Auto BDC

Steve Göring

intro

mcluste

probles

relate

mdc

conclusio

ideas

conecpt

precondition

abstraction

transformati

cvcles

full auto

ovamnlo

CALCITO

.....

tpch

bechmarks

Automatisierter Entwurf eines Schemas für Bitwise Dimensional Clustering

-Verteidigung-

Steve Göring

16. Februar 2012 $V0.\overline{9}$

Inhalt

1 intro

2 related

3 ideas

4 tests

intro

mclust

proble

relate

mdc

conclusio

ideas

preconditio

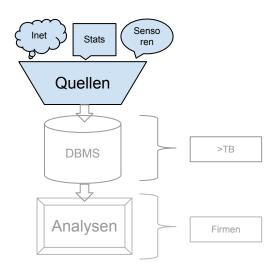
u D J L I U C L I O I

-,----

tpch

. bechmarks

Einleitung (0)



intro

mclust

proble

relate

mdc

conclusio

ideas

conecni

preconditio

abstraction

-,----

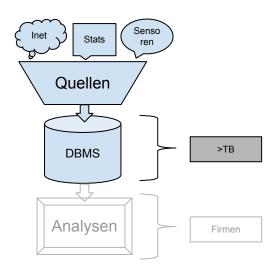
extens

test

tpch

hechmark

Einleitung (1)



intro

mclust

proble

relate

mdc

conclusio

ideas

u D J L I U C L I O I I

transformat

full auto

ovamala

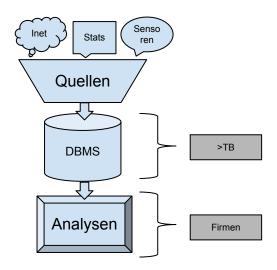
ovtoncio

tect

tpch

bechmark

Einleitung (0)



intro

mcluste

proble

relate

mdc

conclusio

idea

conecpt

preconditio

abstraction

transforma

cycles

full auto

example

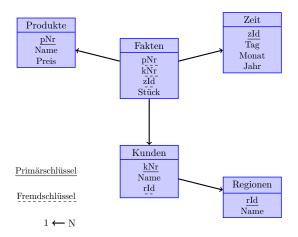
extens

tosts

toch

bechmarks

Einleitung - (Online Versand)



$$\begin{aligned} |\textit{Kunden}| &= |\textit{Fakten}| = |\textit{Zeit}| = 1000 \\ &\rightarrow \mathsf{Kunden} \times \mathsf{Fakten} \times \mathsf{Zeit} = 1 \; \mathsf{Mio}. \end{aligned}$$

Multidimensionales Clustern

mcluster

Daten eingeteilt in:

▷ Dimensionen:

beschreibende Daten

▶ Fakten: Kern/ Ergebniszahlen

Multidimensionales Clustern

intr

mcluster

proble

relate

mdc

conclusio

idea

precondition

abstraction

transformation

cycles

full auto

extension

tests

tpch

bechmarks

Daten eingeteilt in:

- Dimensionen:
- beschreibende Daten

Kern/ Ergebniszahlen

- geclustert durch Dimensionen
- Daten mit ähnlichen Dims dicht beieinander
- Clusterinfos zur Datenreduktion

Bitwise Dimensional Clustering

intr

mclust bdc

relate

other

conclus

ideas

conecpt

abstraction.

transformati

cycles

full auto

example

extensi

tests

tpch

Bitwise Dimensional Clustering (BDC):

- multidimensionales Speichersystem für Columnstores (Vektorwise)
- ightharpoonup Clusterungsproblem ightarrow Ordnungsproblem

Grundidee von BDC:

- \triangleright Dimensionen \rightarrow Integerzahlen
- ▶ Tabellen geclustert nach Dimensionsauswahl

BDC-Schema

intro

bdc

probler

mdc

otners

COITCIUSI

idea

precondition

abstraction

transformatio

cycles

full auto

example

extens

oechmarks

Festlegung von Dimensionen:

mclust bdc

mdc

conclusi

idea

conecpt

abstraction

.....

transionnatio

cycles

full auto

example

extens

tests

bechmark

Festlegung von Dimensionen:

- ▶ Tabellenname
- > Dimensionsspalte
- ▷ cbits

bdc

proble

relati

mdc

conclusi

ideas

precondi

abstraction

transformatio

cvcles

full auto

evamnle

. . . .

bechmarks

Festlegung von Dimensionen:

- ▶ Tabellenname
- Dimensionsspalte
- ▷ cbits

BDC-geclusterte Tabelle:

mclust bdc

. .

mdc

conclusi

idea

conecpt

abstraction

a Contraction

cycles

full outo

full auto

extensi

+nch

bechmarks

Festlegung von Dimensionen:

- ▶ Tabellenname
- Dimensionsspalte
- ▷ cbits

BDC-geclusterte Tabelle:

- ▷ Pfad zur jeweiligen Dimension
- Bitmaske für jede Dimension
- ▷ iobits

BDC- Ablauf

mcluster

bdc

related

others conclusio

idea

conecpt

abstraction

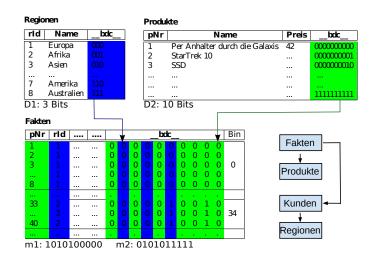
transformati cvcles

full auto example

extension

tests

bechmarks



iobits(Fakten) = 10



Steve Goring

intro

mcluste

problem

relate

mdc

conclusio

1444

idea

conecpt

preconditio

abstraction

transformati

cycles

full auto

example

tests

tpch

oechmarks

Problemstellung

> komplexere Schemas \rightarrow Dimensionsauswahl schwerer

Auto BDC

Steve Göring

intro

mciuste

problem

relate

mdc

conclusio

conclusio

idea

conecpt

1.0

abstraction

transformatio

cycles

full auto

example

toch

bechmark

Problemstellung

- ightharpoonup komplexere Schemas ightarrow Dimensionsauswahl schwerer
- $\, \triangleright \, \, \text{viele Faktoren beinflussen Clusterung} \,$

Auto BDC

Steve Göring

intro

hde

problem

relat

others

conclusi

idea

conecpt

abstraction

transionnatio

cycles

full auto

example

....

bechmark

Problemstellung

- ⇒ automatisches Erzeugen des Schemas

Verwandte Themen - MDC

intro

mcluste bdc

relate

mdc others

conclusi

idea

precondition

abstraction

transformati

cycles

full auto

example

toch

bechmark

MDC in DB2 ¹

- ⊳ Lightstone, Bhattacharjee ²
- > automatisiertes Verfahren
- > als Ratgeber implementiert

für Dimensionen Analyse:

- ▷ Daten
- ▷ Anfragen (GROUP BY, ORDER BY, WHERE)

¹"Database Partitioning, Table Partitioning, and MDC for DB2 9", 2009

²"Automated Design of Multidimensional Clustering Tables for Relational Databases". 2004

Auto BDC

Steve Göring

intro

mcluste

proble

relate

mdc

others

conciu

idea:

conecpt

.

transionnati

cycles

full auto

example

extens

tests

tpch

bechmarks

Verwandte Themen - andere Verfahren

Fazit

mclust bdc

and the second

Ciaco

others

conclusion

idea

conecpt

preconditio

abstraction

transformation

cycles

full auto

example

....

hechmark

 $\, \triangleright \, \, \mathsf{Datenanalyse} \,$

mclust

. . .

mdc

conclusion

idea:

conecpi

abstraction

transformat

cuclos

full auto

evamnle

.

. . . .

- Datenanalyse :)
- ▶ Anfrageanalyse

IIILIC

mclust hdc

and a second

mdc

conclusion

idea

conocn

precondition

abstraction

transformati

.

cycles

full auto

example

LCSLS

bechmark

- ▷ Datenanalyse :)
- ▷ Anfrageanalyse :(

Fazit

IIILIY

mcluste

relate

mdc

conclusion

idea

conecpt

.

.

cycles

full auto

example

extensio

. .

tpch

bechmarks

- Datenanalyse :)
- ▶ Anfrageanalyse :(

Ziel:

- ▷ innerhalb von Vectorwise
- $\, \triangleright \, \, \mathsf{auf} \, \, \mathsf{BDC} \, \, \mathsf{abgestimmt} \, \,$

Steve dom

intro

bdc

mdc

. .

conclusio

idea

conecpt

precondition

abstraction

transformation

cuclos

ovamnlo

extens

toch

bechmark

Lösungsidee

Dimensionstabellen:

intro

mcluste

.....

mdc

conclusi

ideas

conecpt

precondition

abstraction

transformat

.

.

.

....

bechmark:

Lösungsidee

- ▷ Dimensionstabellen:
 - Tabellen, die keine Fremdschlüsselbeziehungen zu anderen besitzen

otere com

intro

mclust

100

mdc

conclus

ideas

.

conecpt

.

transiormati

cycles

full auto

example

tests

bechmark

Lösungsidee

▷ Dimensionstabellen:

- Tabellen, die keine Fremdschlüsselbeziehungen zu anderen besitzen
- benutzen zur Clusterung eigene Dimension

conecpt

Lösungsidee

- Dimensionstabellen:
 - ⋄ Tabellen, die keine Fremdschlüsselbeziehungen zu anderen besitzen
 - benutzen zur Clusterung eigene Dimension
- ▷ andere Tabellen:

Lösungsidee

intr

mclust

proble

....

mdc

conclusi

idoo

conecpt

precondition

abstraction

transformatio

.

cycles

full auto

example

extens

test

. .

▷ Dimensionstabellen:

- ♦ Tabellen, die keine Fremdschlüsselbeziehungen zu anderen besitzen
- benutzen zur Clusterung eigene Dimension
- □ andere Tabellen:
 - benutzen Informationen der Nachbartabellen zum clustern

Steve domi

intr

hdc

proble

relat

mdc

otners

idea

conecpt

precondition

abstraction

transformati

cycles

full auto

evamnle

evtensio

tpch

bechmark

Lösungsidee

- ▷ Dimensionstabellen:
 - Tabellen, die keine Fremdschlüsselbeziehungen zu anderen besitzen
 - benutzen zur Clusterung eigene Dimension
- - benutzen Informationen der Nachbartabellen zum clustern
- ▷ Abarbeitung von "Außen nach Innen"

Lösungsidee

intr

mclust hdc

proble

relati

mdc

conclus

idead

conecpt

precondition

abstraction

transformati

cycles

full auto

example

extensi

tests

tpch

hochmark

- ▷ Dimensionstabellen:
 - Tabellen, die keine Fremdschlüsselbeziehungen zu anderen besitzen
 - benutzen zur Clusterung eigene Dimension
- - benutzen Informationen der Nachbartabellen zum clustern
- ▷ Abarbeitung von "Außen nach Innen"
- ▷ eventuell vorher Transformation nötig

Voraussetzungen

intro

mcluste hdc

. . .

mdc

conclusio

CONCIUSIO

ideas

conecht

precondition

abstraction

transformati

full auto

example

extensio

tests

tpch

bechmark:

▷ Informationen über Schema:

Voraussetzungen

intro

mcluste

100

mdc

otners

conclusio

idea

conecpt

precondition

.

transionnat

cycles

full auto

Coumpi

. .

tpch

bechmark:

- ▶ Informationen über Schema:
 - ♦ Tabellennamen

Voraussetzungen

intro

mcluste

p.

relate

mdc

conclusio

idea

precondition

•

abstraction

transforma

cycles

full auto

extens

....

bechmark

- ▶ Informationen über Schema:
 - ⋄ Tabellennamen
 - ⋄ Fremdschlüssel / Primärschlüssel Beziehungen

Voraussetzungen

intro

mclust

probi

relate

mdc

conclusio

. .

idea

precondition

abetraction

. .

.

full auto

example

.....

tpch

bechmark

- ▶ Informationen über Schema:
 - ⋄ Tabellennamen
 - ⋄ Fremdschlüssel / Primärschlüssel Beziehungen
- ▷ Statistische Informationen: (Histogramm/ Anfragen)

mure

hde

proble

relate

mdc

conclusi

luea

precondition

u DJET UCETOT

full auto

....

Lests

bechmark

Voraussetzungen

- ▶ Informationen über Schema:
 - ⋄ Tabellennamen
 - ⋄ Fremdschlüssel / Primärschlüssel Beziehungen
- ▷ Statistische Informationen: (Histogramm/ Anfragen)
 - ♦ Min/Max/Count

mclust

proble

relate

mdc

conclusi

luea

precondition

. .

.

Tun uuc

example

CALCUIS

tocto

tpch

bechmark

Voraussetzungen

- ▶ Informationen über Schema:
 - ⋄ Tabellennamen
 - Fremdschlüssel / Primärschlüssel Beziehungen
- Statistische Informationen: (Histogramm/ Anfragen)
 - Min/Max/Count
 - ♦ Blockgröße

$\mathsf{Schema} \to \mathsf{Graph}$

intro

mcluste

mdc

conclusio

ideas

precondition

abstraction

transformat

cycles

full auto

ovamnlo

extensio

extensio

....

hechmark

⊳ Schema

$\mathsf{Schema} \to \mathsf{Graph}$

intro

mcluste

relate

others

conclusio

ideas

preconditio

abstraction

transformatic

.

full auto

Tull auto

extension

CACCITOTO

....

hochmark

 \triangleright Schema \rightarrow

intro

mclust

probi

relate

mdc

conclusio

ideas

conecp

precondition

abstraction

transformati

cycles

full auto

example

extens

tpch

. bechmarks \triangleright Schema \rightarrow erweiterter gerichteter Multigraph $G = (V, E, \Delta)$, mit:

Schema \rightarrow Graph

intr

mclust

proble

relate

mdc

conclusi

idea

procondit

abstraction

transforma

cycles

full auto

ovamalo

extensio

extensio

toch

hochmarl

▷ Schema \rightarrow erweiterter gerichteter Multigraph $G = (V, E, \Delta)$, mit: $\Diamond V = \{Tabellennamen\}$

intro

mclust

proble

relate

mdc

conclusi

idea

precondit

abstraction

.

cyclos

cycles

full auto

extensio

tpch

bechmarks

Schema → erweiterter gerichteter Multigraph

$$G = (V, E, \Delta)$$
, mit:

$$\diamond V = \{Tabellennamen\}$$

$$\diamond$$
 $E = \{(x, y) | x \text{ hat FK zum PK von } y \}$!Multimenge!

intro

mclust

probl

relate

mdc

conclusi

idea

conecpt

abstraction

transformat

full auto

ovamnlo

EXTERIS

+nch

bechmarks

▷ Schema \rightarrow erweiterter gerichteter Multigraph $G = (V, E, \Delta)$, mit:

 $\wedge V = \int T_2 h_0 ||f_2||_2$

- $\diamond V = \{Tabellennamen\}$
- $\diamond E = \{(x,y)|x \text{ hat FK zum PK von } y\} \text{ !Multimenge!}$
- ♦ Δ Funktion: Pfadspeicherungsfunktion $\Delta : E \mapsto Pot(Pfade)$

<ロ > ← □

intro

mclust

problen

relate

mdc

conclusio

ideas

conecpt

abstraction

austraction

transforma

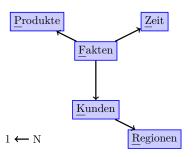
cycles

Tull auto

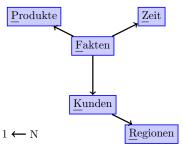
example

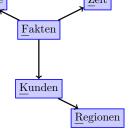
toch

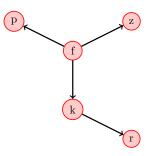
bechmarks



abstraction







$$\triangleright V = \{p, f, z, k, r\}$$

intro

bdc

related

mdc

conclusio

idea:

conecpt

abstraction

. .

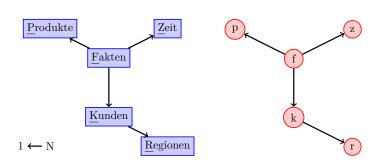
-, -.--

coumpic

tests

tpch

bechmark



$$▷ V = \{p, f, z, k, r\}$$

$$▷ E = \{(f, p), (f, z), (f, k), (k, r)\}$$

intro

mcluste bdc

. .

mdc

conclusio

ideas

preconditi

abstraction

transformat

cycles

full sute

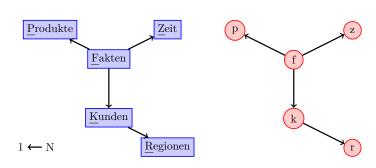
example

extension

tests

tpch

bechmark



$$\triangleright V = \{p, f, z, k, r\}$$

$$\triangleright E = \{(f,p), (f,z), (f,k), (k,r)\}$$

$$\triangleright \Delta(x) = x, \forall x \in E$$

transformation

intro

mcluste

relate

mdc

conclusi

ideas

conecut

. . .

transformation

cycles

full auto

toch

tpch

bechmarks

Transformationskriterien

intro

mcluste

relate

mdc

conclusio

ideas

conecpt

precondition

abstraction

transformation

full auto

evamnle

extensi

tpch

bechmarks

Transformationskriterien

 $\, \triangleright \, \, \mathsf{zusammenh\"{a}ngend} \, \to \, \mathsf{mehrmaliges} \, \, \mathsf{Anwenden} \,$

mcluste

mdc

others

COTTCIUS

ideas

conecpt

abstraction

.

transformation

cyclos

. ..

evamnle

extens

tpch

bechmark:

- $\, \triangleright \, \, \mathsf{zusammenh\"{a}ngend} \, \to \, \mathsf{mehrmaliges} \, \, \mathsf{Anwenden} \,$
- ⊳ Eigenkanten

intro

mcluste

relate

mdc

conclusio

ideas

conecpt

.

abstraction

transformation

cycles

Tull auto

example

tests

bechmark

- hd zusammenhängend ightarrow mehrmaliges Anwenden
- ightarrow Eigenkanten ightarrow entfallen

intro

mcluste

relate

mdc

conclus

ideas

conecpt

transformation

transformatio

cvcles

full auto

example

extens

tests

tpch

- ightarrow zusammenhängend ightarrow mehrmaliges Anwenden
- ${\,\vartriangleright\,} \mathsf{Eigenkanten} \to \mathsf{entfallen}$

intro

mcluste

relate

mdc

conclus

ideas

conecpt

.

transformation

full auto

ovamnlo

toch

bechmarks

- \triangleright zusammenhängend \rightarrow mehrmaliges Anwenden
- ${\,\vartriangleright\,} \mathsf{Eigenkanten} \to \mathsf{entfallen}$
- ightharpoonup Kreise ightarrow transformiert

intro

mcluste

proble

relate

mdc

conclus

ideas

conecpt

abstraction

transformation

transformati

c 11 .

Tull auto

. . .

bechmark

- ightharpoonup zusammenhängend ightharpoonup mehrmaliges Anwenden
- ${\,\vartriangleright\,} \mathsf{Eigenkanten} \to \mathsf{entfallen}$
- \triangleright Kreise \rightarrow transformiert
- \Rightarrow zyklusfreier, zusammenhängender Multigraph $G=(V,E,\Delta)$

Kreise - (allgemein)

intr

mcluste bdc

.....

mdc

conclusi

idea:

concept

.

abstraction

cycles

full auto

example

extens

toch

oechmark:

▷ Tiefensuche (DFS) um Kreise zu finden

intro

mcluste bdc

100

mdc

. .

COTTCTUST

idea

conecpt

preconarcio

abstraction .

transformat

cycles

full auto

example

tpch

bechmark:

- ▷ Tiefensuche (DFS) um Kreise zu finden
- ▷ für jeden Knoten eines Kreises:

Steve Göring

intro

mcluste

. . . .

relate

mdc

conclusi

idooc

raca.

preconditio

abstraction

transformatio

cvcles

cycles

run uuto

example

tests

bechmark

- ▷ für jeden Knoten eines Kreises:
 - DFS vom Knoten ausführen um neue Pfade zu den Blättern zu finden

intro

mcluste

proble

relate

mdc

conclusi

idea

proconditio

abstraction

transformation

cvcles

full auto

example

extens

tocto

tpch

bechmark

- ▷ für jeden Knoten eines Kreises:
 - DFS vom Knoten ausführen um neue Pfade zu den Blättern zu finden
 - ⋄ für jeden Pfad: füge Abkürzung in *E* ein

intro

mclust

probl

relate

mdc

conclusi

idea

conecpt

abstraction

transformatio

cvcles

full auto

example

CALCUI

tests

bechmar

- ▷ für jeden Knoten eines Kreises:
 - DFS vom Knoten ausführen um neue Pfade zu den Blättern zu finden
 - \diamond für jeden Pfad: füge Abkürzung in E ein
 - \diamond merken des Pfades via Δ -Funktion

intro

mcluste

probl

relate

mdc

conclusi

ideas

conecpt

precondition

abstraction

.

cycles

Tull auto

example

CALCII

Lests

bechmar

- ▷ für jeden Knoten eines Kreises:
 - DFS vom Knoten ausführen um neue Pfade zu den Blättern zu finden
 - \diamond für jeden Pfad: füge Abkürzung in E ein
 - \diamond merken des Pfades via Δ -Funktion

intro

mclust

proble

relate

mdc

conclusi

ideas

conecpt

abstraction

.....

cycles

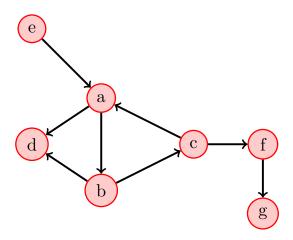
ovtoncio

test

tpch

bechmarks

Kreise - (0)



Steve Göring

intro

mclust

proble

rolato

mdc

conclusi

ideas

conecpt

abstraction

transformatio

cycles

full auto

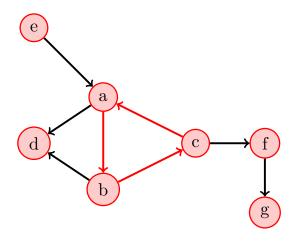
example

test

tpch

bechmarks

Kreise - (1)



Steve Göring

intro

mclust bdc

proble

relate

mdc

conclusio

idea

conecpt

abstraction

transionnati

cycles

example

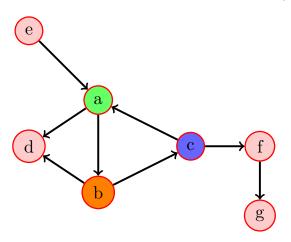
extensio

test

tpch

bechmarks

Kreise - (2)



Tiefensuche: Knoten a

Steve Göring

intro

mclust

problen

relate

mdc

conclusi

ideas

iuea

conecpt

abstraction

transformatio

cycles

full auto

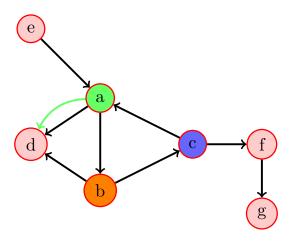
example

test

tpch

bechmarks

Kreise - (2)



Tiefensuche: Knoten a: Pfad: a, d



intro

mclust

probler

relate

mdc

conclusi

idea

conecpt

abstraction

transformatio

cycles

full auto

example

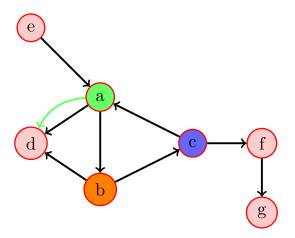
CALCITA

test

tpch

bechmarks

Kreise - (2)



Tiefensuche: Knoten a: Pfad: $a, d \rightarrow$ orginale Kante

Steve Göring

intro

mclust

proble

relate

mdc

conclusi

ideas

conecpt

abstraction

transformati

cycles

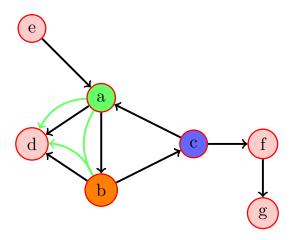
extensio

test

tpch

bechmarks

Kreise - (2)



Tiefensuche: Knoten a: Pfad: a, b, d

Steve Göring

intro

mclust

probler

relate

mdc

conclusio

ideas

conecpt

. . .

cycles

full auto

example

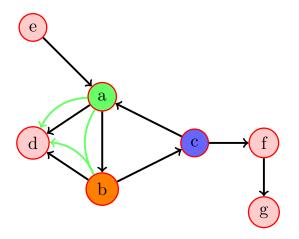
extensio

tests

tpch

bechmarks

Kreise - (2)



Tiefensuche: Knoten a: Pfad: $a, b, d \rightarrow \text{neu}(a, d)$

Steve Göring

intro

mclust

proble

relate

mdc

conclusi

idea

conecpt

precondition

abstraction.

cycles

full auto

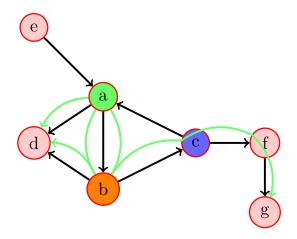
example

.

tpch

bechmarks

Kreise - (2)



Tiefensuche: Knoten a: Pfad: a, b, c, f, g

Steve Göring

intro

mclust

probler

mdc

conclusio

ideas

precond

abstraction

transformati

cycles

evamnle

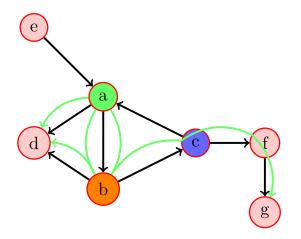
extension

test

tpch

bechmarks

Kreise - (2)



Tiefensuche: Knoten a: Pfad: $a, b, c, f, g \rightarrow \text{neu}(a, g)$

intro

mclust

proble

relate

mdc

conclusi

ideas

conecpt

preconditio

abstraction

transformati

cycles

ovamnla

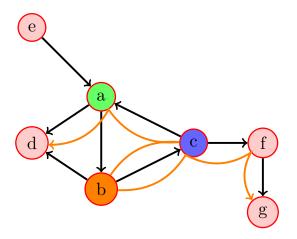
extensio

test

tpch

bechmarks

Kreise - (3)



Tiefensuche: Knoten $b \rightarrow \text{neu } (b,d), (b,g)$

intro

mclust

probler

relate

mdc

conclusi

idea

conec

precondition

abstraction

transformati

cycles

Tull auto

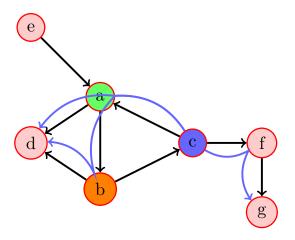
example

tests

tpch

bechmarks

Kreise - (4)



Tiefensuche: Knoten $c o ext{neu}(c,d), (c,d), (c,g)$

Steve Göring

intro

mclust

probler

relate

mdc

conclusio

idea

conecpt

abstraction

cycles

full auto

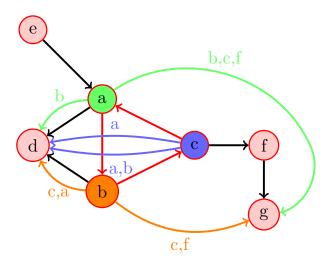
extension

test

tpch

bechmarks

Kreise - (5)



alle neuen Kanten

Auto BDC

Steve Göring

intro

mclust

problen

relate

mdc

conclusio

ideas

conecpt

abstraction.

cycles

full auto

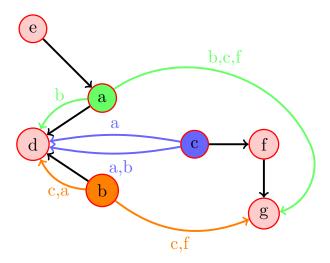
example

test

tpch

bechmarks

Kreise - (6)



ohne Kreiskanten

Auto BDC

Steve Göring

intro

mclust

probler

relate

mdc

conclusio

idea

conecpt

.

abstraction

transformati

cycles

evample

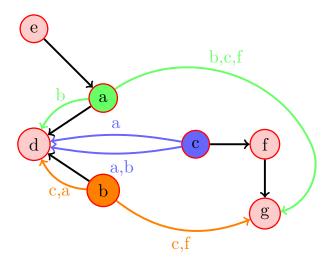
extensio

test

tpch

bochmark

Kreise - (6)



ohne Kreiskanten \Rightarrow zyklusfreier Graph

intro

mclust hdc

proble

. .

mdc

OLIICI S

COTTCTUST

idea:

conecpt

abstraction

cycles

example

evtension

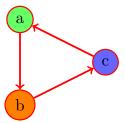
test

tpch

bechmarks

Sonderfall

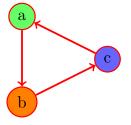
Graph = kompletter Kreis

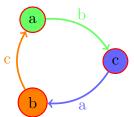


cycles

Sonderfall

Graph = kompletter Kreis





Steve Goring

intro

bdc

problem

relate

mdc

conclusio

idea

conecpt

abstraction

.....

transionnati

cycles

Tull auto

example

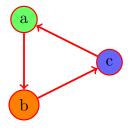
tests

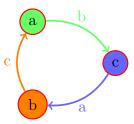
tpch

hochmarks

Sonderfall

Graph = kompletter Kreis





 \Rightarrow spezielle Behandlung

Steve Goring

intro

bdc

problem

relate

mdc

conclusio

CONCIUSIO

luea

conecpt

abstraction

.....

transiormati

cycles

Tull auto

example.

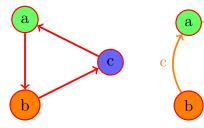
test

tpch

hechmarks

Sonderfall

Graph = kompletter Kreis



 \Rightarrow spezielle Behandlung \rightarrow nicht weiter betrachtet

Grundvariante

ıntr

mcluste bdc

mdc

orners .

ideas

conecpt

preconditio

abstraction

transformation

cycles

full auto

examp

....

hochmark

 $= \mathsf{komplett} \ \mathsf{automatisierte} \ \mathsf{Variante}$

Grundvariante

intr

mcluste bdc

. . .

mdc

conclusio

ideas

conecpt

abstraction

.

transformati

cyclos

full auto

example

exten:

tests

tpch

bechmark:

= komplett automatisierte Variante Ablauf:

Grundvariante

intr

mcluste bdc

mdc

conclusio

conclusi

idea

conecpi

abstraction

.....

full auto

....

......

CALCITY

.

bechmark

= komplett automatisierte Variante Ablauf:

 $\, \triangleright \, \, \mathsf{Schema} \, \to \, \mathsf{Graph} \,$

full auto

Grundvariante

- = komplett automatisierte Variante Ablauf:
 - \triangleright Schema \rightarrow Graph
 - Dimensionen finden /registrieren

Grundvariante

intr

mclust bdc

relate

mdc

conclusi

:dana

luea

abstraction

transformatio

cycles

full auto

example

extens

tests

bechmark

= komplett automatisierte Variante Ablauf:

- ightharpoonup Schema ightarrow Graph
- Dimensionen finden /registrieren
- ▷ BDC Tabellen erzeugen: von außen nach innen

Grundvariante

intr

mclust bdc

relate

mdc

conclusi

:dana

luea

abstraction

transformatio

cycles

full auto

example

extens

tests

bechmark

= komplett automatisierte Variante Ablauf:

- ightharpoonup Schema ightarrow Graph
- Dimensionen finden /registrieren
- ▷ BDC Tabellen erzeugen: von außen nach innen

Grundvariante - (1)

mcluste bdc

. .

mdc

conclusio

idea

conecpt

preconditio

abstraction

transformatio

cvcles

full auto

examp

toch

oechmarks

 $Dimensionen\ finden/registrieren:$

Grundvariante - (1)

intro

mcluste bdc

. . .

mdc

conclusio

ideas

conecpt

.

abstraction

transformation

.

full auto

.......

tpch

bechmarks

Dimensionen finden/registrieren:

▷ Blätter finden

Grundvariante - (1)

intr

mcluste bdc

100

mdc

conclusio

ideas

conecpt

abstraction

transiormatic

cycles

full auto

evtensi

toch

bechmark

Dimensionen finden/registrieren:

- ▷ Blätter finden
- ▷ für jedes Blatt:

Steve Goring

intr

mcluste bdc

relate

mdc

conclusio

idea

conecpt

abstraction

transionnatio

full auto

run uuc

extensi

toch

bechmark

Grundvariante - (1)

Dimensionen finden/registrieren:

- ▷ Blätter finden
- ▷ für jedes Blatt:
 - \diamond Dimension registrieren

Grundvariante - (1)

intr

mclust bdc

relate

mdc

conclusio

ideas

conecpt

1.0

.

transiormatic

full auto

.....

ovtone

tpch

bechmark

Dimensionen finden/registrieren:

- ▷ Blätter finden
- ▷ für jedes Blatt:
 - ⋄ Dimension registrieren
 - ♦ BDC-Tabelle erzeugen

intro

mcluste

proble

mdc

. .

idea:

proconditio

abstraction

transformati

transformati

cycles

full auto

exampi

CALCII

toch

bechmark

Grundvariante - (2)

BDC-Tabellen erzeugen:

> ermittle Dimensionen der abgearbeiteten Nachbartabellen

intro

mcluste bdc

proble

rolati

mdc

others

conclusio

idea

conecpt

.

transformat

cycles

full auto

Country

toch

bechmark

Grundvariante - (2)

- \triangleright sammle Pfade zu Dimtab mit Δ + Nachbar-Dimpfad

intro

mcluste

probler

relate

mdc

conclusi

idea

precondition

.

.

cycles

full auto

example

extensi

. . . .

bechmarl

Grundvariante - (2)

- ▷ ermittle Dimensionen der abgearbeiteten Nachbartabellen
- ight
 angle sammle Pfade zu Dimtab mit $\Delta + \mathsf{Nachbar} ext{-}\mathsf{Dimpfad}$
- benötigte *iobits*, aus System ermittelt

intro

mcluste bdc

proble

relat

mdc

conclusi

idea

conecpt

abstraction

transformatio

.

full auto

ovamala

extensi

toch

bechmark

Grundvariante - (2)

- ▷ ermittle Dimensionen der abgearbeiteten Nachbartabellen
- hd sammle Pfade zu Dimtab mit Δ + Nachbar-Dimpfad
- ▷ benötigte iobits, aus System ermittelt
- ▶ Masken = RoundRobin(Dims, iobits)

intro

mcluste bdc

proble

relat

mdc

conclusi

idea

conecpt

transformatio

transformatio

full auto

evamnle

evtens

tpch

bechmark

Grundvariante - (2)

- ▷ ermittle Dimensionen der abgearbeiteten Nachbartabellen
- ight
 angle sammle Pfade zu Dimtab mit Δ + Nachbar-Dimpfad
- ▷ benötigte iobits, aus System ermittelt
- ▶ Masken = RoundRobin(Dims, iobits)
- ▷ erzeuge BDC-Tabelle

intro

mcluste bdc

proble

relat

mdc

conclusi

idea

conecp

precondition

abstraction

transionnatio

cycles

full auto

exampl

. .

tpch

bechmark

Grundvariante - (2)

- ▷ ermittle Dimensionen der abgearbeiteten Nachbartabellen
- ight
 angle sammle Pfade zu Dimtab mit Δ + Nachbar-Dimpfad
- ▷ benötigte iobits, aus System ermittelt
- ▶ Masken = RoundRobin(Dims, iobits)
- ▷ speichere Diminfos

intro

mcluste bdc

. . . .

mdc

others

idea

conecpt

abstraction

transformatio

cycles

full auto

exampl

extensio

tests

bechmar

Grundvariante - (2)

- ▷ ermittle Dimensionen der abgearbeiteten Nachbartabellen
- ight
 angle sammle Pfade zu Dimtab mit $\Delta + \mathsf{Nachbar} ext{-}\mathsf{Dimpfad}$
- ▷ benötigte iobits, aus System ermittelt
- ▶ Masken = RoundRobin(Dims, iobits)
- speichere Diminfos

intro

mcluste bdc

relat

others

conclusi

idea

conecpt

abstraction

transformatio

i

full auto

......

test

bechmarl

Grundvariante - (2)

- ▷ ermittle Dimensionen der abgearbeiteten Nachbartabellen
- ight
 angle sammle Pfade zu Dimtab mit $\Delta + \mathsf{Nachbar} ext{-}\mathsf{Dimpfad}$
- ▷ benötigte iobits, aus System ermittelt
- ▶ Masken = RoundRobin(Dims, iobits)
- speichere Diminfos

intro

mclust bdc

problen

relate

mdc

conclusio

ideas

conecut

preconditio

abetraction

.

cvcles

C II .

example

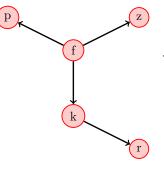
extensio

. .

tpch

bechmarks

Beispiel (Graph)



 $Transformier ter\ Schemagraph$

intro

mclust

proble

relate

mdc

conclusio

ideas

conecpt

1.0

transionnati

cycles

example

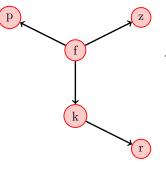
extensi

. .

tpch

bechmark

Beispiel (Graph)



Transformierter Schemagraph

- zusammenhängend
- ▷ ohne Kreise

intro

mclust

proble

relate

mdc

conclusio

ideas

conecpt

precondition

abstraction

transformati

cvcles

full auto

example

evtensio

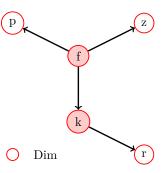
tests

tpch

bechmarks

Beispiel - (Dimensionen)

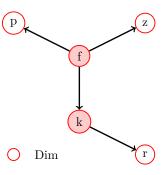
Dimensionen = Blätter



example

Beispiel - (Dimensionen)

Dimensionen = Blätter Dim-Registierung:

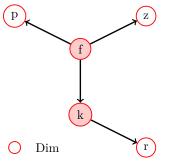


example

Beispiel - (Dimensionen)

Dimensionen = BlätterDim-Registierung:

▷ Dimspalte = Primärschlüssel



intro

bdc

problen

mdc

others conclusion

idea

conecpt

abstraction

transformation

Liuisioimuu

full auto

example

extensio

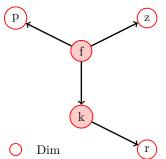
tests

bechmark

Beispiel - (Dimensionen)

Dimensionen = Blätter Dim-Registierung:

- \triangleright Dimspalte = Primärschlüssel
- $box{cbits} = \\ \min\{\lfloor \log_2(count(x)) \rfloor, max_{cbits}\} \\ \forall x \ \mathsf{Dimtab}$



intro

bdc

probler

mdc others

conclusion

idea

conecpt

abstraction

transformation

cycles

example

extensio

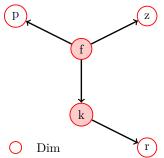
tests

bechmark

Beispiel - (Dimensionen)

Dimensionen = Blätter Dim-Registierung:

- \triangleright Dimspalte = Primärschlüssel
- $box{coits} = \\ \min\{\lfloor \log_2(count(x)) \rfloor, max_{cbits}\} \\ \forall x \ \mathsf{Dimtab}$



intro

mcluste bdc

probler

related

others

conclusi

idea

conecpt

abstraction

transformation

cycles

Tull aut

example

extensio

extension

tpch

bechmark

Beispiel - (Dimensionen)

 $\begin{array}{ll} {\sf Dimensionen} = {\sf Bl\"{a}tter} \\ {\sf Dim-Registierung:} \end{array}$

- \triangleright Dimspalte = Primärschlüssel
- $box{coits} = \\ \min\{\lfloor \log_2(count(x)) \rfloor, max_{cbits}\} \\ \forall x \ \mathsf{Dimtab}$

BDC-Tabelle erzeugen:

 \triangleright iobits = cbits



intro

mcluste bdc

problen

related

others

conclusio

idea

conecpt

abstraction

transformation

cycles

Tull aut

example

extensi

tests

tpch

Beispiel - (Dimensionen)

 $\begin{array}{l} {\sf Dimensionen} = {\sf Bl\"{a}tter} \\ {\sf Dim-Registierung:} \end{array}$

- \triangleright Dimspalte = Primärschlüssel
- $box{cbits} = \\ \min\{\lfloor \log_2(count(x)) \rfloor, max_{cbits}\} \\ \forall x \ \mathsf{Dimtab}$

BDC-Tabelle erzeugen:

- \triangleright *iobits* = *cbits*
- \triangleright mask = 1...1, iobit viele

example

Beispiel - (Dimensionen)

Dimensionen = BlätterDim-Registierung:

- ▷ Dimspalte = Primärschlüssel
- \triangleright chits = $\min\{|\log_2(count(x))|, max_{cbits}\}$ $\forall x$ Dimtab

BDC-Tabelle erzeugen:

- \triangleright iobits = chits
- \triangleright mask = 1...1, iobit viele
- $\triangleright Pfad(x) = Dimspalte$

intro

mcluste bdc

proble

related

others

idea

conecp

precondition

transformati

cycles

full aut

example

tpch

bechmark

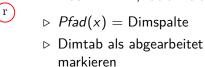
Beispiel - (Dimensionen)

 $\begin{array}{ll} {\sf Dimensionen} = {\sf Bl\"{a}tter} \\ {\sf Dim-Registierung:} \end{array}$

- \triangleright Dimspalte = Primärschlüssel
- $box{cbits} = \\ \min\{\lfloor \log_2(count(x)) \rfloor, max_{cbits}\} \\ \forall x \ \mathsf{Dimtab}$

BDC-Tabelle erzeugen:

- \triangleright *iobits* = *cbits*
- \triangleright mask = 1...1, iobit viele



intro

bdc

problen

relate

mdc

conclusio

ideac

iuea

conecpt.

abstraction

.

transformation

cycles

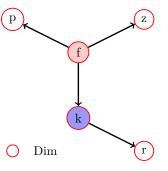
full auto

example

extensi

tpch

bechmark:



intro

mcluste bdc

relate

mdc

conclusio

ideas

conecpt

abstraction

transformati

cvcles

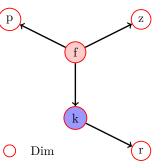
full auto

example

extensio

tocto

tpch



$$ightharpoonup Pfad(k) = (k,r) + Pfad(r)$$

intro

mcluste bdc

mdc

conclusio

idea

conecpt

1.0

abstraction

transformation

cvcles

Tull auto

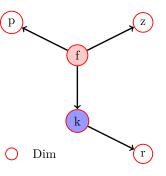
example

extensi

tpch

bechmark

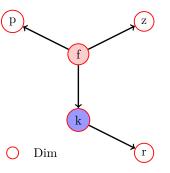
Beispiel - (Tabelle k)



$$\triangleright$$
 $Pfad(k) = (k, r) + Pfad(r)$

▷ iobits vom System ausgelesen

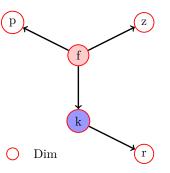
example



$$\triangleright$$
 $Pfad(k) = (k, r) + Pfad(r)$

- $mask = RoundRobin(\{r\}, iobits)$

example



$$\triangleright$$
 $Pfad(k) = (k, r) + Pfad(r)$

- $mask = RoundRobin(\{r\}, iobits)$
- k abgearbeitet

intro

bdc

relate

mdc

conclusio

ideas

conecut

preconditi

abstraction

transformatic

example

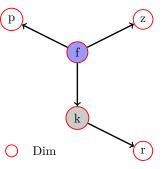
evtensio

CACCITATO

tests

tpch

bechmarks



intro

bdc

relate

mdc

conclusio

ideas

conecpt

abstraction

transformation

cycles

Tull auc

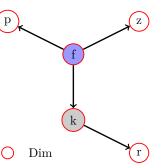
example

ovtonci

tecto

tpch

bechmarks



$$hd Pfad_1(f) = (f,p) + Pfad(p)$$

intro

mclust bdc

relate

mdc

conclusio

idea

conecpt

abstraction

transformati

cycles

Tull auc

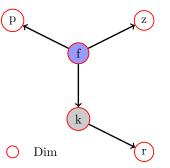
example

extensio

tocto

tpch

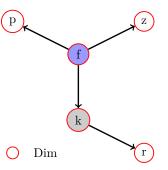
bechmark



$$\triangleright$$
 $Pfad_1(f) = (f, p) + Pfad(p)$

$$Pfad_2(f) = (f, z) + Pfad(z)$$

example



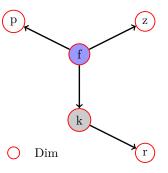
$$\triangleright$$
 $Pfad_1(f) = (f, p) + Pfad(p)$

$$\triangleright$$
 $Pfad_2(f) = (f, z) + Pfad(z)$

$$\triangleright$$
 $Pfad_3(f) = (f, k) + Pfad(k)$

example

Beispiel - (Tabelle f)



$$\triangleright$$
 $Pfad_1(f) = (f, p) + Pfad(p)$

$$\triangleright Pfad_2(f) = (f,z) + Pfad(z)$$

$$\triangleright$$
 $Pfad_3(f) = (f, k) + Pfad(k)$

abstraction

.

cycles

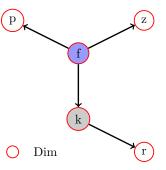
Tull auto

example

extensio

toch

bechmark



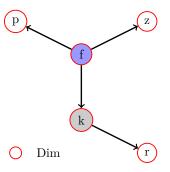
$$\triangleright$$
 $Pfad_1(f) = (f, p) + Pfad(p)$

$$\triangleright$$
 $Pfad_2(f) = (f, z) + Pfad(z)$

$$\triangleright$$
 $Pfad_3(f) = (f, k) + Pfad(k)$

- > masks = RoundRobin({p,z,k},iobits)

bechmark



$$\triangleright$$
 $Pfad_1(f) = (f, p) + Pfad(p)$

$$\triangleright Pfad_2(f) = (f,z) + Pfad(z)$$

$$\triangleright$$
 $Pfad_3(f) = (f, k) + Pfad(k)$

- ▷ iobits vom System ausgelesen
- > masks =
 RoundRobin({p, z, k}, iobits)

transforma

cycles

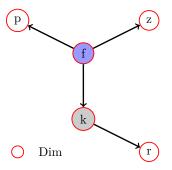
full aut

example

extensio

toch

bechmark



$$\triangleright$$
 $Pfad_1(f) = (f, p) + Pfad(p)$

$$\triangleright$$
 $Pfad_2(f) = (f, z) + Pfad(z)$

$$\triangleright$$
 $Pfad_3(f) = (f, k) + Pfad(k)$

- ▷ iobits vom System ausgelesen
- > masks = RoundRobin({p, z, k}, iobits)
- \triangleright f abgearbeitet \rightarrow fertig

Erweiterungen

intr

mclust bdc

relate

mdc

conclusi

idea

....

preconditio

abstraction

transformatio

cycles

full auto

example

extensions

tests

tpch

oechmark:

Beschrieben:

> Wunschdimensionen

Offen:

- ▶ Anfrageanalyse vorschalten
- \triangleright Schema = Kreis

Auto BDC

Steve Göring

intro

mcluste

probler

relate

mdc

conclusio

ideas

conecpt

ubstruction

transformat

.

cycles

example

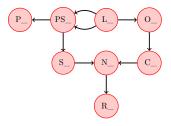
extensio

test

tpch

bechmark

TPC-H Test



DB Test für Data-Warehouse Szenarien

Auto BDC

Steve Göring

intro

mclust hdc

proble

relate

mdc

conclusio

ideas

conecpt

.....

transformati

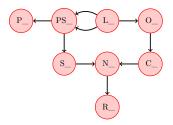
c 11 .

.

tpch

bechmark:

TPC-H Test



- > DB Test für Data-Warehouse Szenarien
- ▶ Anfragen
 - ♦ komplex
 - ohne Vorwissen
 - ⋄ große Datenmengen (abhängig von SF)

intro

mclust

proble

relate

mdc

conclusi

idea

conecpt

abstraction

trunsionnut

cycles

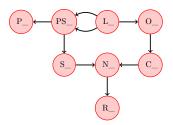
Tull auc

evtensio

tpch

bechmark

TPC-H Test



- DB Test für Data-Warehouse Szenarien
- ▶ Anfragen
 - ⋄ komplex
 - ohne Vorwissen
 - ⋄ große Datenmengen (abhängig von SF)
- ▷ gefundene Dimensionen: R_, P_
- \triangleright $SF=1 \rightarrow \mathsf{ca.}\ 1\ \mathsf{GB}$

Auto BDC

Steve Göring

intro

mclust

proble

rolato

mdc others

conclusio

ideas

conecpt

abetraction

abstraction

cycles

full auto

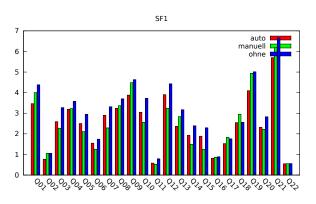
example

extension

toch

bechmarks

Benchmark



3 Wdh., pro Query Kaltstart, Anfragen nicht auf Schema optimiert Q5 Umsätze der lokalen Anbieter in S



Fragen?

bechmarks



Vielen Dank für Ihre Aufmerksamkeit.