

УДК 519.688

*С.В. Аникеев, М.Г. Костиков, А.В. Маркин***МАТЕМАТИЧЕСКАЯ МОДЕЛЬ ОПРЕДЕЛЕНИЯ АКТИВНЫХ УСЛУГ
НА ОСНОВЕ МАТРИЧНОЙ АЛГЕБРЫ**

Предложена параметрическая модель определения активных услуг абонента. Разработан алгоритм реализации модели на основе операций над многомерными матрицами. Сделан вывод о повышении быстродействия выполнения расчета при выполнении операций с использованием распараллеливания вычислений в нескольких потоках.

Ключевые слова: многомерные матрицы, многопоточность, алгоритм расчета.

Введение. При расчете сумм начислений за оказываемые жилищно-коммунальные услуги и потребленные коммунальные ресурсы в информационных системах, оперирующих большими объемами данных, вопросы быстродействия выходят на одно из первых мест. При этом, как показывает практика, на длительность выполнения операций наибольшее влияние оказывают не арифметические действия, как правило, операции умножения и суммирования, а операции логические, связанные с выбором исходных данных, на основе которых выполняется расчет. К таким операциям можно отнести: процесс выбора услуги для начисления, проверка ее активности у каждого отдельного абонента, определение режима потребления, определение значения тарифов и цен на указанную дату. Наибольшего времени выполнения требуют задачи определения действующих режимов потребления и услуг. **Цель работы** – разработать общий алгоритм выбора активных услуг абонента, удовлетворяющий требованиям масштабируемости и гибкости.

Постановка задачи. Коммунальная услуга (КУ) – деятельность исполнителя (поставщика) коммунальных услуг по холодному водоснабжению, горячему водоснабжению, водоотведению, электроснабжению, газоснабжению и отоплению, обеспечивающая комфортные условия проживания граждан в жилых помещениях. Под коммунальным ресурсом (КР) понимается холодная вода, горячая вода, электрическая энергия, газ, бытовой газ в баллонах, тепловая энергия, твердое топливо, используемое для предоставления услуг [1]. Таким образом, КУ – деятельность поставщика по предоставлению абонентам коммунального ресурса. Под расчетом понимается выполнение месячных начислений и

разовых перерасчетов по различным КУ в натуральном и денежном выражении, расчет пени и т.д. Расчет включает в себя, в частности, следующие операции по вычислению для каждого абонента:

- размера платы за пользование коммунальными услугами;
- объема расходуемого ресурса;
- пени за просрочку платежей;
- ведение сальдо в связи с возможными изменениями характеристик за прошлые учетные периоды.

В крупных расчетных центрах, обслуживающих большое количество абонентов, на практике приходится сталкиваться с тем, что один и тот же ресурс могут предоставлять несколько организаций-поставщиков. Каждая организация-поставщик может предоставлять целый спектр разнообразных услуг. Ни поставщики, ни ресурсы не являются константами. С течением времени могут появляться новые и изменяться существующие поставщики и ресурсы как у конкретного абонента, так и в целом в системе. Активной услугой является та услуга, которой абонент пользуется. В общем виде, в каждый конкретный момент времени, у каждого абонента, по каждому КР может как не действовать ни одной услуги, так и одновременно действовать несколько различных КУ. При этом с течением времени активность услуг может меняться. Любой абонент характеризуется некоторой совокупностью признаков (количественных и качественных характеристик), являющихся исходными данными для ведения расчета. Данные признаки привязаны к временной шкале. Таким образом, имеется возможность определения их текущих значений в каждый конкретный момент времени. В связи с большим объемом обрабатываемой информации

(число абонентов может достигать десятков и сотен тысяч) особую актуальность приобретает задача минимизации скорости выполнения всех вышеперечисленных вычислительных операций. Одним из путей решения этой задачи, учитывая групповой характер расчета, является использование операций не отдельно для каждого, а одновременно для множества абонентов. Такой подход позволит более эффективно использовать ресурсы ЭВМ и повысить общую скорость выполнения вычислительных операций.

Задача определения активной услуги сводится к анализу признаков группы абонентов, выделению тех совокупностей признаков, которые влияют на выбор КУ и сопоставления значений данных признаков со значениями, необходимыми для того, чтобы услуга являлась активной.

Теоретическая часть. Каждому абоненту соответствует набор количественных и качественных характеристик. Под количественными характеристиками абонента понимаются характеристики, непосредственно влияющие на размер начисления. К ним относятся число зарегистрированных граждан и животных, размер занимаемой жилой и нежилой площади и т.д. Качественная характеристика – величина, определяющая характер потребления КУ [2]. К ним относятся такие логические (номинальные) характеристики, как режимы работы оборудования, признак наличия или отсутствия горячего водоснабжения и т.д.

Качественные характеристики служат для определения активных услуг и режимов потребления. При этом активность каждой услуги определяется некоторой совокупностью значений качественных характеристик. Таким образом, каждая услуга так же, как и абонент, характеризуется некоторой совокупностью допустимых значений качественных характеристик, за тем исключением, что данные значения постоянны и не изменяются с течением времени. В момент времени t услуга является активной, если совокупность значений качественных характеристик абонента в момент времени t принадлежит множеству допустимых значений услуги.

Любая качественная характеристика может принимать конечное количество целых значений, которые можно представить с помощью вектора. Каждому значению характеристики соответствует некоторый код. Учитывая, что количество таких значений конечно, возможна замена кода характеристики его бинарным представлением. Подобная замена позволит представить качественную характеристику с произвольным количеством значений в виде совокупности не-

скольких характеристик, принимающих только значения 0 и 1 [3]. После выполнения данной операции появляется возможность использовать логические операции для определения активности услуги в каждом временном интервале. Если услуга активна, для нее будет производиться начисление, в противном случае - начисление не производится.

Для повышения быстродействия используется способ организации вычислений, основанный на групповом выполнении вычислительных операций над реляционной моделью данных посредством языка реляционной алгебры. Каждая вышеозначенная вычислительная операция при таком способе должна быть сведена к одному или нескольким выражениям реляционной алгебры, преобразующим набор отношений с исходными данными в отношение, содержащее результат вычислений необходимого значения для каждого абонента. Возможность такого подхода основывается на том факте, что для каждого абонента вычисление требуемой величины можно свести к расчету соответствующих значений одной и той же функции, но с различными значениями аргументов.

Известно, что в современных РСУБД операции над отношениями (таблицами) выполняются посредством высокоуровневого декларативного языка обработки данных. Один из таких языков – язык структурированных запросов SQL.

Одним из недостатков, ограничивающих применение реляционных СУБД для создания масштабируемых систем, является недостаточная гибкость структуры данных. Многомерное представление реляционной модели данных позволяет устранить этот недостаток [4]. Главным отличием многомерного представления данных от табличного представления, используемого в реляционных СУБД, является то, что можно в любой момент времени добавлять новые элементы во множества, определяющие оси многомерного пространства [5]. Тем самым происходит изменение модели данных. Табличное представление реляционной модели не позволяет изменять саму модель при функционировании автоматизированной системы.

Многомерное пространство представления данных можно разделить на три уровня: оси многомерного пространства, декартово произведение множеств значений осей и значения точек пространства [4]. Такая структура позволяет вводить как дополнительные оси и размерности, так и дополнять множества элементов существующих осей в любой момент времени.

Учитывая, что значения качественных характеристик как абонента, так и услуги могут

принимать только значения 0 и 1, имеет смысл перейти к логическим многомерным матрицам.

Собственно логическая матрица представляет собой следующую систему (1):

$$M = \langle V, O, D \rangle, \quad (1)$$

где V – непустое множество истинностных значений;

O – множество матричных операций, определенных на множестве V ;

D – множество выделенных значений такое, что D является собственным подмножеством V , т.е. $D \subset V$, $0 \in V$, но $0 \notin D$ [6].

Примером простейшей логической матрицы может быть следующая матрица (2):

$$M = \langle \{1,0\}, \neg, \supset, \vee, \wedge, \{1\} \rangle \quad (2)$$

$$\text{где } x \wedge y = \min(x, y)$$

$$x \vee y = \max(x, y)$$

$$x \supset y = \max(1 - x, y)$$

$$\neg x = 1 - x.$$

Учитывая то, что необходимо выполнять сравнение элементов, введем в определение логической матрицы еще одну операцию – эквивалентность (XNOR) (3). Таблица истинности данной матрицы представлена в таблице 1.

$$x \equiv y = \max(\min(x, y), \min(1 - x, 1 - y)). \quad (3)$$

Таблица 1 – Таблица истинности для операции эквивалентности

x	y	$x \equiv y$
0	0	1
0	1	0
1	0	0
1	1	1

Таким образом, логическая матрица будет определяться следующим образом (4):

$$M = \langle \{1,0\}, \neg, \supset, \vee, \wedge, \equiv, \{1\} \rangle. \quad (4)$$

Для определения активной услуги будут использоваться две логические матрицы: матрица качественных характеристик абонента LCA и матрица качественных характеристик услуги LCU . Пусть расчетный период разбивается на n равных временных интервалов. Важно при этом выбирать интервалы разбиения таким образом, чтобы в течение одного интервала значения качественных характеристик абонента, а следовательно, и активность услуг не изменялась. Количество столбцов в матрицах LCA и LCU одинаково и соответствует количеству качественных характеристик. Количество строк в матрице LCA равно количеству временных интервалов n . Таким образом, матрица LCA отражает то, какие значения качественных характеристик бы-

ли у абонента в каждый временной интервал. Количество строк в матрице LCU соответствует количеству услуг (5). В каждой строке данной матрицы находятся те значения качественных характеристик, которые должны быть у абонента, чтобы начисление по данной услуге выполнялось.

$$LCA = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \end{bmatrix}, \quad LCU = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix}. \quad (5)$$

Обе матрицы определяются аналогично матрице (2). Выполнив построчное сравнение элементов матриц LCA и LCU (операцию эквивалентности), получим матрицу активности услуг, разложенную по временным интервалам (6).

$$alchar = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}, \quad (6)$$

где $alchar(i, j) = LCA^{<i>} \equiv LCU^{<j>}$.

Из определения коммунальной услуги следует, что КУ – деятельность поставщика по предоставлению абонентам коммунального ресурса [1]. Таким образом, каждая услуга характеризуется совокупностью двух значений – поставщик и ресурс. Все услуги, по которым ведется расчет, можно представить в виде двумерной матрицы BAL (7), где количество строк матрицы соответствует количеству ресурсов, количество столбцов – числу поставщиков. Элемент матрицы $BAL(i, j)$ принимает значение 1, если j поставщик предоставляет i ресурс, и 0 в обратном случае.

$$BAL = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}. \quad (7)$$

Данная матрица является слабоизменяемой и с течением времени остается практически неизменной. Матрица (элементы матрицы) могут изменяться только в случаях:

- добавления новых строк, столбцов (происходит при появлении новых ресурсов или организаций – поставщиков);
 - изменения значения элемента матрицы.
- При этом значение 0 может измениться на зна-

чение 1. Обратная операция невозможна, так как даже в случае прекращения поставщиком деятельности по предоставлению ресурса, при наличии начислений в истории абонента, существует необходимость в хранении информации о том, что деятельность по данному ресурсу в определенный момент времени имела место.

Матрица BAL задает правила, по которым можно определить, предоставляет ли j поставщик деятельность по i ресурсу. Количество элементов в матрице, чьи значения равны 1, соответствует общему количеству услуг в системе. В каждый момент времени абонент характеризуется собственной матрицей активных услуг $ABAL$. Матрица имеет те же размер и размерность, что и матрица BAL , однако элементы матрицы $ABAL$ принимают значение, равное 1, только если данная услуга (совокупность поставщика и ресурса) активна в данный момент. При этом необходимо строгое соблюдение правил, заданных в матрице BAL : если i поставщик не предоставляет j ресурс, то и в матрице $ABAL$ соответствующий элемент равен 0. Таким образом, всегда выполняется тождество (8):

$$BAL(i, j) \cap ABAL(i, j) = ABAL(i, j). \quad (8)$$

Матрица, получившаяся при выполнении поэлементной конъюнкции между матрицами $BAL(i, j)$ и $ABAL(i, j)$, должна быть равна матрице $ABAL(i, j)$.

Пусть расчетный период разбивается на t равных временных интервалов. В каждом из этих интервалов возможно изменение значения активности услуги. Для большинства расчетных задач таким интервалом может служить день. В редких случаях используются часы [3]. В общем случае использование масштабирования значения t позволяет выполнять определение активности услуг с заданной точностью. Добавив к матрице $ABAL$ еще одно измерение, получим трехмерный массив $ABALT$ характеризующий активные услуги абонента в каждый временной интервал t .

Учитывая замену понятия услуги на совокупность поставщик-ресурс, преобразуем двумерную матрицу LCU к трехмерной матрице $LCPR$ размерности i, j, k , где i – количество поставщиков, j – количество ресурсов, k – количество характеристик. Очевидно, что если представить данную матрицу как вектор размерности k , где каждый элемент вектора представляет собой матрицу размерности i, j , то будет справедливо равенство (9), являющееся следствием выражения (8):

$$LCPR_k \cap BAL = LCPR_k. \quad (9)$$

Выполнив построчное сравнение элементов матриц LCU и $LCPR$ (операцию эквивалентности), получим матрицу активности услуг (поставщиков-ресурсов), разложенную по временным интервалам (9). Методологически подход к решению рассматриваемой задачи соотносим с рассмотренным в работе [8].

Таким образом, матрица активностей услуг для данного абонента будет находиться:

$$ABALT(i, j, t) = (LCU(k, t)_t \equiv LCPR(i, j, k)_{i, j})_t \cap BAL(i, j). \quad (10)$$

Практическая часть. Из определения операции матричного умножения и конъюнкции следует, что вычисление всех элементов матрицы может быть выполнено независимо друг от друга. Возможный подход для организации параллельных вычислений состоит в использовании в качестве базовой подзадачи процедуры определения одного элемента результирующей матрицы. Очевидно, что при проведении практических расчетов такое количество сформированных подзадач превышает число имеющихся процессоров и делает неизбежным этап укрупнения базовых задач. В этом плане может оказаться полезной агрегация вычислений уже на шаге выделения базовых подзадач [7]. Ускорение при использовании параллельных алгоритмов можно определить как отношение времени выполнения операций при использовании одного процессора к времени выполнения при использовании многопроцессорной системы (11).

$$S_p = \frac{T_s}{T_p}, \quad (11)$$

где T_s – время выполнения операций без распараллеливания;

T_p – время выполнения операций при использовании алгоритмов с параллельными вычислениями.

Ускорение от использования параллельных вычислений напрямую растет с увеличением количества процессоров (11). Пусть на выполнение всех вычислительных операций на одном процессоре требуется n секунд. Тогда при использовании p процессоров на выполнение расчетов потребуется $\frac{n}{p}$ секунд. Таким образом, ускорение от использования расчета на многомерных матрицах составит (12):

$$S_p = \frac{n}{n/p} = p. \quad (12)$$

На практике данное значение недостижимо.

Помимо вычислительных операций над многомерными матрицами, имеются дополнительные накладные расходы, связанные с операциями, не поддающиеся распараллеливанию. К таким операциям относятся чтение исходных данных в оперативную память, сохранение результатов вычислений, операции распределения вычислений между процессами, объединение полученных результатов на различных процессорах и т.д. Однако использование параллельных вычислений даже для определенной части расчета позволяет сделать вывод о том, что эффективность использования данного алгоритма весьма существенна.

Выводы. Рассмотренный подход позволяет формализовать операцию определения активных услуг в общем виде с помощью матричной алгебры. Полученная модель отвечает требованию масштабируемости и может использоваться для любого количества услуг, поставщиков и ресурсов. Переход к матричной алгебре позволяет распараллеливать процесс ведения расчета. Так как расчет элементов матриц ведется независимо друг от друга, то появляется возможность рассчитывать их отдельно в различных потоках, что приводит к существенному увеличению производительности в современных многопроцессорных системах.

Библиографический список

1. Постановление правительства Российской Фе-

дерации о порядке предоставления коммунальных услуг гражданам от 6 мая 2011 г. № 354.

2. Аникеев С.В., Костиков М.Г., Маркин А.В. Автоматизация настройки расчета за коммунальные услуги в системе «Абонент» // Сборник Математическое и программное обеспечение вычислительных систем. – Рязань: РГРТУ, 2011.

3. Аникеев С.В., Костиков М.Г. О математической модели ведения расчета за жилищно-коммунальные услуги // Сборник Математическое и программное обеспечение вычислительных систем. – Рязань: РГРТУ, 2013.

4. Адаптируемая модель данных на основе многомерного пространства [Электронный ресурс] / Балдин А.В., Елисеев Д.В. – Электрон. журн. – М.: Наука и образование: электронное научно-техническое издание, 2010. Режим доступа: <http://technomag.edu.ru/doc/161410.html>

5. Гарсиа-Молина, Гектор, Ульман, Джефффри, Д., Уидом, Дженифер Г., Системы баз данных. Полный курс. — М.: Издательский дом "Вильямс", 2003. — 1088 с.

6. Карпенко А.С. Развитие многозначной логики. — М.: Издательство ЛКИ, 2010. — 444 с.

7. Немнюгин С., Стесик О. Параллельное программирование для многопроцессорных вычислительных систем — СПб.: БХВ-Петербург, 2002. — 400 с.

8. Миронов В.В., Чураков Е.П. Гарантированная обработка дискретных измерений при слабом контроле помех // Вестник Рязанской государственной радиотехнической академии. 1999. № 6. С. 54-62.