А.А. Ахрем, В.З. Рахманкулов, К.В. Южанин

О сложности редукции моделей многомерных данных[1]

Аннотация. В работе исследованы методы декомпозиции гиперкубов многомерных данных аналитических OLAP-систем на подкубовые компоненты. Установлены критерии уменьшения вычислительной сложности методов декомпозиции по сравнению с традиционными не редукционными методами решения задач анализа многомерных данных. Рассмотрены примеры применения указанных критериев при исследовании динамики изменения величины вычислительной сложности конкретных типов задач редукции.

Ключевые слова: гиперкуб, многомерные данные гиперкуба, вычислительная сложность, методы декомпозиции.

Введение

Аналитические OLAP-системы предназначены для анализа и обобщения детальных данных, накапливаемых в базах и хранилищах данных. Для таких систем характерен быстрый рост объемов обрабатываемых данных. OLAP-кубы могут включать полный объем исходных детальных данных. Объем данных увеличивается в результате денормализации и дублирования части детальных данных, кроме того, он может расти при добавлении к детальным данным агрегированной информации, особенно, при полном переборе сочетаний размерностей и их значений. С увеличением объемов данных падает производительность вычислений кубовых структур, при этом иногда возникают эффекты «взрывного» характера данных, когда в аналитической системе резко прекращаются вычисления. Устранению подобных проблем способствуют методы редукции больших кубов данных на подкубы с меньшими объемами. Очевидно, что уменьшение объемов данных в подкубах способствует повышению производительности вычислений, однако после обработки данных подкубов необходимо получить обобщенные результаты, которые не отличались бы от результатов вычислений на полном исходном кубе. Это условие часто создает проблемы, вызванные редукцией больших кубов на подкубы. Например, данные OLAP-кубов могут быть сегментированы не аддитивно и тогда возможны нарушения целостности данных. Если редукция исходного полного куба одновременно затрагивает число размерностей и число значений на каждой шкале размерностей, это может привести к совершенно неожиданным результатам с точки зрения производительности вычислений. Следовательно, решая задачи редукции OLAP-кубов, необходимо исследовать влияние выбранного метода редукции на вычислительную сложность задач и целостность обобщенных результатов.

В данной статье, являющейся продолжением цикла работ авторов [1, 2] по системам многомерного анализа данных и аналитическим OLAP-системам, исследованы методы редукции моделей многомерных данных в виде гиперкубов, способствующих уменьшению вычислительной сложности задач с большими и сверхбольшими исходными данными. Для ряда специальных классов декомпозиционных задач получены границы изменения вычислительной сложности при варьировании основных параметров гиперкуба. На их основе установлены необходимые и достаточные условия эффективного применения методов декомпозиции

[1] Работа выполнена при финансовой поддержке РФФИ (проекты №15-29-04827; 16-07-00644).

задач анализа гиперкубов многомерных данных по сравнению с традиционными нередукционными методами их решения.

Изложенные в статье результаты по редукции OLAP-кубов используются в междисциплинарном проекте РФФИ, выполняемом совместно медиками и специалистами в области управления и информационной технологии, исследующими принципы и методы виртуального моделирования искусственных биологических органов на основе моделей OLAP и Data Mining. Подобные подходы применяют также отечественные и зарубежные исследователи [3, 4, 7].

Проблемы редукции моделей и анализа вычислительной сложности алгоритмов традиционно интересуют исследователей-математиков и компьютерных специалистов при практической реализации методов моделирования [5, 6, 8-11]. В частности, проблеме снижения размерности признакового пространства в задачах ранжирования и многокритериальной классификации посвящена работа [12], в которой критерии агрегируют с помощью экспертов, что упрощает порядковую классификацию многокритериальных альтернатив в статическом пространстве оценок.

В данной работе задача ставится иначе – агрегирование критериев уже определено решеткой куба, а редукция куба на меньшие по размерности кубы нужна для снижения времени вычислений полной решетки при динамическом изменении данных в кубе.

1. Постановка задачи

Предположим, что задан куб H_n многомерных данных, имеющий n размерностей [3-5]. Допустим, что для решения некоторой задачи (класса задач) на H_n используется подмножество $K_n(m)$ гиперкуба H_n, образованное данными с m (m<n) размерностями с фиксированными значениями. Предположим дополнительно, что сложность решения задачи допускает следующие оценки:

$$\text{M(n, a)} = n^a; \quad (1)$$

$$\text{A(m, n, a)} < m^a; \quad (2)$$

$$\text{B(n-m, n, a)} < (n-m)^a, \quad (3)$$

где M – сложность решения исходной задачи на гиперкубе H_n; A, B – сложность решения задачи на подмножествах $K_n(m)$ и $H_n/K_n(m)$ соответственно; m(m < n) – фиксированное натуральное число; a(a > 1) – коэффициент сложности решения исходной задачи на гиперкубе H_n.

Справедливо следующее утверджение.

Утверждение 1.1. Величины A, B, M удовлетворяют неравенству:

$$\text{A + B < M}. \quad (4)$$

Доказательство Утверждения 1.1. Имеет место следующая необходимая для доказательства Утверждения 1

Лемма 1.1. Пусть даны числа x, y удовлетворяющие неравенствам:

$$\begin{aligned} &0 < x < 1,\ 0 < y < 1; \\ &x^{a_1} + y^{b_1} < 1,\ a_1 > 0,\ b_1 > 0. \end{aligned} \quad (5)$$

Тогда для любых чисел $c_1 > a_1$, $d_1 > b_1$ выполнено соотношение:

$$x^{c_1} + y^{d_1} < 1.$$

Доказательство леммы 1.1. Для чисел x, y, c_1, d_1 имеют место неравенства:

$$x^{c_1} = x^{c_1-a_1} \cdot x^{a_1} < x^{a_1}; \quad (6)$$

$$y^{d_1} = y^{d_1-b_1} \cdot y^{b_1} < y^{b_1}. \quad (7)$$

Используя (5)-(7), окончательно получаем:

$$x^{c_1} + y^{d_1-b_1} < x^{a_1} + y^{b_1} < 1,$$

что и требовалось доказать.

Доказательство Утверждения 1.1. Принимая во внимание неравенства (1)-(3), находим, что

$$\text{A + B} < m^a + (n-m)^a; \quad (8)$$

$$\text{M} = n^a. \quad (9)$$

Так как a > 1, $\frac{m}{n} + \left(1 - \frac{m}{n}\right) = 1$, то применяя результат Леммы 1.1 при $a_1 = b_1 = 1$, $c_1 = d_1 = a$ будем иметь:

$$n^a \cdot \left[\left(\frac{m}{m}\right)^a + \left(1 - \frac{m}{n}\right)^a\right] < n^a = M. \quad (10)$$

Из цепочки неравенств (10) окончательно заключаем, что A + B < M, что и требовалось установить.

Утверждение 1.1 полностью доказано.

Отметим, что неравенство (4) устанавливает верхнюю границу изменения величины сложности декомпозиции решения задачи на гиперкубе данных по сравнению со сложностью исходной задачи, решаемой традиционным методом перебора данных. При m = 1 из неравенства (4) Утверждения 1.1 заключаем, что

$$\text{A + B} < 1 + (n-1)^a < n^a = \text{M}. \quad (11)$$

Из неравенства (11) находим, что при m=1 (т.е. когда выделяется лишь одна размерность гиперкуба данных H_n) справедливы соотношения:

$$M^{-1} \cdot (A + B) < n^{-a} \cdot (1+(n-1)^{a}) = \\ = n^{-a} + (1-\frac{1}{n})^{a}. \quad (12)$$

Неравенство (12) показывает, во сколько раз уменьшается сложность решения исходной задачи на гиперкубе H_n при выделенной одной из его размерностей. Допустим теперь, что сложности A, B, M удовлетворяют оценкам:

$$A(m, n, a) = A(p, n, a) = p^{a} \cdot n^{a}; \quad (13)$$

$$B(n - m, n, a) = A((1-p)n, n, a) = (1 - p)a \cdot n^{a}; \quad (14)$$

$$M(n, a) = n^{a}, \quad (15)$$

где m=p·n, p=const, 0<p<1, a>1.

Для сложностей A, B, M в данном случае справедливо следующее утверждение.

Утверждение 1.2. Сложности A, B, M при любых значениях p, 0<p<1, удовлетворяют неравенству

$$M \cdot 2^{1-a} < A + B \quad (16)$$

Доказательство. Используя (13)-(15), находим, что

$$A + B = [p^{a} + (1-p)^{a}] \cdot n^{a} = s(p)\ M, \quad (17)$$

где s(p) ≡ p^{a}+ (1 - p)a.

Для функции s(p) справедливо следующее соотношение:

$$s(p) > 2^{1-a} \quad (18)$$

для любых p, 0<p<1; a<1 .

Для доказательства (18) найдем сначала точку минимума p_{min} функции (18). Точка p_{min} должна удовлетворять следующей системе неравенств [10]:

$$ds/dp(p_{min}) = \\ = a[(p_{min})^{a-1} - (1-p_{min})^{a-1}] = 0; \quad (19)$$

$$d^2s/dp^2(p_{min}) = \\ = a(a-1)[(p_{min})^{a-2} + (1-p_{min})^{a-2}] > 0.$$

Нетрудно показать, что система неравенств (19) имеет единственное решение p_{min}:

$$p_{min} = 1/2. \quad (20)$$

Используя (17) окончательно получаем:

$$A + B = n^{a} \cdot s(p) > n^{a} \cdot s(p_{min}) = \\ = 2^{1-a} \cdot n^{a} = 2^{1-a} \cdot M,$$

что и требовалось доказать.

Отметим, что соотношение (16) дает нижнюю границу изменения величины сложности декомпозиции задачи на H_n по сравнению со сложностью исходной задачи, решаемой традиционным методом перебора данных.

Рассмотрим, наконец, случай, когда сложности A, B, M удовлетворяют соотношениям:

$$A = A(m, n, b) = m^{b};$$
$$B = B(n - m, n, c) < (n-m)^{c};$$
$$M = M(n, c) = n^{c};$$
$$l = \frac{m}{n}, \quad (21)$$

где m = 1,…,n , 0< l <1 , b=const, c>b.

Этот случай описывает ситуацию, при которой размерности гиперкуба отличаются по сложности. В данной ситуации при больших значениях натурального параметра n будем иметь:

$$A + B = m^{b} + (n-m)^{c} = \\ = \left[\left(\frac{m}{n}\right)^{b} \cdot n^{b-c} + \left(1-\frac{m}{n}\right)^{c}\right] \cdot n^{c} = \\ = [l^{b} \cdot n^{b-c} + (1-l)^{c}] \cdot n^{c} = (1-l)^{c} \cdot M \quad (22)$$

Из цепочки соотношений (22) заключаем, что:

а) сложность решения декомпозиционного варианта задачи на гиперкубе многомерных данных H_n уменьшается по сравнению со сложностью традиционного переборного метода ее решения в $(1-l)^{-c}$ раз;

б) коэффициент уменьшения сложности k = $(1-l)^{-c}$ является монотонно убывающей функцией своих аргументов l, c:

k(l’, c’) > k(l ”, c”) при l ” > l’ , c” > c’ , (0 < l < 1).

Таким образом, в данном разделе показано, что уменьшение числа размерностей гиперкуба, как правило, снижает вычислительную сложность декомпозиции гиперкубов OLAP-систем.

2. Верхняя и нижняя оценки сложности решения декомпозиционных задач на многомерных гиперкубах данных, принадлежащих классу задач с экспоненциальной сложностью

Пусть сложности M, A, B традиционного и декомпозиционного методов решения задач на гиперкубе H_n принадлежат классу задач с экспоненциальной сложностью:

$$M(n, a) = a^{n};$$
$$A(m, n, a) = a^{m};$$

$$B(n - m, n, a) = a^{n-m}. \quad (23)$$

Для сложностей M, A, B из экспоненциального класса (23) имеет место следующее утверждение, дающее верхнюю и нижнюю оценки сложности задач декомпозиции по сравнению с исходной нередукционной задачей на H_n.

Теорема 2.1. Для величин M, A, B из класса (23) имеют место следующие неравенства:

1. Для любых натуральных n, m, n>m, вещественном a ≥ 2

$$A(m, n, a) + B(n - m, n, a) < M(n, a). \quad (24)$$

2. Для любых натуральных n, m, n>m, n=2m, вещественном a ≥ 2

$$A(m, n, a)+B(n - m, n, a)>2a^{-n/2}M(n, a). \quad (25)$$

Доказательство. Положим:

$$A(m, n, a) = A,\ B(n - m, n, a) = B,\ M(n, a) = M. \quad (26)$$

Имеем:

$$(a^{n-m} + a^{m}) \cdot a^{-n} = a^{-m} + a^{m-n} \quad (27)$$

Так как n>m+1, то из соотношения (26) находим

$$(a^{n-m} + a^{m}) \cdot a^{-n} < a^{-1} + a^{-1} = 2a^{-1} \leq 1. \quad (28)$$

Учитывая (26), (28), окончательно получаем, что

$$A(m, n, a) + B(n - m, n, a) < M(n, a) \quad (29)$$

при любых натуральных n, m, n>m, a ≥ 2.

Неравенство (24) доказано.

Докажем теперь неравенство (25). Положим:

$$r(x) = a^{x} + a^{n-x}. \quad (30)$$

Найдем экстремум функции r(x). Имеем [10]:

$$\dot{r}(x) = (a^{x} - a^{n-x}) ln\, a; \quad (31)$$

$$\ddot{r}(x) = (a^{x} + a^{n-x}) ln^2 a. \quad (32)$$

Используя соотношения (31), (32), заключаем, что функция *r(x)* имеет минимум в точке x_0, являющейся решением уравнения:

$$x = n - x. \quad (33)$$

Из (33) находим, что

$$x_{min} = \frac{n}{2} = m;$$

$$r_{min}(m) = a^{n-m} + a^{m} = 2a^{m}. \quad (34)$$

Принимая во внимание (26), (34), получаем

$$[A(m, n, a) + B(n - m, n, a)]M^{-1}(n, a) =$$
$$=(A+B) \cdot M^{-1} > r_{min}(m) \cdot M^{-1} =$$
$$= 2a^{m} \cdot a^{-n} \cdot M(n, a) = 2a^{-n/2} \cdot M(n, a) =$$
$$= 2a^{-m} \cdot M(n, a).$$

Неравенство (25) установлено. Теорема 2.1 полностью доказана.

Отметим, что Теорема 2.1 показывает, в каких диапазонах изменяется сложность декомпозиционных задач анализа данных на гиперкубе по сравнению со сложностью исходной традиционной задачи.

3. О сложности декомпозиционных задач на двумерной плоскости

В данном разделе будем оценивать сложность задач редукции, заданных на прямоугольниках со сторонами a > 0, b > 0, $a^{-1} \cdot b = c \geq 1$. Предположим, что грань гиперкуба из этого класса разбита на k непересекающихся одинаковых подкубов со сторонами a, $b \cdot k^{-1}$ (Рис. 1).

Будем считать, что сложности $S,\ S'$ решения задачи традиционным и декомпозиционным методами равны:

$$S = a^2 + b^2; \quad (35)$$

$$S_i' = d \cdot (a^2 + b^2 \cdot k^{-2}),\ i = 1, \dots, k, \quad (36)$$

где $d \geq 1$ – некоторая константа.

Найдем критерии эффективности применения редукционного метода для этого класса двумерных гиперкубов. Имеет место следующее утверждение.

Теорема 3.1:

1. Пусть $d = 1$, тогда $S' < S \Leftrightarrow c > \sqrt{k}$.

2. Пусть $d > 1$, тогда $S' < S \Leftrightarrow k + k^{-1} \cdot c^2 < d^{-1}(1 + c^2)$.

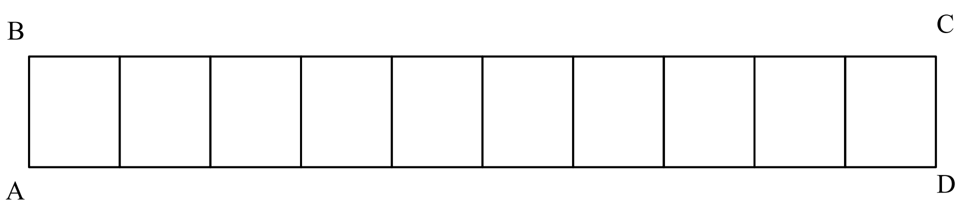

Рис.1. Грань гиперкуба ABCD, AB = a, BC = b, k = 10

Доказательство:

1. Пусть вначале $d = 1$. Используя (35), (36), находим

$$(S') \cdot S^{-1} = k(1 + c^2 \cdot k^{-2}) \cdot (1 + c^2)^{-1}. \quad (37)$$

Принимая во внимание (37), получаем

$$S' < S \Leftrightarrow k(1 + k^{-2} \cdot c^2) < 1 + c^2. \quad (38)$$

Учитывая (38), находим

$$S' < S \Leftrightarrow c^2 > k \Leftrightarrow c > \sqrt{k},$$

что и требовалось доказать.

2. Пусть теперь $d > 1$. Имеем

$$(S') \cdot S^{-1} = d \cdot k(1 + k^{-2} \cdot c^2) \cdot (c^2 + 1)^{-1}. \quad (39)$$

Используя (39), окончательно находим

$$S' < S \Leftrightarrow k + k^{-1} \cdot c^2 < d^{-1} \cdot (c^2 + 1),$$

что и требовалось доказать. Теорема 3.1 доказана полностью.

Отметим, что Теорема 3.1 дает необходимые и достаточные условия эффективного применения редукционного метода решения заданного множества задач на двухмерных гранях гиперкуба.

В заключение приведем ряд числовых примеров оценки эффективности редукционного метода по сравнению с традиционным методом решения задач.

Пример 1. Предположим, что $d = 1, c = 3, k = 10$. Имеем

$$S' = 10(1 + 0{,}09) = 10{,}9; \quad (40)$$

$$S = 1 + 9 = 10. \quad (41)$$

Сравнивая (40), (41), получаем, что в этой ситуации редукционный метод оказывается менее эффективным, чем традиционный метод решения задачи сложности S на гиперкубе-прямоугольнике.

Пример 2. Пусть $d = 1, c = 4, k = 10$. Так как $4 > \sqrt{10}$, то в этом случае получаем, что $S' < S$, что доказывает эффективность применения редукционного метода решения задачи по сравнению с традиционным.

Пример 3. Пусть $d = 2, c = 5, k = 6$. В этом случае будем иметь:

$$k + k^{-1} \cdot c^2 = 6 + \frac{25}{6} = \frac{61}{6} = 10\frac{1}{6};$$

$$(1 + c^2) \cdot d^{-1} = 26 \cdot \tfrac{1}{2} = 13.$$

В данной ситуации получаем, что редукционный метод S' оказывается более эффективным традиционного метода решения S, несмотря на то, что вычислительная сложность решения задачи на каждом «малом» подкубе гиперкуба ABCD увеличилась в 2 раза.

Пример 4. Пусть $d = 3, c = 5, k = 6$. Имеем:

$$k + k^{-1} \cdot c^2 = 10\frac{1}{6};$$

$$(1 + c^2) \cdot d^{-1} = \frac{26}{3} = 8\frac{2}{3}.$$

В этом случае $S' > S$, т.е. традиционный метод решения задачи S является более эффективным, чем редукционный метод S' решения задачи на грани гиперкуба ABCD.

4. О сложности декомпозиционных задач на гиперкубах

Оценим вычислительную сложность задач редукции на трехмерных кубах, заданных в виде прямоугольных параллелепипедов со сторонами a; $b = k_1 a$, $k_{1 < 1}$; $c = k_2 a$, $k_{2 < 1}$. На Рис. 2 показан такой гиперкуб: AD = a; $BB_1 = b$; AB = c.

Пусть сторона AD параллелепипеда разделена на три равные части, а сам куб декомпозирован на три одинаковых непересекающихся параллелепипеда. Предположим также, что сложности S, S’ решения задачи традиционным и декомпозиционным методами равны соответственно:

$$S\text{т} = a^2 + b^2 + c^2 = a^2(1 + k^2{}_1 + k^2{}_2);$$

$$S_d = S_1 + S_2 + S_3 = \frac{a^2}{3} + 3a^2 k_1 + 3a^2 k_2,$$

где $S_i = \frac{a^2}{9} + b^2 + c^2$ – сложность декомпозированной задачи на i –ом малом параллелепипеде, i =1,2,3.

Используя эти формулы, получаем:

$$S\text{т} > S_d \leftrightarrow S > S_1 + S_2 + S_3 \leftrightarrow$$

$$a^2(1 + k^2{}_1 + k^2{}_2) > \frac{a^2}{3} + 3a^2 k_1 + 3a^2 k_2.$$

отсюда

$$S\text{т} > S_d \leftrightarrow 1/3 > k^2{}_1 + k^2{}_2. \quad (42)$$

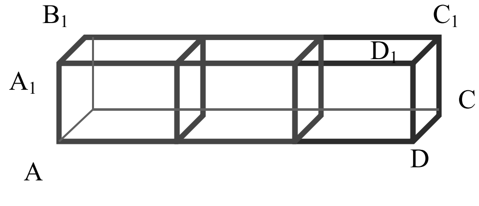

Рис.2. Гиперкуб

Соотношение (42) выполняется, например, в следующих случаях:

k_1 = 1/3; k_2 = 1/3 (${k^2}_1 + {k^2}_2$ = 2/9 < 1/3);

k_1 = 1/2; k_2 = 1/4 (${k^2}_1 + {k^2}_2$ = 5/16 < 1/3).

Приведенные в статье примеры говорят о том, что не любая редукция кубовой структуры OLAP ведет к снижению времени вычислений полной решетки куба, как это кажется на первый взгляд. При определенных сочетаниях шкал размерностей и значений на шкалах время вычислений по редуцированным моделям в сумме может превысить время вычисления по полной модели.

Заключение

Полученные в работе результаты используются в исследованиях по междисциплинарному проекту РФФИ в области медицины, связанному с созданием искусственной поджелудочной железы в виде транспозиционной виртуальной модели. Модель предназначена для замещения нарушенных функций биологической поджелудочной железы по регулированию уровня глюкозы в крови при сахарном диабете. Уровень глюкозы в крови пациента является обобщенной характеристикой влияния большого числа факторов, образующих разветвленную иерархическую структуру, которую можно описать динамической агрегатной сетью OLAP. Современные медицинские приборы для лечения сахарного диабета позволяют обновлять информацию о гликемии каждые 5 минут, поэтому полная агрегатная сеть должна пересчитываться с таким же периодом времени. Задержки в вычислениях недопустимы, так как данные быстро теряют актуальность и могут нарушить ход лечения. Поэтому затронутые в настоящей статье проблемы имеют важное значение для исследования новых способов лечения сахарного диабета.

Таким образом, в статье исследованы математические методы редукции больших гиперкубов многомерных данных на подкубовые компоненты. Введено понятие вычислительной сложности таких методов решения задач анализа данных. Получены критерии успешного применения декомпозиционных методов по сравнению с традиционными нередукционными методами решения задач анализа больших массивов информации.

Литература

1. Макаров И.М., Рахманкулов В.З., Ахрем А.А., Ровкин И.О. Исследование свойств гиперкубовых структур в OLAP-системах // Информационные технологии и вычислительные системы.2005.№2. С. 4-9.
2. Макаров И.М., Рахманкулов В.З., Ахрем А.А., Ровкин И.О. Свойства гиперкубовых структур в OLAP-системах // Труды 1 международной научной конференции САИТ-2005. т.1.-М.:КомКнига. 2005,с.56-60.
3. Чубукова И.А. Data Mining.-М.: Бином. 2008, 382 с.
4. Palaniappan S., Ling C. Clinical Decision Support Using OLAP with Data Mining // Intern. J. of Computer Science and Network Security. 2008. Vol. 8. No9, pp. 290-296.
5. Крупский В.Н., Плиско В.Е. Теория алгоритмов. М.: Академия. 2009, 208 с.
6. Абрамов С.А. Лекции о сложности алгоритмов. М.:МЦНМО. 2009, 256 с.
7. Коробко А.В., Пенькова Т.Г. Метод концептуального OLAP-моделирования на основе формального концептуального анализа // Вестник Сибирского государственного аэрокосмического университета им. акад. М.Ф. Решетнева. Вып.4. Т.30. 2010, С. 74 -78.
8. Yuan C., Lim H., Littman M.L. Most Relevant Explanation: computational complexity and approximation methods. Annals of Mathematics and Artificial Intelligence. 2011, Vol. 61, No.3, pp. 159-183.
9. Gashkov S. B. The arithmetic computational complexity of linear transforms. Moscow University Mathematics Bulletin. 2014. Vol. 6. No.6, pp. 251-257.
10. Ларкин Е.В., Привалов А.Н., Клепиков А.К. Алгоритм последовательных упрощений при оценке временной сложности алгоритмов с применением полумарковских моделей // Известия Тульского государственного университета. Технические науки. 2014, № 9 (2).
11. Ильин В.А., Куркина А.В. Высшая математика. М.: Проспект. 2015, 608 с.
12. Петровский А.Б., Ройзензон Г.В. Многокритериальный выбор с уменьшением размерности пространства признаков: многоэтапная технология ПАКС // Искусственный интеллект и принятие решений. 2012. № 4, С. 88–103.

Рахманкулов Виль Закирович. Заведующий лабораторией ИСА ФИЦ ИУ РАН. Окончил Московский авиационный институт в 1960 году. Доктор технических наук, профессор. Автор 175 печатных работ. Область научных интересов: системный анализ, виртуальное моделирование, интеллектуальный анализ данных, системы бизнес-аналитики, теория управления, компьютерная автоматизация производства. E-mail: vilrakh@mail.ru

Ахрем Андрей Афанасьевич. Старший научный сотрудник ИСА ФИЦ ИУ РАН. Окончил МГУ им. М.В. Ломоносова в 1977 году. Кандидат физико-математических наук. Автор 110 печатных работ. Область научных интересов: математическая теория систем, математическое и виртуальное моделирование сложных технических систем. E-mail: vilrakh@mail.ru

Южанин Кирилл Викторович. Инженер-исследователь ИСА ФИЦ ИУ РАН. Окончил Московский инженерно-физический институт в 2012 году. Автор 1 печатной работы. Область научных интересов: теория распознавания образов, системный анализ. E-mail: vilrakh@mail.ru

On complexity of reduction of multidimensional data models

A.A. Akhrem, V.Z. Rakhmankulov, K.V. Yuzhanin
Institute for System Analysis FRC CSC RAS, Moscow, Russia

Abstract. In this paper decomposition methods for multidimensional data hypercubes of OLAP-systems are studied. The criteria for reducing the computational complexity of the decomposition methods are presented and comparisons have made to the traditional solutions of multidimensional data analysis problems. The examples of application of these criteria in the study of the dynamics of computational complexity changes for the specific types of reduction problems have been considered.
Keywords: hypercube, hypercube multidimensional data, computational complexity, decomposition methods.