

## Методы синтеза физических структур распределенных баз данных

*Харьковский национальный университет радиоэлектроники*

Разработаны методы многокритериального синтеза физических структур распределенных баз данных по критериям минимума затрат, минимального времени доступа и минимального объема передаваемой информации, реализуемых на компьютерных сетях произвольной топологии. Произведена сравнительная оценка эффективности методов, даны рекомендации по их эффективному использованию.

**Ключевые слова:** распределенные базы данных, проектирование, физическая структура, синтез, многокритериальная оптимизация.

### Введение

Одной из наиболее рациональных форм организации, хранения и обработки информации в территориально распределенных системах транспорта, связи, мониторинга, управления, проектирования являются распределенные базы данных (РБД) [1–3]. В процессе их проектирования используется блочно-иерархический подход, позволяющий декомпозировать проблему их синтеза на комплекс частных задач, таких как, синтез концептуальной, логической и физической структур, определение оптимального количества локальных баз данных (ЛБД), прикрепление пользователей к РБД и т.д. В результате системологического анализа проблемы определены взаимосвязи выделенных основных задач по входным и выходным данным [4]. Большинство задач синтеза, входящих в состав проблемы, имеют комбинаторный характер и высокую размерность. Сложность их совместного решения обуславливает итерационный характер процедур проектирования РБД, гарантирующих получение лишь рациональных проектных решений. При этом затраты на создание и эффективность функционирования РБД во многом зависят от их физической структуры, что определяет особое место задач синтеза физических структур в технологиях проектирования РБД.

### Анализ публикаций по теме исследования

Физическая структура РБД определяет физическую реализацию ее информационной структуры на существующей или создаваемой компьютерной сети. В общем случае в рамках этой задачи определяются количество мест хранения информации (локальных баз), их расположение в сети, распределение информационных ресурсов (ИР) по местам хранения, емкости запоминающих устройств для хранения ИР, а также пропускные способности каналов связи [5].

В подавляющем большинстве публикаций, посвященных проектированию РБД, рассматриваются частные однокритериальные задачи синтеза: синтез логической структуры [6]; определение количества информационных фондов [3]; размещение информационных фондов [3, 7–9]; прикрепление пользователей к РБД [10]. Для решения таких задач используются методы перебора [10–11], Балаша [12–14]), генетические алгоритмы [7–9]).

Основной идеей точных методов решения задачи является полный или усеченный перебор вариантов построения физической структуры РБД. Так, при

использовании комбинаторного метода [10] реализуется полный перебор распределений ИР по возможным местам их размещения на компьютерной сети заданной топологии по стоимостному критерию.

Метод Балаша (аддитивный алгоритм) используется для решения задачи по критерию минимума затрат для случая равенства количества узлов количеству ИР  $n=m$  [12] и позволяет за конечное число шагов найти оптимальный вариант физической структуры РБД. На первом этапе формируется исходное размещение ИР и вычисляются начальные значения рекорда (искусственно введенного параметра, который определяет оптимальность решения). С учетом распределения ИР в РБД, первый уровень дерева решений определяется путем размещения каждого ИР в каждом узле компьютерной сети, т.е.  $x_{ij} = 1, \forall i = j$  и  $x_{ij} = 0, \forall i \neq j, i, j = \overline{1, n}$ . В направлении доминирования вычисляется значение целевой функции задачи и производится проверка ограничений. Следующий уровень ветвления определяется на основании полученных значений целевой функции, при сравнении со значением рекорда. Если при ветвлении было найдено оптимальное решение, а поиск вариантов физической структуры не окончен, то вводится новое значение рекорда. Операция ветвления продолжается до тех пор, пока множество вариантов физической структуры не окажется полным [13]. Суть модификации аддитивного алгоритма, представленной в [14], состоит в том, что на основании ограничения количества копий ИР формируется допустимый вектор-столбец частичного решения, который в дальнейшем зондируется с помощью соответствующих тестов.

Основным недостатком точных методов решения задач синтеза физических структур РБД, ограничивающим их применение для решения практически интересных задач, является их высокая сложность. Альтернативой точным методам в последнее время служат эволюционные методы, построенные на основе генетических алгоритмов [7–8].

Суть этой группы методов применительно к рассматриваемой задаче состоит в следующем. На первом этапе случайным образом генерируется множество исходных вариантов физических структур РБД, рассматриваемых в качестве популяции родителей (хромосом). Полученная популяция хромосом упорядочивается по убыванию значений функции их приспособленности (целевой функции математической модели). К популяции с заданными вероятностями применяются операторы кроссовера (скрещивания) и мутации. Оператор кроссовера скрещивает между собой пары хромосом, создавая популяцию потомков (новых вариантов физических структур РБД). Оператор мутации вносит случайные изменения в хромосомы популяции. Процесс прекращается по достижении популяцией состояния адаптации [9, 15].

Эволюционные методы, построенные на основе генетических алгоритмов, не гарантируют получение оптимального решения. Кроме того, их отличает относительно высокая трудоемкость.

С учетом многокритериального характера задачи синтеза физических структур РБД [16] требуются сравнительный анализ эффективности существующих методов и разработка более эффективных по показателям точности и трудоемкости (сложности) методов решения задач большей размерности.

*Целью статьи* является разработка эффективного по комплексному показателю «точность-сложность» метода решения задачи многокритериального синтеза физических структур распределенных баз данных.

## Математическая модель многокритериальной задачи синтеза физических структур РБД

Общая задача многокритериального синтеза физических структур РБД рассматривается в следующей постановке. Заданы: множество потенциальных территориально рассредоточенных пользователей базы данных  $I = \{i\}$ ,  $i = \overline{1, n}$ , связанных однородной компьютерной сетью  $G = (I, R)$  (где  $R = [r_{ik}]$ ,  $i, k = \overline{1, n}$  – матрица смежности, определяющая множество каналов связи между узлами сети); множество ИР  $J = \{j\}$ ,  $j = \overline{1, m}$ ;  $X = \{x\}$  – множество допустимых реализаций физических структур.

Необходимо определить наилучший вариант физической структуры РБД  $x^o \in X$  (количество ЛБД, распределение ИР по ЛБД  $x = [x_{ij}]$ , размещение ЛБД по узлам сети, объемы запоминающих устройств для хранения ЛБД  $b = [b_i]$ ,  $i = \overline{1, n}$  и пропускные способности каналов связи между узлами сети  $h = [h_{ik}]$ ,  $i, k = \overline{1, n}$ ) с учетом экстремизации выбранных частных критериев, определяющих затраты на синтез РБД, время доступа к РБД, а также объем передаваемых данных.

Рассматривается задача синтеза физической структуры РБД динамического типа с возможностью реинжиниринга компьютерной сети. При размещении ИР в соответствующем узле формируется ЛБД. Информационный обмен и обновление ИР между ЛБД осуществляются с помощью запросов. Состав и количество ИР определяется на этапе концептуального проектирования.

Затраты на реализацию физической структуры РБД определяются соотношением

$$c(x) = \sum_{i=1}^n \sum_{j=1}^m c_{ij}(x) x_{ij} + \sum_{i=1}^n \sum_{j=1}^m c_t (\alpha_{ij} + \beta_{ij}) x_{ij} + \sum_{i=1}^n \sum_{j=1}^m c_t l'_j z_{ij}, \quad (1)$$

где  $c_{ij}(x)$  – затраты на хранение j-го ИР в i-м узле сети;

$c_t$  – затраты на передачу единицы информации;

$\alpha_{ij}$  – суммарный объем запросов к j-му ИР из i-го узла;

$\beta_{ij}$  – суммарный объем ответов на запросы к j-му ИР из i-го узла сети;

$x = [x_{ij}]$  – матрица размещения ИР по узлам сети ( $x_{ij}$  – булева переменная:  $x_{ij} = 1$ , если j-й ИР хранится в i-м в узле сети;  $x_{ij} = 0$  в противном случае);

$l'_j$  – суммарный объем информации, передаваемой при обновлении j-го ИР;

$z = [z_{ij}]$  – матрица обновлений ИР ( $z_{ij}$  – булева переменная:  $z_{ij} = 1$ , если j-й ИР обновляется из i-го узла сети;  $z_{ij} = 0$  в противном случае).

Время доступа к информационным ресурсам РБД определяется соотношением

$$t(x) = \frac{\sum_{i=1}^n \sum_{j=1}^m [t_{ij}^{tr}(x) + t_{ij}^{pr}(x) + t_{ij}^{qp}(x) + t_{ij}^{rp}(x)] x_{ij}}{n \cdot m}, \quad (2)$$

где  $t_{ij}^{tr}(x)$  – время передачи запроса из  $i$ -го узла сети к  $j$ -му ИР;

$t_{ij}^{pr}(x)$  – время ожидания запроса из  $i$ -го узла сети по  $j$ -му ИР;

$t_{ij}^{qp}(x)$  – время обработки запроса из  $i$ -го узла сети по  $j$ -му ИР;

$t_{ij}^{rp}(x)$  – время передачи ответа на запрос из  $i$ -го узла сети к  $j$ -му ИР.

Объем передаваемой информации определяется соотношением

$$v(x) = \sum_{i=1}^n \sum_{j=1}^m (\alpha_{ij} + \beta_{ij}) d_{ij}(x) x_{ij} + \sum_{i=1}^n \sum_{j=1}^m l_j d_{ij}(x) z_{ij}, \quad (3)$$

где  $\alpha_{ij}$  – суммарный объем запросов к  $j$ -му ИР из  $i$ -го узла;

$\beta_{ij}$  – суммарный объем ответов на запросы к  $j$ -му ИР из  $i$ -го узла сети;

$d_{ij}(x)$  – расстояние от места хранения  $j$ -го ИР до  $i$ -го узла сети;

$l$  – объем ИР ( $l = [l_j], j = \overline{1, m}$ ).

Для обеспечения полноты РБД должно выполняться одно из ограничений: в базе хранится более одной копии каждого ИР (тиражирование)  $\sum_{i=1}^n x_{ij} \geq 1, \forall j = 1, n$

или в базе хранится только по одной копии каждого ИР  $\sum_{i=1}^n x_{ij} = 1, \forall j = 1, n$ . При

этом все ИР должны присутствовать в РБД  $\sum_{i=1}^n \sum_{j=1}^m x_{ij} = m$ .

Компромиссное решение многокритериальной задачи  $x^0 \in X$  формально определяется системой трех частных критериев:

$$c(x) \rightarrow \min_{x \in X}, \quad t(x) \rightarrow \min_{x \in X}, \quad v(x) \rightarrow \min_{x \in X}. \quad (4)$$

Используемые показатели качества физических структур (1)–(3) являются разнородными, имеют различные размерность и интервал измерения. Для формирования обобщенного критерия были использованы функции полезности частных критериев вида

$$\xi_i(x) = \left( \frac{k_i(x) - k_i^-}{k_i^+ - k_i^-} \right)^{\beta_i}, \quad (5)$$

где  $k_i(x)$ ,  $k_i^+$ ,  $k_i^-$  – соответственно текущее, наилучшее и наихудшее значения

$i$ -го частного критерия;

$\beta_i$  – параметр, определяющий вид зависимости (5).

Выбор наилучшего компромиссного решения  $x^o \in X$  предлагается производить в рамках кардиналистического подхода с использованием аддитивно-мультипликативной функции общей полезности:

$$P(x) = \delta \cdot \sum_{i=1}^3 \eta_i \xi_i(x) + (1 - \delta) \prod_{i=1}^3 [\xi_i(x)]^{\eta_i}, \quad (6)$$

где  $\eta_i$ ,  $i = \overline{1,3}$  – коэффициенты важности частного критерия  $k_i(x)$ ,  $\eta_i \geq 0$ ,

$$\sum_{i=1}^3 \eta_i = 1;$$

$\xi_i(x)$  – функция полезности частного критерия  $k_i(x)$ ;

$\delta$  – параметр модели, определяющий конкретный вид схемы компромиссов,  $0 \leq \delta \leq 1$ : при  $\delta = 1$  функция (6) принимает форму аддитивной, а при  $\delta = 0$  – форму мультипликативной функции общей полезности.

Для выбора наилучшего компромиссного варианта физической структуры РБД требуется решить задачу оптимизации вида

$$x^o = \arg \max_{x \in X} P(x). \quad (7)$$

### Методы решения задачи

Проектируемые РБД существенно различаются по количествам потенциальных пользователей, узлов компьютерных сетей, в которых они функционируют, ИР, локальных баз. Это требует использования в системах автоматизации проектирования множества методов синтеза физических структур, различающихся по точности и временной сложности. Наличие базы методов позволит выбирать лучший из них, исходя из требуемой точности решения, а также имеющихся временных и вычислительных ресурсов.

Дискретный характер решаемой задачи определяется ее исходными данными: множеством потенциальных территориально рассредоточенных пользователей РБД  $I = \{i\}$ ,  $i = \overline{1, n}$ ; структурой и топологией компьютерной сети  $G = (I, R)$  (где  $R = [r_{ij}]$ ,  $i, j = \overline{1, n}$  – матрица смежности, определяющая множество каналов связи между узлами сети); множеством ИР  $J = \{j\}$ ,  $j = \overline{1, m}$ ; множеством допустимых реализаций физических структур  $X = \{x\}$ .

В процессе оптимизации *методом перебора сочетаний* при заданных  $I, G, J$  необходимо определить размещение ИР в РБД  $x^o$ , а также значения пропускных способностей каналов связи  $h = [h_{ij}]$ ,  $i, j = \overline{1, n}$  и объем ЗУ узлов  $b = [b_i]$ ,  $i = \overline{1, n}$ . С учетом этого при решении задачи формируется множество возможных размещений ИР по РБД  $X = \{x\}$ , которое представляет собой множество соче-

таний  $C_n^m$  (где  $n$  – количество узлов в компьютерной сети,  $m$  – количество ИР в РБД). Метод основан на упорядочении области допустимых решений  $X = \{x\}$  таким образом, что оптимальное решение задачи находится путем направленного перебора по количеству мест размещения локальных баз и наборам ИР в них [17–19]. Поиск наилучшего варианта размещения ИР в РБД идет путем последовательного увеличения количества мест размещения ИР (локальных баз) до достижения экстремума функции цели.

При решении задачи *методом Балаша* необходимо определить нижнюю границу целевой функции  $\min P(x)$  (недопустимое значение критерия (6)), которая соответствует наихудшему варианту. Значение критерия (1) вычисляется для случая размещения всех ИР в каждом узле компьютерной сети ( $\max c(x)$ ). Наихудшие значения критериев (2)-(3) определяются путем задания максимально допустимых уровней времени ( $\max t(x)$ ) доступа и объема передаваемых данных ( $\max v(x)$ ).

С учетом распределения ИР в РБД первый уровень дерева решений определяется путем размещения одного ИР в одном узле компьютерной сети. Для решаемой задачи количество ветвей дерева равно количеству уровней ветвлений, т.е.  $n = m$ . На каждом последующем уровне ветвления происходит доразмещение ИР до тех пор, пока не будут распределены все ИР. Выбор направления ветвления в дереве решений осуществляется на основании максимального значения выражения (6). В результате работы алгоритма получаются  $n$  максимальных значений обобщенного критерия  $P(x)$ , среди которых необходимо выбрать максимальное, соответствующее оптимальной физической структуре РБД.

*Эволюционный метод* решения задачи, реализуемый с помощью генетического алгоритма, представляет комбинацию переборного и локально-градиентного методов: механизмы кроссовера и мутации реализуют переборную часть метода, а отбор лучших хромосом – градиентный подъем. Формирование начального набора хромосом осуществляется путем генерации случайной последовательности нулей и единиц. Количество генов в хромосоме соответствует количеству узлов компьютерной сети  $n$ , а аллели генов принимают значения 0 или 1, в зависимости от наличия или отсутствия ИР в данном узле сети. Локус определяет номер узла в компьютерной сети. Проверка веса каждой хромосомы  $u$  осуществляется путем анализа сгенерированного битового массива с учетом задаваемого количества ИР ( $u = m$ ). При невыполнении условия  $u = m$  случайно сгенерированная хромосома отбрасывается и генерация продолжается до выполнения этого условия.

Технологию решения задачи синтеза физической структуры РБД с использованием этого метода можно представить последовательностью из трех этапов: на первом этапе происходит генерация размещений ИР по ЛБД (получение физической структуры  $X = [x_{ij}]$ ); на втором этапе, с учетом применения основных генетических операторов, формируются варианты размещения ИР по ЛБД; на третьем этапе происходит оценка эффективности сформированных вариантов (согласно функции приспособленности хромосом (6)) и их ранжирование. Результатом решения является физическая структура РБД с наибольшим значением обобщенного критерия (6).

Суть предлагаемой модификации *метода направленного перебора* вариантов состоит в использовании схемы покоординатной оптимизации для выбора наилучшего размещения заданного количества ИР  $m$  на заданном количестве узлов компьютерной сети  $n$  [17]. Покоординатная оптимизация состоит в улучшении некоторого начального размещения ИР в узлах компьютерной сети путем после-

довательного перемещения одного из ресурсов при фиксированном размещении  $m-1$  остальных. При наличии достаточных вычислительных ресурсов точность этого метода может быть повышена путем использования мультистарта (многопроходности).

### Анализ эффективности методов решения задачи

Сравнительная оценка эффективности рассмотренных методов производилась экспериментальным путем на основании оценок времени решения контрольных задач и точности получаемых решений. В ходе исследования на персональном компьютере с тактовой частотой процессора 4,5 ГГц было решено 150 задач синтеза физической структуры РБД при различных исходных данных и структурах компьютерной сети.

Для комбинаторных методов определялась их временная сложность (время решения задачи как функция от размерности задачи: количества узлов компьютерной сети  $n$  и количества размещаемых информационных ресурсов  $m$ ). Ввиду быстрого увеличения количества способов размещения  $m$  информационных ресурсов на  $n$  узлах компьютерной сети (сочетаний  $C_n^m$ ) с увеличением  $m$  и  $n$  время решения задачи также резко возрастает (рис. 1). Этот метод имеет неполиномиальную временную сложность, его целесообразно применять для  $n \leq 20$  и  $m \leq 20$ . Но при относительно небольших значениях  $m$  и  $n$  его временная сложность с удовлетворительной точностью (с достоверностями  $R=0,99$  и  $R=0,98$ ) аппроксимируется полиномами второй степени.

Метод Балаша позволяет за приемлемое время получать решения задач большей размерности (до  $n = 30$  и  $m = 30$ ). По сравнению с методом перебора сочетаний для  $m = 30$  и  $n = 30$  уменьшение времени счета для метода Балаша составляет порядка 67%.

Оценка эффективности эволюционного метода и метода покоординатной оптимизации проводилась по показателям точности (максимальной погрешности  $\sigma_{max}$ ) и времени решения задач (рис. 2–4).

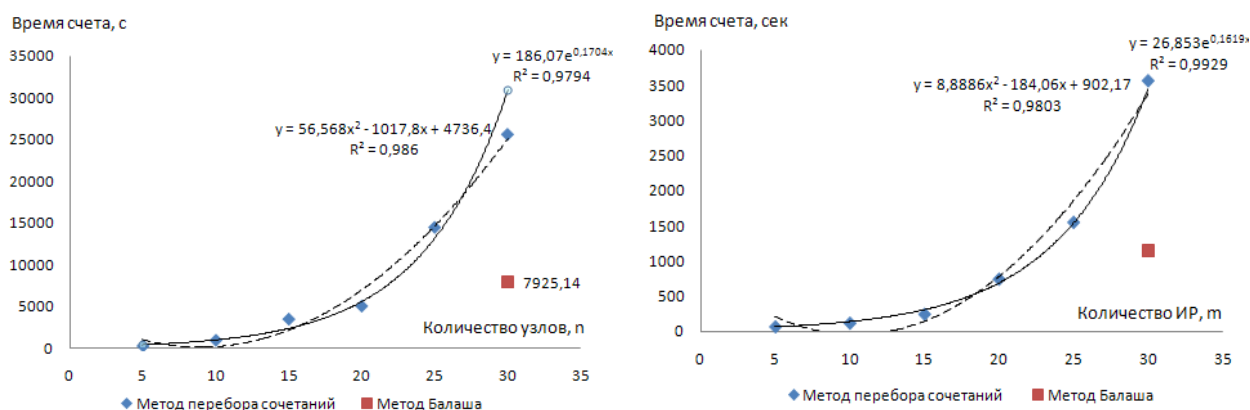


Рис. 1. Аппроксимация временной сложности комбинаторных методов

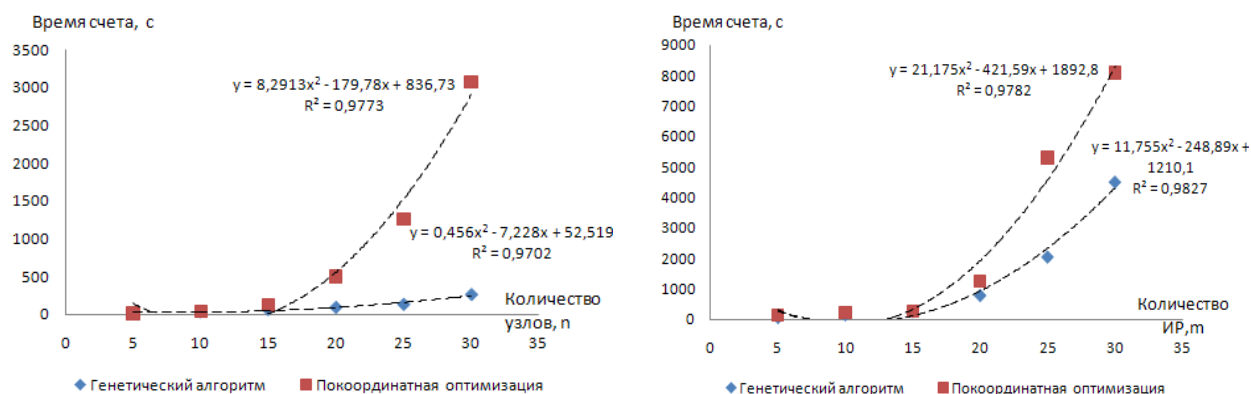


Рис. 2. Аппроксимация временной сложности эвристических методов

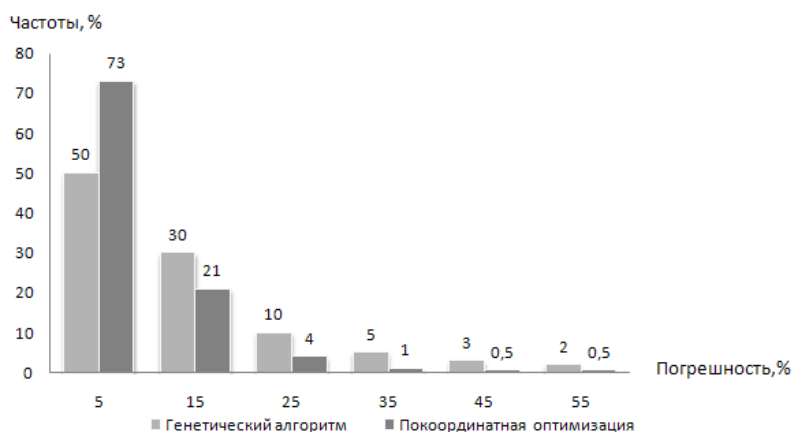


Рис. 3. Распределение частот  $\sigma_{max}$  для приближенных методов

Эксперименты проводились путем изменения количества узлов компьютерной сети (при фиксированном количестве ИР в РБД  $m = 30$ ) и информационных ресурсов (при фиксированном количестве узлов компьютерной сети  $n = 30$ ). Временные сложности эволюционного метода и метода покоординатной оптимизации с удовлетворительной точностью (с достоверностями  $R=0,97$  и  $R=0,98$ ) аппроксимируются полиномами второй степени.

При решении задач многокритериальной синтеза физической структуры РБД с  $n \times m \leq 100$  целесообразным оказалось применение комбинаторных методов. При  $n \times m > 100 \div 120$  ввиду значительного времени решения задачи применение комбинаторных методов становится нерациональным. Для задач такой и большей размерности целесообразным оказывается применение приближенных методов. При значительно меньшем времени решения задачи методом покоординатной оптимизации средняя относительная погрешность относительно глобального оптимума составила  $\bar{\sigma} = 6\%$ , а максимальная относительная погрешность  $\sigma_{max} = 9,8\%$ .

Для эволюционного метода в ходе экспериментального исследования было установлено, что при малых размерах хромосом (количество узлов  $n \leq 5$ ), когда направления экстремума покрываются быстрее (процесс ранжирования хромосом задает алгоритму направление исследования) алгоритм может застрять в локальном оптимуме. Это подтверждают оценки метода генетических алгоритмов на точность решения, проводимые при сравнении с эталонным (полученным комби-



наторным методом перебора сочетаний). В качестве оценок использовались средняя и максимальная относительные погрешности решения, которые составили 22,1 и 40,5% соответственно.

При анализе временной сложности получаемых решений аддитивным алгоритмом при  $n = m$  время счета алгоритма меньше времени счета метода перебора сочетаний примерно в 3 раза (3555,35 и 1147,78 с соответственно), что обеспечивает эффективность его применения для фиксированных  $n = m$ . При размерности задач  $n \times m \geq 150$  и однокритериальной их постановке время решения методом Балаша сравнимо со временем решения методом перебора сочетаний.

Применение метода эволюционного поиска оправданно при необходимости получения приближенного решения задачи синтеза физических структур РБД с  $n \times m > 100$  за короткое время, а также использования совокупности разнородных и противоречивых критериев.

При реализации генетических алгоритмов без использования кода Грея для кодирования хромосом ухудшения качества решений не выявлено. Введение дополнительной проверки веса хромосомы с учетом общих временных затрат на поиск многокритериального решения оказалось несущественным. При решении многокритериальных задач и увеличении количества эпох  $t$  (100,100000) время решения возросло в 10 раз. При этом количество хромосом в популяции / на время решения задачи существенного влияния не оказывало.

Сравнительная оценка эволюционного метода и метода покоординатной оптимизации проводилась для заданных значений времени поиска решения. При этом определялись средняя и максимальная относительные погрешности получаемых решений (рис. 4).

На основании этих оценок можно сделать вывод, что для размерности задачи

$n \leq 20, m \leq 20$  метод покоординатной оптимизации является более эффективным по комплексному показателю «точность-сложность». Применение метода генетических алгоритмов эффективно при размерности задачи синтеза физической

структуры при  $n > 20, m > 20$ .

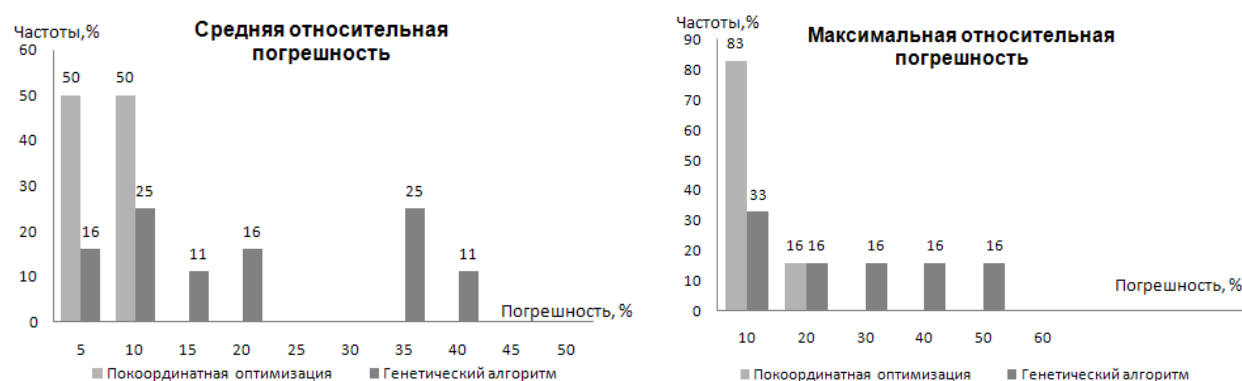


Рис. 4. Оценка эффективности приближенных методов

### Выводы

На основе анализа постановки общей многокритериальной задачи синтеза физических структур РБД предложен набор методов ее решения для использова-

ния в системе автоматизации проектирования. Выполнена экспериментальная оценка эффективности предложенных методов, даны рекомендации по их эффективному использованию. Предложенные методы существенно различаются по времени решения задачи и точности. Использование в системе проектирования множества методов позволит выбирать наиболее рациональный из них, исходя из размерности решаемой задачи, требуемой точности решения и имеющихся временных и вычислительных ресурсов.

Направлением дальнейших исследований может быть разработка методов формирования подмножеств недоминируемых альтернатив и выбора параметров моделей многофакторного оценивания вариантов физических структур РБД.

### Список литературы

1. Арсеньев В.П. Интегрированные распределенные базы данных [Текст] / В.П. Арсеньев. – СПб.: Изд.-полигр. центр СПбГЭТУ (ЛЭТИ), 2004. – 498 с.
2. Коннолли Т. Базы данных: проектирование, реализация, сопровождение. Теория и практика [Текст] / Т.Коннолли, К. Бегг; [пер. с англ. В.А. Иванов]. – М.: Вильямс, 2003. – 720 с.
3. Теоретические основы проектирования оптимальных структур распределенных баз данных [Текст] / [Кульба В.В., Ковалевский С.С., Косяченко С.А., Сиротюк В.О.]. – М.: Синтег, 1999. – 660 с.
4. Бескоровайный, В.В. Системологический анализ проблемы автоматизированного проектирования распределенных баз данных [Текст] / В.В. Бескоровайный, В.В. Евсеев, О.С. Ульянова // Вестн. Херсон. нац. техн. ун-та. – 2010. – № 2 (38). – С.120 - 125.
5. Корпоративные хранилища данных. Планирование. Разработка. Реализация [Текст]: [под. ред. В.К. Фролова]. – М.: Вильямс, 2004. – 658 с.
6. Силин, А.В. Методы и модели проектирования структур территориально-распределенных баз данных [Текст] / А.В. Силин, В.И. Воробьев, Г.И. Ревунков // Деп. рук. ВИНТИ № 3282-00В. – 2005. – С.21 - 25.
7. Зарецкий, К.А. Модель оптимального размещения информационных ресурсов в неоднородной компьютерной сети [Текст] / К.А. Зарецкий, В.И. Мейкшан // Вестн. СибГУТИ. – 2006. – № 1. – С. 18 - 21.
8. Лаздынь, С.В. Оптимизация распределенных корпоративных информационных сетей с использованием генетических алгоритмов и объектного моделирования [Текст] / С.В. Лаздынь, С.Ю. Землянская // Наук. пр. ДонНТУ. – 2009. – № 147. – С.83 - 95.
9. Новосельский, В.Б. Методы автоматизации проектирования распределенной реляционной базы данных [Текст] / В.Б. Новосельский // Программные продукты и системы. – 2008. – № 3. – С. 45 - 48.
10. Цегелик, Г.Г. Системы распределенных баз данных [Текст] / Г.Г. Цегелик. – Львов: Світ, 1990. – 168 с.
11. Колесников, Д.Г. Построение модели оптимального размещения копий информационных файлов по узлам сети ЭВМ [Текст] / Д.Г. Колесников, Е.Н. Остроух // Материалы конф. “Компьютерные технологии в инженерной и управленческой деятельности”. Тезисы докладов. – Таганрог: ТГРУ, 1998. – С.105 - 107.
12. Евсеев, В.В. Модели и алгоритмы размещения информационных фондов в интегрированных АСУ городом [Текст] / В.В. Евсеев, В.И. Барский // Современные методы и средства создания и развития интегрированных АСУ городом. – М.: Мир, 1986. – С.114 - 115.

13. Кофман, А. Методы и модели исследования операций [Текст] / А. Кофман, А. Анри-Лабурдер. – М.: Мир, 1977. – 423 с.
14. Синтез информационно-вычислительного обеспечения распределенных АСПИ. [Текст] / [Петров Э.Г., Аннамухамедов, О.Б., Евсеев В.В. и др.]; под ред. Э.Г. Петрова. – Ашхабад: Ылым, 1991. – ч.2. – 168 с.
15. Кулагин, О.А. Многокритериальная модель проектирования схемы размещения данных в информационной сети [Текст] / О.А. Кулагин // Приборостроение. – 1997. – № 9. – С. 12 - 15.
16. Лаздынь С.В. Оптимизация распределенных баз данных использованием генетических алгоритмов [Текст] / С.В.Лаздынь, А.О.Телятников // Вестн. Херсон. нац. техн. ун-та. – 2004. – № 1(19). – С. 236 - 239.
17. Бескоровайный В.В. Математические модели многокритериального синтеза физических структур распределенных баз данных [Текст] / В.В. Бескоровайный, О.С. Ульянова // Восточно-европейский журнал передовых технологий. – 2010. – № 4. – С.44 – 48.
18. Бескоровайный, В.В. Модификация метода направленного перебора для синтеза топологии систем с радиально-узловыми структурами [Текст] / В.В. Бескоровайный // Автоматизированные системы управления и приборы автоматизации. – 2003. – Вып. 123. – С. 110 - 116.
19. Бескоровайный, В.В. Модификация методов направленного перебора для оптимизации топологии систем с радиально-узловыми структурами [Текст] / В.В. Бескоровайный, Е.В. Соболева // Системи обробки інформації. – 2008. – № 5 (72). – С. 25 - 30.

**Рецензент:** д-р техн. наук, проф. С.Ю.Шабанов-Кушнаренко, Харьковский национальный университет радиоэлектроники, Харьков

Поступила в редакцию 09.06.10

## **Методи синтезу фізичних структур розподілених баз даних**

Розроблено методи багатокритеріального синтезу фізичних структур розподілених баз даних за критеріями мінімуму витрат, мінімального часу доступу і мінімального обсягу інформації, яка передається, що реалізуються на комп'ютерних мережах довільної топології. Зроблено порівняльне оцінювання ефективності методів, наведено рекомендації щодо їх ефективного використання.

**Ключові слова:** розподілені бази даних, проектування, фізична структура, синтез, багатокритеріальна оптимізація.

## **Methods of synthesis of physical structures of the distributed databases**

The methods for multi-criteria synthesis of physical structures of the distributed databases including on the criteria of minimum cost, minimum access time and the minimum amount of information transmitted, solved on the computer networks of arbitrary topology are developed. A comparative estimation of the effectiveness of the methods, recommendations for their efficiency of use is made.

**Keywords:** the distributed databases, design, the physical structure, synthesis, multicriteria optimization.