

2. Chmora A. Sovremennaia prikladnaia kriptografiia / A. Chmora. – М. : Gelios ARV, 2005. – 232 p.
3. Kernigan B. Praktika programmirovaniia / B. Kernigan, R. Paik. – SPb. : Nevskii dialect, 2006. – 340 p.
4. Knut A. Iskusstvo programmirovaniia / A. Knut. – М. : Viliams, 2007. – Т. 2: Poluchislennii algoritmi. – 278 p.
5. Kormen T. Algoritmi: postroenie i analiz / T. Kormen, Ch. Leizerson, R. Rivest. – М. : MCNMO, 2001. – 321 p.
6. Kultin N. A. Delphi 7 programmirovanie na Object Pascal / N. A. Kultin. – SPb. : Nevskii dialect, 2007. – 370 p.
7. Levitin G. N. Algoritmi. Vvedenie v razrabotku i analiz / G. N. Levitin. – М. : Viliams, 2008. – 287 p.
8. Virt N. Algoritmi + structuri dannih = programmi / N. Virt. – М. : Mir, 1985. – 227 p.

УДК 004.652.5

ПРИМЕР РАБОТЫ АЛГОРИТМА ТРАНСФОРМАЦИИ КУБА ДАННЫХ С ИСПОЛЬЗОВАНИЕМ МАТРИЦ ТРАНСФОРМАЦИИ

Спандерашвили Дмитрий Викторович, кандидат технических наук, Астраханский государственный университет, 414056, Россия, г. Астрахань, ул. Татищева, 20а; начальник отдела организации эксплуатации систем коммутации и сетей доступа Астраханского филиала ОАО «Ростелеком», 414000, Россия, г. Астрахань, Театральный пер., 7/8, e-mail: spanderashvili@gmail.com.

С течением времени в многомерных структурах данных могут происходить изменения. Для проведения достоверного анализа данных необходимо отслеживать все изменения в структуре и элементах измерений. В статье в краткой форме дано формальное описание механизма трансформации куба данных из одной структурной версии в другую, для обеспечения возможности проводить анализ данных из куба данных, претерпевшего структурные изменения. В первой части статьи дается описание механизма использования матриц трансформации для приведения куба данных к заданной структурной версии, а во второй части приводится пример использования алгоритма трансформации на тестовых данных. В краткой форме в статье дано формальное описание механизма темпорального хранения данных. Сформулированы предпосылки применения темпорально-многомерного хранилища данных, вводится формальное описание темпорально-многомерного хранилища данных, определяется механизм формирования соответствия между структурными версиями, описан механизм применения матриц трансформации. В качестве примера работы алгоритма трансформации куба данных из одной структурной версии в другую рассматривается тестовый пример, в котором требуется сравнить данные куба одной структурной версии с данными куба данных из другой структурной версии. Путем предлагаемого метода проводятся преобразования, благодаря которым такое сравнение становится возможным.

Ключевые слова: *многомерные структуры данных, изменения, анализ данных, измерение, трансформация, куб данных, структурная версия, матрицы трансформации, алгоритм трансформации, темпорально-многомерные данные.*

**EXAMPLE OF THE TRANSFORMATION ALGORITHM
OF DATA CUBE USING TRANSFORMATION MATRIX**

Spanderashvili Dmitry V., Ph.D., Astrakhan State University, 20a Tatishchev str., Astrakhan, 414056, Russia, Head of Department of organization of operating switching systems and access networks, Astrakhan branch of OJSC "Rostelecom", 7/8 Teatralny lane, Astrakhan, 414000, Russia, e-mail: spanderashvili@gmail.com.

Over time, the multi-dimensional data structures can be changed. For a reliable analysis of the data it is necessary to track all changes in the structure and elements of the measurement. The article briefly gives a formal description of the mechanism of transformation of the data cube from one structure version to another one, in order to hold the analysis of data from the data cube that has undergone structural changes. The first part of the article describes the mechanism of transformation matrix to align the data cube to a given structural version, while the second part is an example of the transformation algorithm on test data. A formal description of temporal data warehousing mechanism is given shortly in the article. The premises of spatial-temporal data warehouse application are formulated, a formal description of spatial-temporal data warehouse is introduced, the mechanism of structural versions mapping formation is determined, the transformation matrices usage mechanism is described. As an example of the algorithm of transformation of the data cube from one version to another structure is considered a test case: you want to compare the data from data cube in a one structural version with the data from data cube structure of another version. Transformations are carried out the by the proposed method due to which such a comparison is possible.

Key words: *multi-dimensional data structures, changes, data analysis, measurement, transformation, data cube, structure version, transformation matrix, transformation algorithm, spatial-temporal data.*

Механизм использования матриц трансформации для приведения куба данных к заданной структурной версии. Цель трансформации хранилища данных в том, чтобы отобразить куб структурной версии SV_i в структуру SV_j .

Определим операции трансформации над хранилищем данных:
 $T(SV_i, SV_j, D_{id}, D'_{id})$ – матрица трансформации измерения D_k при отображении куба данных структурной версии SV_i в структуру структурной версии SV_j [3]. Матрица строится следующим образом: по вертикали откладываются элементы измерения [3] $D_k(SV_i)$, а по горизонтали – элементы $D'_k(SV_j)$, на пересечении выставляются коэффициенты w ;

$C(SV_i)$ – куб в структурной версии SV_i ;

$C(SV_j)$ – куб в структурной версии SV_j ;

$C(SV_i, SV_i)$ – куб $C(SV_i)$ в структурной версии SV_i ;

$C(SV_i, SV_j)$ – куб $C(SV_i)$ в структурной версии SV_j ;

$C(SV_i, SV_i, D_n)$ – двумерная матрица, представляющая собой развертку куба данных $C(SV_i, SV_i)$, такую, что по горизонтали располагаются элементы D_n , а по вертикали – все возможные комбинации элементов остальных измерений;

$C(SV_i, SV_j, D'_n)$ – двумерная матрица, представляющая собой развертку куба данных $C(SV_i, SV_j)$, такую, что по горизонтали располагаются элементы D'_n , а по вертикали – все возможные комбинации элементов остальных измерений; тогда можно определить следующую формулу [2]:

$$C(SV_i, SV_j, D'_n) = C(SV_i, SV_i, D_n) * T(SV_i, SV_j, D_n, D'_n) \quad (1)$$

Операция перемножения в данной формуле означает обычное перемножение матриц (так при перемножении $C(D_n)[m \times n]$ на $T(SV_1, SV_2, D_n)[n \times k]$ получаем $C'(D'_n)[m \times k]$; m – количество комбинаций членов измерений без D_n , n – количество элементов D_n , k – количество элементов измерения D'_n .

Возможны следующие варианты:

$n = k$ – перегруппировка мощности элементов измерения;

$n > k$ – объединение элементов измерения;

$n < k$ – дробление элементов измерения;

$n = 1, k > 1$ – порождение нового измерения с k элементами;

$n > 1, k = 1$ – вырождение одного измерения (в частном случае – агрегация).

Формально из этого следует, что в любом кубе может существовать бесконечное множество измерений, которые могут свободно вырождаться и порождаться при правильно подобранных коэффициентах w без потери качества данных.

Для преобразования всего куба необходимо произвести поочередное перемножение разверток куба по измерениям на соответствующие матрицы преобразования этих измерений.

Преобразование всего куба данных $C(SV_i, SV_j)$ в $C(SV_i, SV_j)$ можно описать следующей формулой (2)[2]:

$$C(SV_i, SV_j) = C(SV_i, SV_i, D_n) * T(SV_i, SV_j, D_n, D'_n); n \in [1..N] \quad (2)$$

где N – количество различных измерений двух структурных версий, т.е. если N_D и $N_{D'}$ – количество измерений в структурных версиях SV_i и SV_j , то $N = N_D \cup D' - N_D \cap D'$.

При организации матриц трансформации двух структурных версий необходимо создать в каждой структурной версии недостающие измерения (псевдоизмерения), не содержащие ни одного элемента.

Пример работы алгоритма трансформации данных с использованием матриц трансформации (с использованием тестовых данных). В качестве примера работы алгоритма трансформации куба данных из одной структурной версии в другую рассмотрим тестовый пример.

Пусть требуется сравнить данные куба $C(SV_2)$ из структурной версии SV_2 с данными куба $C(SV_1)$ из структурной версии SV_1 . В качестве точки отсчета выбираем T_Q , которая лежит в интервале валидности SV_2 .

Обозначения:

$C(SV_1)$ – куб данных с данными, актуальными в период валидности структурной версии SV_1 ;

$C(SV_2)$ – куб данных с данными, актуальными в период валидности структурной версии SV_2 ;

$C(SV_1, SV_1)$ – куб данных $C(SV_1)$, рассматриваемый в структурной версии SV_1 ;

ИНФОРМАЦИОННО-ТЕЛЕКОММУНИКАЦИОННЫЕ СИСТЕМЫ И ТЕХНОЛОГИИ

$C(SV_2, SV_2)$ – куб данных C_{SV_2} , рассматриваемый в структурной версии SV_2 .

Сравнивать данные $C(SV_2)$ и $C(SV_1)$ напрямую нельзя, так как они имеют различную структуру. Так как в качестве точки отсчета T_Q выбираем точку, в которой валидна SV_2 , будем рассматривать данные в структурной версии SV_2 . Данные $C(SV_2)$ не нуждаются в преобразовании, так как уже находятся в требуемой структурной версии. Данные же $C(SV_1)$ необходимо привести к версии SV_2 . Для такого преобразования необходимо иметь матрицы трансформации из SV_1 в SV_2 по всем измерениям SV_1 и SV_2 .

В SV_1 содержатся измерения: $X(X_1, X_2); Z(Z_1, Z_2, Z_3, Z_4); Y(Y_1, Y_2, Y_3)$. В SV_2 содержатся измерения $X(X_{11}, X_{12}, X_{21}, X_{22}); Z(Z_{11}, Z_{22}, Z_{33}); W(W_{11}, W_{12}, W_{13})$.

Таким образом, происходит перегруппировка по измерениям X и Z, вырождается измерение Y и порождается измерение W.

Итак, для сравнения имеем следующие данные. $C(SV_2, SV_2)$ – для наглядности на рисунке 1 приводим развертку куба $C(SV_2, SV_2, W)$.

		W11	W12	W13
Z 11	Z11-X11	6,18	9,25	12
	Z11-X12	6,22	9,4	15
	Z11-X21	7,25	10,2	16
	Z11-X22	7,3	11,5	17
Z 22	Z22-X11	2	3,5	3,2
	Z22-X12	2,2	3,6	3,3
	Z22-X21	2,5	6,5	5,9
	Z22-X22	2,6	5,3	5,8
Z 33	Z33-X11	7	13,2	16,5
	Z33-X12	8	9,5	15,3
	Z33-X21	9,5	12,5	19,2
	Z33-X22	9,6	9,4	20,1

Матрица трансформации $T(SV_1, SV_2, X)$ представлена на рисунке 3.

	X11	X12	X21	X22
X1	0,5	0,5	0	0
X2	0	0	0,5	0,5

ИНФОРМАЦИОННО-ТЕЛЕКОММУНИКАЦИОННЫЕ СИСТЕМЫ И ТЕХНОЛОГИИ

Далее, приведя полученную развертку к развертке по Z , получаем результат, приведенный на рисунке 8.

		Z1	Z2	Z3	Z4
Y1	X11	0,5	3,5	6,5	9,5
	X12	0,5	3,5	6,5	9,5
	X21	1	4	7	10
	X22	1	4	7	10
Y2	X11	1,5	4,5	7,5	10,5
	X12	1,5	4,5	7,5	10,5
	X21	2	5	8	11
	X22	2	5	8	11
Y3	X11	2,5	5,5	8,5	11,5
	X12	2,5	5,5	8,5	11,5
	X21	3	6	9	12
	X22	3	6	9	12

ПРИКАСПИЙСКИЙ ЖУРНАЛ:
управление и высокие технологии № 3 (15) 2011

Переведа результат к развертке по W и помножив на $T(SV_1, SV_2, W)$, получаем результат, приведенный на рисунке 11.

		W11	W12	W13
Z11	X11	5,58	8,37	13,95
	X12	5,58	8,37	13,95
	X21	6,06	9,09	15,15
	X22	6,06	9,09	15,15
Z22	X11	1,89	2,84	4,725
	X12	1,89	2,84	4,725
	X21	2,1	3,15	5,25
	X22	2,1	3,15	5,25
Z33	X11	6,93	10,4	17,33
	X12	6,93	10,4	17,33
	X21	7,44	11,2	18,6
	X22	7,44	11,2	18,6

**ИНФОРМАЦИОННО-ТЕЛЕКОММУНИКАЦИОННЫЕ
СИСТЕМЫ И ТЕХНОЛОГИИ**

		W11	W12	W13
Z 11	Z11-X11	5,58	8,37	13,95
	Z11-X12	5,58	8,37	13,95
	Z11-X21	6,06	9,09	15,15
	Z11-X22	6,06	9,09	15,15
Z 22	Z22-X11	1,89	2,835	4,725
	Z22-X12	1,89	2,835	4,725
	Z22-X21	2,1	3,15	5,25
	Z22-X22	2,1	3,15	5,25
Z 33	Z33-X11	6,93	10,4	17,33
	Z33-X12	6,93	10,4	17,33
	Z33-X21	7,44	11,16	18,6
	Z33-X22	7,44	11,16	18,6

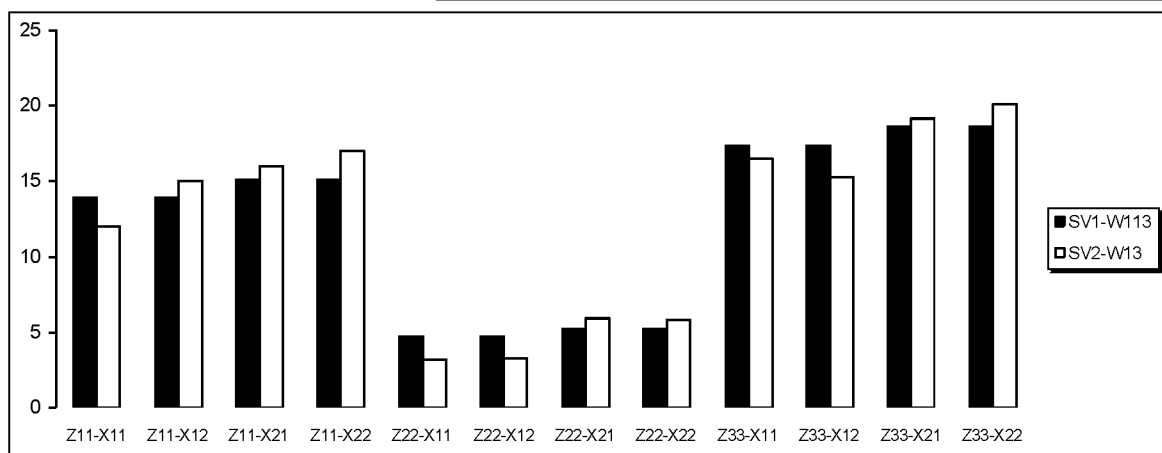


Рис. 16. Сравнение W13

На рисунках 14–16 приведены результаты сравнения.

Список литературы

1. Спандерашвили Д. В. Механизмы отслеживания изменений в многомерных структурах данных / Д. В. Спандерашвили // Инфокоммуникационные технологии в науке, производстве и образовании : мат-лы II Междунар. науч.-техн. конф. – Ставрополь : СКГТУ, 2006. – Ч. 1. – С. 160–62.
2. Спандерашвили Д. В. Формальное описание модели и алгоритмы трансформации темпорального хранилища данных / Д. В. Спандерашвили // Южно-Российский вестник геологии, географии и глобальной энергии. – Астрахань, 2006. – № 7 (20). – С. 136–142.
3. Eder J. Analysing Slices of Data Warehouses to Detect Structural Modifications / J. Eder, C. Koncilia, D. Mitsche // Advanced Information Systems Engineering : 16th International Conference (CAiSE – 2004). – 2004. – P. 492–505.
4. Eder J. Changes of Dimension Data in Temporal Data Warehouses / J. Eder, C. Koncilia // Data Warehousing and Knowledge Discovery : Third International Conference (DaWaK – 2001). – 2001. – P. 284–293.

References

1. Eder J. Analysing Slices of Data Warehouses to Detect Structural Modifications / J. Eder, C. Koncilia, D. Mitsche // Advanced Information Systems Engineering : 16th International Conference (CAiSE – 2004). – 2004. – P. 492–505.
2. Eder J. Changes of Dimension Data in Temporal Data Warehouses / J. Eder, C. Koncilia // Data Warehousing and Knowledge Discovery : Third International Conference (DaWaK – 2001). – 2001. – P. 284–293.
3. Spanderashvili D. V. Formalnoe opisaniye modeli i algoritmi transformacii temporalnogo hranilishcha danih / D. V. Spanderashvili // Uzhno-Rossiisky vestnik geologii, geografii i globalnoi energii. – Astrakhan, 2006. – № 7 (20). – P. 136–142.
4. Spanderashvili D. V. Mehanizmi otslezhivaniya izmenenii v mnogomernih strukturah danih / D. V. Spanderashvili // Infokommunikacionnie tehnologii v nauke, proizvodstve i obrazovanii : materialy II mezhdunarodnaia nauchno-tehnicheskaya konferenciya. – Stavropol : SKGTU, 2006. – Ch. 1. – P. 160–162.