(0), \dots, STEP(S)$  sind solche Werte aus der Tupelmengemenge mit  $A < STEP(I)$ , die kleiner oder gleich  $I/S$  sind (Beispiel  $S = 10$ , siehe Abb. 1).

$STEP(0)$  ist der minimale Wert der Attributmenge  $A$  und  $STEP(10)$  ist der maximale Attributwert.  $STEP(3)$  bedeutet, daß für 30% der Attributwerte der gegebenen Tupelmengemenge  $A < STEP(3)$  gilt.

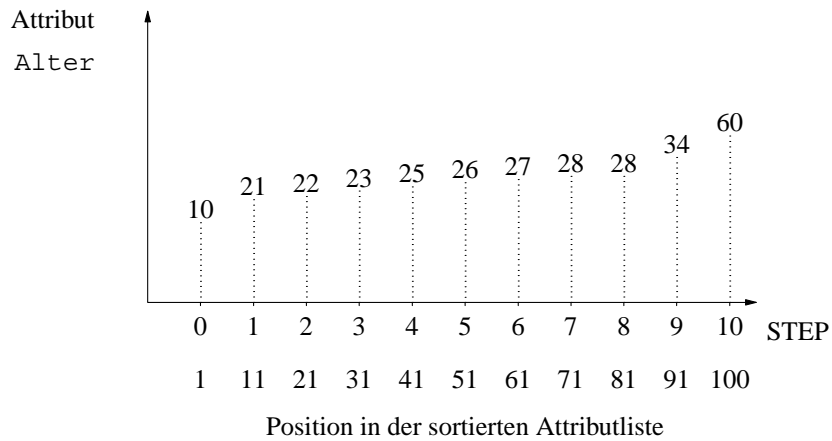
Nachfolgend wird die Herleitung der Intervallschritte beschrieben [PiCo84]:

1. Sortierung der Tupel der Relation nach dem betreffenden Attribut (where-Klausel) in aufsteigender Reihenfolge.
2. Auswahl der Anzahl der Verteilungsschritte  $S$ , jeweils in Abhängigkeit von der betrachteten Genauigkeit. Festlegung von  $S + 1$  Positionen (einschließlich der ersten und letzten) in der sortierten Liste der Attributwerte, so daß die Anzahl der Attributwerte zwischen zwei beliebigen aufeinanderfolgenden Positionen gleich sind. Diese Positionen sind  $1, 1 + N, 1 + 2N, \dots$ , mit  $N = \frac{(T - 1)}{S}$ .
3. Die Attributwerte, die sich an den Positionen der Verteilungsschritte  $STEP(0), \dots, STEP(S)$  befinden, bilden die jeweiligen Ordinatenwerte (siehe Abb. 1)

Ausgehend von diesem Konstruktionsprozeß, ist erkennbar, daß für jeden Schritt  $I$ , die Anzahl der Tupel mit  $A < STEP(I) \leq I * N$  gilt.

#### Beispiel:

In der Abbildung 1 wird die Attributwertverteilung bzgl. eines Attributes `Alter` einer Relation `Angestellte` mit 100 Tupeleinträgen betrachtet.



**Abb. 1:** Histogramm

Das heißt, für die Selektivität  $SF(\text{Alter} < 29)$  ergibt sich  $0.80 < SF(\text{Alter} < 29) \leq 0.90$ . Als Abschätzungsergebnis kann damit der Mittelpunkt des Intervall 0.85 gewählt werden. Der maximale Abschätzungsfehler beträgt 0.05.

Die konkreten Berechnungsvorschriften für die Bestimmung von  $SF(A < X)$ ,  $SF(A = X)$ <sup>†</sup> können auf der Grundlage einer analytischen Herleitung in Abhängigkeit von der Position von  $X$  innerhalb der einzelnen Intervallschritte bestimmt werden.

Eine Anwendung dieses Konzeptes gleich hoher Histogramme (gleiche Anzahl von Attributwerten innerhalb eines Buckets) für mehrdimensionale Anfragen ist über die Generierung mehrdimensionaler Histogramme ebenso möglich [Mura88].

Als nachteilig bei dem bisher beschriebenen Verfahren erweisen sich die folgenden Punkte:

- Overhead durch Erstellung der jeweiligen Statistiken zu den einzelnen Attributen einer Relation
- Periodisches Einlesen und Sortieren der Attributwerte des gesamten Datenbestandes ist sehr zeitaufwendig.
- Bei hoher Änderungshäufigkeit der Daten kommt es zu einer Verfälschung dieser Statistiken.

### 2.2.2. Einschränkung des Berechnungsaufwandes durch 'Random Sampling'

Im nachfolgenden wird basierend auf der Kolmogorov-Statistik [Gibb76] davon ausgegangen, daß eine Beschränkung auf einen Teil der betrachteten Tupelmengemenge für die Generierung der notwendigen Histogramme möglich ist.

Es sei  $\lambda$  der Anteil der Tupelmengemenge der Relation, der einer bestimmten Bedingung (Selektionsprädikat) entspricht.  $\omega$  sei der Tupelanteil der ausgewählten Teiltupelmengemenge, der ebenfalls innerhalb des Anfragebereiches liegt. Nach den Kolmogorov-Statistiken gilt  $|\lambda - \omega| \leq d$  mit der Wahrscheinlichkeit  $p$  und  $n$  als Teilmengengröße sowie  $d$  als Genauigkeit.

Bei vorgegebenen  $p$  und  $d$  kann der entsprechende Wert für  $n$  den jeweiligen Tabellen [Gibb76] entnommen werden.

<sup>†</sup> Beschränkung auf diese Terme ist möglich, da  $SF(A \leq X)$ ,  $SF(A > X)$ ,  $SF(A \geq X)$  ableitbar sind.

So kann beispielsweise bei einer Teilmenge<sup>†</sup> mit der Tupelanzahl von 1063, entsprechend der Tabellen [Gibb76] davon ausgegangen werden, daß  $|\lambda - \omega| \leq 0.05$  mit einer Wahrscheinlichkeit  $\geq 0.99$  gilt.

### 3. Ausblick

Wir haben einen kurzen Überblick über die Problematik der Selektivitätsabschätzung relationaler Operatoren gegeben und Lösungsmöglichkeiten im Rahmen des HE<sub>A</sub>D-Projektes aufgezeigt. Die weiteren Forschungsarbeiten werden sich auf folgende Punkte konzentrieren:

- Einbeziehung von Adaptive-Sampling-Mechanismen in die Selektivitätsabschätzung
- Bereitstellung von systeminternen Mechanismen und Verfahren zur adaptiven Anpassung der Kostenparameter an Systemveränderungen
- Validierung der Kostenabschätzungsmechanismen

### Literaturverzeichnis

CePe85.

Ceri, S. and Pelagatti, G., "Distributed Databases: Principles and Systems," *McGraw-Hill*, New York (1985).

Flac93.

Flach, G., "Konzeption eines Kostenmodells für ein verteiltes Datenbanksystem," *Diplomarbeit*, Universität Rostock, Fachbereich Informatik, Rostock (Mai 1993).

FILM94.

Flach, G., Langer, U., and Meyer, H., "Kostenbasierte Anfrageoptimierung in HE<sub>A</sub>D," *Rostocker Informatik-Berichte*, 15, Universität Rostock, Fachbereich Informatik (1994).

Gibb76.

Gibbons, J.D., "Nonparametric Methods for Quantitative Analysis" in *Holt, Rinehart, and Winston*, New York (1976).

Ioan93.

Ioannidis, Y.E., "Universality of Serial Histograms," *Proc. of the 19th VLDB Conf.*, Dublin (1993).

Lync88.

Lynch, C.A., "Selectivity Estimation and Query Optimization in Large Databases with highly skewed Distributions of Column Values," *Proc. of the 14th VLDB Conf.*, pp. 240-251, Los Angeles (1988).

Mura88.

Muralikrishna, M., "Optimization of Multiple-Disjunct Queries in a Relational Database System," TR 750, Dept. of CS, University of Wisconsin-Madison (February 1988).

PiCo84.

Piatetsky-Shapiro, G. and Connell, C., "Accurate Estimation of the Number of Tuples Satisfying a Condition," *ACM*, Dept. of CS, New York University (1984).

---

<sup>†</sup> Die Teilmengengröße verhält sich umgekehrt proportional zum Quadrat der Genauigkeit