

ИНФОРМАТИКА, ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА И УПРАВЛЕНИЕ

УДК 519.2:004.421.5:004.7

**В. Н. ЗАДОРЖНЫЙ
Т. Р. ЗАХАРЕНКОВА**Омский государственный
технический университет

МЕТОДЫ ПЛАНИРОВАНИЯ ИМИТАЦИОННЫХ ЭКСПЕРИМЕНТОВ ПРИ МОДЕЛИРОВАНИИ ФРАКТАЛЬНЫХ ОЧЕРЕДЕЙ

Рассматриваются проблемы корректной организации имитационных экспериментов при расчете фрактальных очередей. Фрактальные системы с очередями описываются асимптотически степенными законами распределения интервалов поступления и времени обслуживания заявок и являются адекватными математическими моделями сетевых устройств телекоммуникационных систем с фрактальным трафиком. Выявляются особенности расчета фрактальных очередей. Разрабатываются рекомендации по организации последовательных и многократных «параллельных» прогонов модели. Выводятся формулы для контроля точности получаемых результатов.

Ключевые слова: моделирование, генераторы случайных чисел, системы с очередями, распределения с тяжелыми хвостами.

1. Введение. При описании и анализе фрактального (самоподобного) трафика применяются такие математические понятия, как самоподобный случайный процесс, долговременная зависимость, распределения с тяжелыми хвостами (РТХ) [1, 2]. При проектировании сетевых устройств на системном уровне наиболее подходящими математическими моделями являются системы с очередями [3, 4]. Фрак-

тальными системами мы называем такие системы класса $G|G|n|m$ [3], в которых интервалы поступления заявок и/или время их обслуживания описываются распределениями с тяжелыми хвостами (РТХ). При этом мы полагаем, что РТХ определяется такой функцией распределения вероятностей (ф.р.), которая имеет асимптотически степенной хвост, конечное математическое ожидание (м.о.) и бесконеч-

ную дисперсию. Коэффициент ρ загрузки рассматриваемых систем не превосходит единицы:

$$\rho = \frac{\bar{x}}{n\bar{t}} \leq 1,$$

где \bar{x} — среднее время обслуживания заявки, \bar{t} — среднее время между приходами заявок, n — число каналов в системе.

Такие соотношения параметров РТХ наиболее актуальны в приложениях теории фрактальных очередей к проектированию сетевых устройств телекоммуникационных систем. Системы GI|GI|n|m, задаваемые ф.р. с экспоненциальными хвостами, будем называть классическими системами.

Типичными представителями фрактальных систем с очередями являются системы Pa/M/n/m, M/Pa/n/m и Pa/Pa/n/m. Здесь символом Pa обозначено распределение Парето (РП). Если размер m буфера конечный, то требуется при данном m определить вероятность P потери заявки (прямая задача) или найти наименьший размер m , гарантирующий, что вероятность потери будет не больше P (обратная задача). При $m = \infty$ интерес представляет средняя длина очереди L или среднее время ожидания \bar{w} . Сокращенно РП с параметрами K, α будем обозначать как Pa(K, α). Его ф.р. $F(t)$ имеет вид

$$F(t) = 1 - \left(\frac{K}{t}\right)^\alpha, \quad \alpha > 0, \quad K > 0, \quad t \geq K, \quad (1)$$

где α — параметр формы, K — наименьшее значение случайной величины (с.в.) и, одновременно, масштабный параметр. Типичный для фрактального трафика диапазон значений α определяется неравенством $1 < \alpha \leq 2$. Из (1) нетрудно найти, что при таких α РП имеет конечное м.о., равное $\alpha K / (\alpha - 1)$, и бесконечную дисперсию.

Поскольку исследование фрактальных систем с очередями аналитическими методами затруднено, для их расчета широко используется имитационное моделирование (ИМ). В [5] перечислены следующие основные проблемы ИМ фрактальных очередей:

- 1) *проблема корректной реализации РТХ*, выявленная в [6];
- 2) *проблема медленной сходимости оценок*;
- 3) *проблема длительных переходных процессов в фрактальных системах*;
- 4) *проблема больших затрат времени на ИМ фрактальных систем*.

Эффективное решение первой проблемы, найденное в работах [7, 8], состоит в достаточно простом предварительном преобразовании равномерно распределенной базовой с.в. (БСВ), используемой для реализации всех прочих с.в. Это преобразование (метод ARAND) обеспечивает сохранение количества значащих цифр БСВ при любом выпадающем значении БСВ, сколь бы близком к нулю оно ни оказалось. Тем самым обеспечивается достаточно точная реализация РТХ, осуществляемая методом обратного преобразования хвоста распределения.

Оставшиеся три проблемы остаются актуальными и требуют тщательного планирования имитационных экспериментов с фрактальными очередями. В статье выявляется ряд принципиальных отличий планирования таких экспериментов от планирования экспериментов с классическими очередями и выводятся математические соотношения, позволяющие корректно определять длительность (объем) экспериментов для достижения заданной точности результатов.

2. Построение доверительных интервалов в зависимых испытаниях. В общем случае при моделировании очередей для расчета оценки ξ м.о. $\bar{\xi}$ («показателя») используется выборка ξ_1, \dots, ξ_N зависимых реализаций с.в. ξ на выходе имитационной модели. Оценка ξ рассчитывается как выборочное среднее:

$$\xi = \frac{1}{N} \sum_{i=1}^N \xi_i. \quad (2)$$

Корреляции между реализациями ξ_1, \dots, ξ_N усредняемой с.в. ξ можно связать с дисперсией $D(\xi)$ оценки ξ следующим образом:

$$\begin{aligned} D(\xi) &= D\left(\frac{1}{N} \sum_{i=1}^N \xi_i\right) = \frac{1}{N^2} \left[M\left(\sum_{i=1}^N \xi_i\right)^2 - M^2\left(\sum_{i=1}^N \xi_i\right) \right] = \\ &= \frac{1}{N} \left[\sigma^2 + \frac{2}{N} \sum_{i=1}^{N-1} \sum_{j=i+1}^N r_{ij} \sigma^2 \right] = \frac{\sigma^2}{N} (1 + 2R_N), \quad (3) \end{aligned}$$

где σ^2 — дисперсия выборочного значения ξ_i (не зависит от i , так как мы рассматриваем стационарную последовательность реализаций ξ_1, \dots, ξ_N);

$$r_{ij} = \text{corr}(\xi_i, \xi_j) = \frac{M(\xi_i \xi_j) - M(\xi_i)M(\xi_j)}{\sqrt{D(\xi_i)D(\xi_j)}} =$$

$= \frac{M(\xi_i \xi_j) - M^2(\xi)}{\sigma^2}$ — коэффициент корреляции элементов ξ_i, ξ_j (зависит только от расстояния $s = |j - i|$ между элементами ξ_i, ξ_j);

$R_N = \frac{1}{N} \sum_{j=1}^{N-1} \sum_{i=j+1}^N r_{ij}$ — «коэффициент последствия» (выражает степень «прямого» влияния реализованных элементов выборки ξ_1, \dots, ξ_N на элементы ее продолжения).

Обозначая r_{ij} через $r(s)$, где $s = |j - i|$, выражение коэффициента R_N можно записать в виде

$$R_N = \frac{1}{N} \sum_{j=1}^{N-1} \sum_{i=j+1}^N r_{ij} = \frac{1}{N} \sum_{s=1}^{N-1} (N - s)r(s), \quad (4)$$

где $r(s)$ — коэффициент корреляции пары элементов выборки, отстоящих друг от друга на s шагов; $(N - s)$ — число таких пар в выборке.

Учитывая, что $|r(s)| \leq 1$ и $D(\xi) \geq 0$, из (3) нетрудно вывести, что коэффициент $(1 + 2R_N)$ повышения дисперсии лежит в пределах от 0 до N . Как правило, показатель $\bar{\xi}$ при моделировании очередей имеет такой конкретный смысл (как, например, среднее время \bar{w} ожидания заявки), что все коэффициенты корреляции $r(s)$ оказываются положительными, и, соответственно, коэффициент повышения дисперсии $1 + 2R_N$ принимает значение, большее единицы. Поэтому, согласно (3), корреляции элементов выборки приводят к увеличению дисперсии $D(\xi) = \sigma_\xi^2$ и расширению доверительного интервала (по сравнению с независимыми ξ_i), построенного по правилу трех сигм:

$$\bar{\xi} \pm 3\sigma_\xi, \quad (5)$$

где $\sigma_\xi = \sqrt{D(\xi)} = \sigma \sqrt{(1 + 2R_N) / N}$,

σ — среднеквадратичное отклонение (с.к.о.) элементов ξ_i выборки,

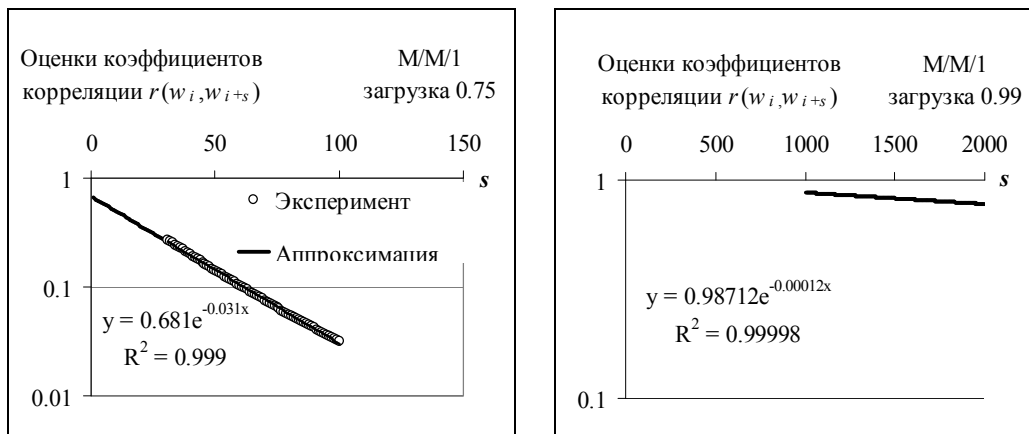


Рис. 1. Корреляция величин w_i и w_{i+s} в системе М/М/1 в зависимости от s

$\sqrt{(1 + 2R_N)}$ — коэффициент расширения доверительного интервала.

Расчет доверительных интервалов (5) при больших объемах N выборок значительно упрощается благодаря тому, что асимптотика коэффициентов последствия R_N при расчете классических и фрактальных очередей не отличается разнообразием: для классических очередей она определяется асимптотически экспоненциальным убыванием коэффициентов корреляции $r(s)$ с ростом s , для фрактальных очередей — асимптотически степенным убыванием $r(s)$. Это позволяет по выборкам умеренного объема находить приближенные аналитические выражения, пригодные для построения доверительных интервалов как функций от N и, соответственно, планировать при последовательном прогоне [9] модели такую длину выборки, которая обеспечивает необходимую точность результатов.

3. Доверительные интервалы в расчетах классических очередей. В качестве примера построения доверительных интервалов при моделировании классических очередей рассмотрим расчет стационарного среднего времени ожидания \bar{w} заявок в системе М/М/1 с интенсивностью входящего потока $\lambda = 1$ при коэффициентах загрузки $\rho = 0,75$ и $\rho = 0,99$.

Статистическую оценку $\hat{r}(s)$ коэффициента корреляции $r(s)$ между временем w_i ожидания i -й заявки и временем w_{i+s} ожидания $(i+s)$ -й заявки можно рассчитать по стационарной выборке w_1, \dots, w_N достаточно большого объема N , применяя формулу

$$\hat{r}(s) = \frac{\hat{M}(w_i w_{i+s}) - \hat{M}^2(w)}{\hat{D}(w)}, \quad (6)$$

$$\text{где } \hat{M}(w_i w_{i+s}) = \frac{1}{N-s} \sum_{i=1}^{N-s} w_i w_{i+s}, \quad \hat{M}(w) = \frac{1}{N} \sum_{i=1}^N w_i,$$

$$\hat{D}(w) = \hat{M}(w^2) - \hat{M}^2(w), \quad \hat{M}(w^2) = \frac{1}{N} \sum_{i=1}^N w_i^2.$$

Оценки $\hat{r}(s)$ коэффициентов корреляции $r(s) = r(w_i, w_{i+s})$, рассчитанные по формуле (6) при длине прогона модели $N = 10$ млн заявок, показаны в виде графиков на построенном средствами Excel рис. 1. Слева показана зависимость $r(s)$ при коэффициенте загрузки $\rho = 0,75$, справа — при $\rho = 0,99$.

В общем случае в классических системах коэффициенты корреляции $r(s)$ между двумя сдвинутыми на s шагов элементами выборки w_1, \dots, w_N описываются выражением

$$r(s) \sim ae^{-bs}. \quad (7)$$

Отклонения функции $r(s)$ от экспоненты (7) быстро уменьшаются с ростом s и на практике ими можно пренебречь. Учитывая (4) и (7), коэффициент последствия R_N можно выразить следующим образом:

$$\begin{aligned} R_N &= \frac{1}{N} \sum_{s=1}^{N-1} (N-s)r(s) \sim \frac{1}{N} \sum_{s=1}^{N-1} (N-s)ae^{-bs} \sim \\ &\sim \frac{a}{N} \int_1^{N-1} (N-s)e^{-bs} = \dots \\ &= \frac{a}{Nb} \left\{ e^{-b(N-1)} \left(\frac{1}{b} - 1 \right) + e^{-b} \left(N - 1 - \frac{1}{b} \right) \right\} \rightarrow \frac{a}{b} e^{-b}. \end{aligned} \quad (8)$$

Для прогонов длины $N = 100$ тыс. заявок и более при тех a, b , что представлены на рис. 1, можно использовать последнее (предельное) выражение из (8). Соответственно, при $\rho = 0,75$ получаем $R_N \approx 22$ и доверительный интервал

$$\bar{w} = \hat{w} \pm 3\sigma\sqrt{(1 + 2R_N)} / \sqrt{N} = \hat{w} \pm \frac{3\hat{\sigma}}{\sqrt{N}} \cdot 6,7, \quad (9)$$

где \hat{w} и $\hat{\sigma}$ — выборочные оценки м.о. и с.к.о. времени ожидания заявки в стационарном режиме. Аналогично для $\rho = 0,75$ по формуле (8) находим, что при больших N коэффициент $R_N \approx 8000$ и доверительный интервал имеет вид

$$\bar{w} = \hat{w} \pm \frac{3\hat{\sigma}}{\sqrt{N}} \cdot 130. \quad (10)$$

Выражения (9) и (10) доверительных интервалов подтверждаются дополнительными проверками.

4. Доверительные интервалы в расчете фрактальных очередей. При расчете фрактальных систем с очередями коэффициенты $r(s)$ корреляции элементов ξ_i обрабатываемых выборок являются асимптотически степенными функциями от s . Это приводит к существенному отличию методов планирования экспериментов с фрактальными очередями по сравнению со случаем классических систем, где соответствующая асимптотика экспоненциальная. На рис. 2 показан график функции $r(s)$, рассчитанной путем ИМ системы Ра/Ра/1 в стационарном режиме ее функционирования, наступающем после прохождения через систему 40 млн заявок. В качестве показателя ξ_i выступал индикатор γ_i отказа i -й заявки $\gamma_i \in \{0, 1\}$, среднее значение $\bar{\gamma}$ которого в стационар-

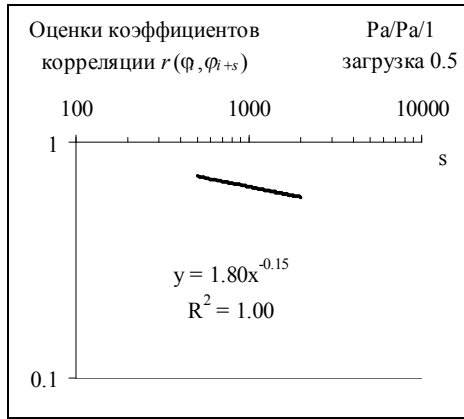


Рис. 2. Асимптотика коэффициентов корреляции $r(s)$ между индикаторами отказа i -й и $(i+s)$ -й заявок в системе Pa/Pa/1 при $\alpha_1=\alpha_2=1, K_1=1, K_2=0,5$ и размере буфера $m=100$

ном режиме равно искомой вероятности P потери заявки. Длительность переходного процесса (ПП) определялась посредством достаточно большого числа независимых прогонов модели.

В общем случае при расчете любых показателей типа вероятности отказа, среднего времени ожидания и т.д. функция $r(s)$ имеет при расчете фрактальных очередей следующий вид:

$$r(s) = as^{-b}, \quad (11)$$

где a и b — некоторые константы, определяемые с помощью пробных прогонов модели ($a > 0$ и $0 < b < 1$).

Используя точную формулу (4) и асимптотическое представление (11), для доверительных интервалов находим:

$$\begin{aligned} R_N &= \frac{1}{N} \sum_{s=1}^{N-1} (N-s)r(s) = \frac{1}{N} \sum_{s=1}^{N-1} (N-s)as^{-b} \sim \\ &\sim \frac{1}{N} \int_1^{N-1} (N-s)as^{-b} ds = \\ &= \frac{a}{1-b} (N-1)^{1-b} - \frac{a}{1-b} - \frac{a}{N} \frac{1}{2-b} (N-1)^{2-b} + \frac{a}{N} \frac{1}{2-b} \sim \\ &\sim \left(\frac{a}{1-b} - \frac{a}{2-b} \right) N^{1-b}. \end{aligned} \quad (12)$$

Поэтому, в соответствии с (5), доверительный интервал запишется в виде

$$\bar{\xi} = \xi \pm 3\sigma \left(\frac{1+2R_N}{N} \right)^{1/2} \sim \xi \pm 3\sigma \cdot CN^{-b/2}, \quad (13)$$

где константа $C = \sqrt{2a \left(\frac{1}{1-b} - \frac{1}{2-b} \right)}$.

Как видим, величина полуинтервала убывает здесь медленнее, чем $N^{-1/2}$ (поскольку $b < 1$). Например, при расчете вероятности P потери заявки в системе, характеризуемой рис. 2, параметр b в (13) равен 0,15 (см. уравнение линии тренда на рис. 2), и, следовательно, величина полуинтервала убывает пропорционально $N^{-0,075}$. Иными словами, чтобы уменьшить погрешность (полуинтервал) в 10 раз, длину прогона N нужно увеличивать в $10^{1/0,075} \approx 10^{13}$ раз.

Принципиальной при любом b здесь является следующая особенность погрешностей оценок: скорость уменьшения погрешностей с ростом N всегда

меньше, чем $cN^{-1/2}$. В случае же классических очередей, как мы видели в разделе 3 статьи, корреляция элементов выборок приводит лишь к масштабному увеличению полуинтервала, но скорость его уменьшения с ростом N всегда остается равной $cN^{-1/2}$, как в случае независимых испытаний.

Кроме того, параллельные прогоны позволяют решать такие актуальные для моделирования фрактальных очередей задачи, как определение длительности ПП и распознавание отсутствия стационарного режима функционирования у моделируемой системы.

5. Распознавание отсутствия сходимости оценок к конечным значениям. Последовательные прогоны при моделировании фрактальных очередей не только неэкономичны, но и могут приводить к ошибочным заключениям, даже при весьма значительной длине выполняемых прогонов. Так, например, на рис. 3 представлен хвост эмпирического распределения времени ожидания заявок в системе Pa/Pa/1 с параметрами распределения (1) у интервалов поступления заявок $\alpha_1 = 1,1, K_1 = 1$ и параметрами распределения времени обслуживания $\alpha_2 = 1,1, K_2 = 0,1$. В этой системе $\bar{\tau} = 11, \bar{x} = 1,1$, ее коэффициент загрузки $\rho = 0,1 \ll 1$. Эмпирическое распределение, хвост которого представлен на рисунке, получено в последовательном прогоне модели по значениям времени обслуживания 100 млн заявок. Спрямление графика при логарифмическом масштабе по обеим осям координат, высокий коэффициент достоверности и значительная длина выборки приводят к выводу, что хвост является асимптотически степенным и что параметр $\alpha = 1,188$ хвоста (равный показателю степени аппроксимирующего равенния на рис. 3 найден с достаточно высокой точностью. И что, поскольку $\alpha > 1$, распределение имеет конечное м.о., т.е. среднее время ожидания в моделируемой системе конечно.

В действительности это не так. Выполнив и обработав несколько довольно коротких прогонов, можно получить наглядную информацию (рис. 3, справа) о том, что среднее время обслуживания заявок сходится к бесконечности. На этом рисунке показана зависимость от номера i заявки времени ее обслуживания, усредненного по $l = 10$ тыс. независимых реализаций процесса. Стационарный режим в системе существует, но за конечное время не достигается. График роста среднего времени обслуживания заявки имеет характерные скачки, обусловленные тяжелыми хвостами распределений Парето, задающих систему. При этом рост средней величины элемента выборки с ростом его номера имеет характерную логарифмическую скорость (логарифмическая линия тренда — гладкая кривая на рисунке).

6. Об эффективности метода ARAND. Использование параллельных прогонов не обеспечит получения адекватных результатов моделирования фрактальных очередей, если не учитывать проблему корректной реализации РТХ [6]. При этом ошибочные результаты ИМ могут быть получены не только при $\alpha \downarrow 1$, когда хвосты РТХ особенно «тяжелы», но и при $\alpha \uparrow 2$, когда эти хвосты наиболее «легки». Как показано в [10] теоретическими средствами, ИМ системы M/Pa/1 при $\alpha \uparrow 2$ будет при использовании обычных ГСЧ приводить к принципиально ошибочным результатам расчета стационарного среднего времени ожидания. Для исключения подобных ошибок рекомендуется использовать ГСЧ, построенные методом ARAND, который предложен в [8]. На рис. 4 приведен пример результатов ИМ системы M/Pa/1, подтверждающих данную рекомендацию.

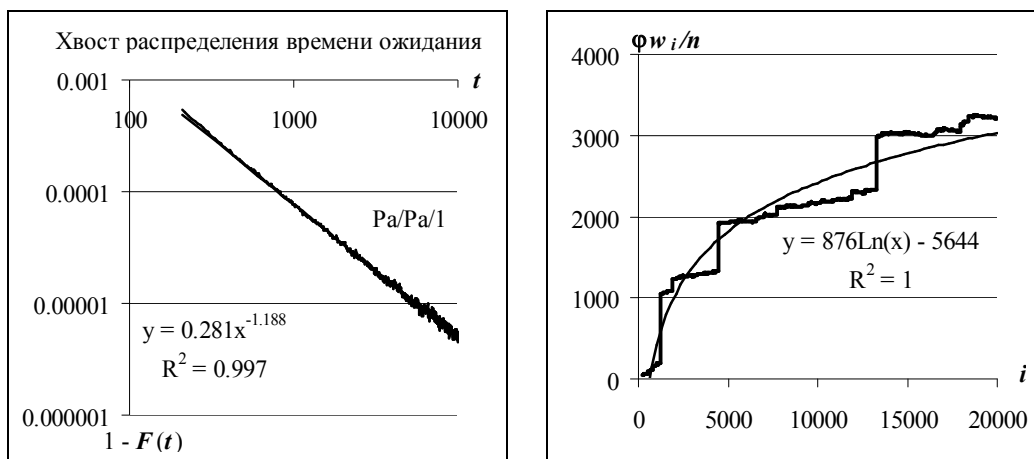


Рис. 3. Результаты обработки данных последовательного прогона (слева) и параллельных прогонов модели Pa/Pa/1 при $\alpha_1=\alpha_2=1,1$, $K_1=1$, $K_2=0,1$

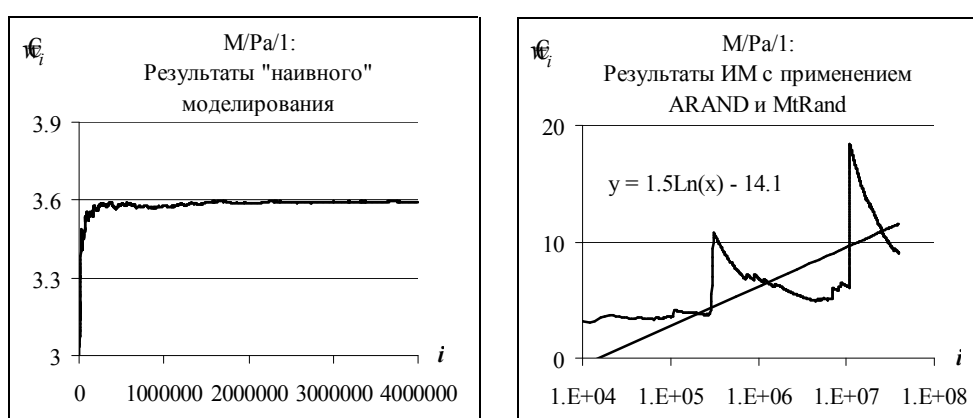


Рис. 4. Результаты ИМ системы M/Pa/1 при $\bar{\tau} = 4$, $K=1$, $\alpha=2$

На рис. 4 слева каждая оценка \bar{w}_i получена усреднением значений случайной величины w_i полученных при данном конкретном i в большом числе независимых прогонов модели. Таким же образом получены оценки, представленные кривой на графике справа. Но результаты на левом рисунке получены при использовании обычного ГСЧ, а результаты на рисунке справа получены при использовании ГСЧ, построенного методом ARAND. Мы видим, что зависимость \bar{w}_i от i не сходится ни к какому конечному значению, кривая на правом рисунке имеет характерные для не сходящейся оценки скачки, которые приводят к росту оценки приблизительно с логарифмической скоростью (масштаб по оси абсцисс на этом рисунке логарифмический). Поскольку коэффициент загрузки системы $\rho = 0,5 < 1$, то стационарный процесс существует, а поскольку оценка \bar{w}_i растет как $\ln(i)$, мы приходим к правильному заключению, что стационарное среднее время ожидания в этой системе существует и равно бесконечности. Этот вывод согласуется с точным анализом системы M/Pa/1.

7. Выводы. Предложенные в статье методы организации последовательных и параллельных прогонов моделей при расчете систем с очередями позволяют свести расчет классических и фрактальных очередей к последовательности операций, типичных для расчетной практики применения классических методов вычислительной математики. Разработанный метод

построения доверительных интервалов на основе экспоненциальных аппроксимаций (для классических очередей) и степенных аппроксимаций (для фрактальных очередей) коэффициентов корреляции элементов выборок позволяет выбирать длину и число прогонов таким образом, чтобы обеспечивать заданную точность результатов за минимальное время.

Библиографический список

1. Leland, W. E., Taqqu, M. S., Willinger W., Wilson, D.V. On the Self-Similar Nature of Ethernet Traffic. IEEE/ACM TRANSACTIONS ON NETWORKING, VOL. 2, № 1, February 1994. — P. 1–15.
2. Crovella, M. E., Taqqu, M., Bestavros, A. Heavy Tailed-Probability distributions in the World Wide Web. — 5(6): 835–846, December, 1997.
3. Kleinrock, L. Queueing Systems: V. II — Computer Applications. — New York : Wiley Interscience, 1976. — 576 p.
4. Zwart, A. P. Queueing Systems with Heavy Tails. Eindhoven University of Technology, 2001. — 227 p.
5. Задорожный, В. Н. Основная задача фрактальной теории массового обслуживания / В. Н. Задорожный // Омский научный вестник. Сер. Приборы, машины и технологии. — 2013. — № 3 (123). — С. 9–13.
6. Задорожный, В. Н. Проблемы генерации случайных величин с фрактальными распределениями // В. Н. Задорожный, О. И. Кутузов // Омский научный вестник. Сер. Приборы, машины и технологии. — 2012. — № 3 (113) — С. 20–24.

7. Zadorozhnyi, V. N. Cascade Method of Realization of Heavy-Tailed Distributions in Data Network Modelling. 2015 International Siberian conference on control and communications SIBCON, sec. Control of the Large-Scale Systems, Russia, Omsk, May 21 – 23, 2015.

8. Задорожный, В. Н. Особенности моделирования систем массового обслуживания с тяжелыми хвостами распределений на GPSS World. Метод ARAND / В. Н. Задорожный // Омский научный вестник. Сер. Приборы, машины и технологии. – 2015. – № 3 (143). – С. 307–311.

9. Клейнен, Дж. Статистические методы в имитационном моделировании : пер. с англ. / Под ред. Ю. П. Адлера и В. Н. Варьгина. – М. : Статистика, 1978. – Вып. 1. – 221 с.

10. Zadorozhnyi, V. N. Fractal Queues Simulation Peculiarities. In Communications in Computer and Information Science, 2015. – P. 413–432.

ЗАДОРЖНЫЙ Владимир Николаевич, доктор технических наук, доцент (Россия), профессор кафедры «Автоматизированные системы обработки информации и управления».

Адрес для переписки: zwn2015@yandex.ru

ЗАХАРЕНКОВА Татьяна Романовна, аспирантка кафедры «Автоматизированные системы обработки информации и управления».

Адрес для переписки: ZakharenkovaTatiana@gmail.com

Статья поступила в редакцию 18.03.2016 г.

© В. Н. Задорожный, Т. Р. Захаренкова

УДК 519.2:004.421.5:004.7

А. М. ПУРТОВ

Институт математики
им. С. Л. Соболева СО РАН,
Омский филиал

ИМИТАЦИЯ СИСТЕМ УПРАВЛЕНИЯ ПОТОКАМИ АВТОМОБИЛЕЙ НА ПЕРЕКРЕСТКЕ

Описана концептуальная модель перекрестка. Разработан новый алгоритм управления потоками автомобилей на перекрестке, основанный на использовании эталонных состояний. Концептуальная модель запрограммирована на GPSSW. Сделан сравнительный анализ четырех алгоритмов управления светофорами на перекрестке. Приведены результаты имитационных экспериментов.

Ключевые слова: перекресток, очереди автомобилей, концептуальная модель, имитационное моделирование, система управления светофорами, результаты экспериментов.

Работа выполнена по базовой теме: «Математические методы распознавания образов и прогнозирования» (034-2014-2017, номер госрегистрации 01201351843).

Введение. В крупных городах развитие сети дорог обычно отстает от темпов роста количества автомобилей. Поэтому все более актуальны задачи разработки, использования методов и средств анализа транспортных сетей, алгоритмов управления потоками автомобилей. В течение нескольких лет автором статьи разрабатывается совокупность методов для анализа задержек на маршрутах города, объединяемая названием GisAuto [1]. В GisAuto предложена следующая последовательность этапов.

1. Построение геоинформационной (ГИС) модели задержек на основных маршрутах города.

Задержки происходят на перекрестках, светофорах, пешеходных переходах.

2. Построение на ГИС-карте графов исследуемых маршрутов.

3. Сбор данных о задержках.

На этом этапе могут быть использованы экспертные, расчетные оценки, результаты наблюдений, имитационного моделирования.

4. Анализ задержек на маршрутах математическими методами (метод свертывания графов с целью получения коэффициентов чувствительности, методы визуальной и автоматической таксономии с целью классификации задержек).

5. Раскраска задержек (вершин графа) на ГИС-карте в зависимости от их значимости.

6. Визуальный анализ полученных результатов, выявление задержек, оказывающих наибольшее влияние на время прохождения исследуемых маршрутов.

7. Микроанализ проблемных участков (дополнительные наблюдения, имитационное моделирование).

Для поддержки этапов 3 и 7 была разработана имитационная модель перекрестка [2]. Позднее выяснилось, что эта модель может быть использована автономно, например, для сравнительного анализа алгоритмов управления автомобилями на перекрестке. Алгоритмы, апробированные на модели перекрестка, можно адаптировать для принятия решений в других областях деятельности. Такая ситуация встречается часто. Например, шахматы используются как модель для работ в областях искусственного интеллекта, психологии, методов принятия решений и др. В статье приведены результаты сравнительного анализа четырех алгоритмов для систем управления светофорами (СУС) на перекрестке.

По алгоритму СУС1, наиболее популярному, светофоры переключаются через заданный интервал времени.

В алгоритме СУС2, в отличие от СУС1, светофоры могут быть переключены при отсутствии автомобилей в одном из потоков.

В алгоритме СУС3 разрешение получает поток с наибольшей очередью автомобилей.

Первые три алгоритма в разных модификациях исследуются в многочисленных работах, связанных с управлением автодорожным движением.

В алгоритме СУС4, разработанном автором статьи, используются эталонные состояния перекрестка для принятия решений о разрешении движения. Обучение, поведение по эталонам свойственно многим живым организмам. Для технических систем имеет значение поиск эффективных способов использования эталонов. В работах Загоруйко Н. Г. [3, 4] приведены методы алгоритмического выбора эталонов и классификации вновь поступивших данных. Ситуацию на перекрестке тоже можно представить набором случайных данных, например, таких, как текущие очереди автомобилей. В зависимости от того, к какому из эталонов ближе текущая ситуация, принимается решение о направлениях, которые открываются для движения автомобилей.

1. Имитационная модель перекрестка. Имитационные модели различаются в первую очередь целями моделирования и используемыми методами. В частности, для перекрестков известны макро-модели и микро-модели, модели на основе агентного подхода, клеточных автоматов, классических и управляемых систем массового обслуживания [5]. В работе используется модель на основе управляемых светофорами очередей автомобилей. Более подробно, чем в статье, обзор работ по имитационному моделированию перекрестков, а также разработанные автором концептуальная и имитационная модели описаны в [2].

Реальные перекрестки имеют множество разных конфигураций, полос и направлений движения, способов размещения светофоров. В работе за основу концептуальной модели принята достаточно универсальная структура одноуровневого перекрестка (рис. 1), которая при некоторых модификациях часто встречается в публикациях различных авторов.

В модели существует четыре потока: а, г, с, е. Для обозначения направлений движения используются пары букв (например, аb, аf). В каждом потоке три полосы движения автомобилей. Каждая полоса предназначена для движения в одном направлении. Окружностями обозначены задержки автомобилей перед перекрестком в каждом из направлений.

Движением на перекрестке управляют четыре светофора (модели совокупности реальных светофоров). Каждый светофор имеет номер, пять секций и управляет одним потоком. Например, светофор 3 управляет потоком «е», светофор 2 управляет потоком «а». Каждая секция имеет два состояния (светится / не светится). Секция 1 светится красным цветом, секция 2 желтым, секции 4, 3, 5 зеленым. Светофор имеет 8 состояний. Каждое состояние светофора задается набором состояний секций, разрешающих или запрещающих движение в различных направлениях. Например, зеленые секции 3, 4, 5 светофора 2 разрешают потоку «а» движение во всех направлениях (аf, аb, аh). Последовательность переходов светофоров в разные состояния задает СУС.

Для каждого направления задается интенсивность поступления автомобилей на перекресток (интенсивность генерации транзакций).

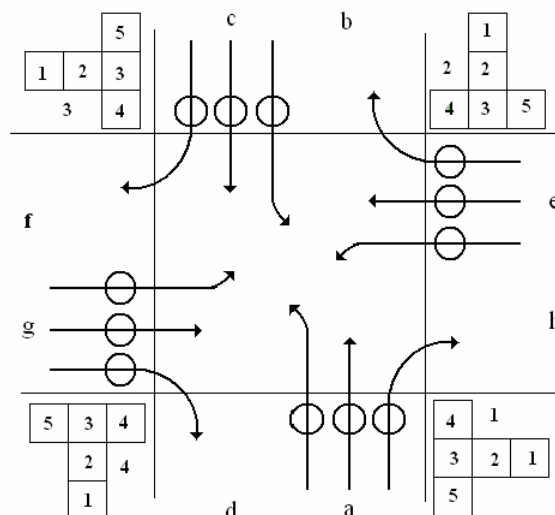


Рис. 1. Модель перекрестка

Имитационная модель перекрестка реализована на языке GPSSW. Разные версии GPSS используются с восьмидесятых годов двадцатого века. Язык GPSS хорошо зарекомендовал себя при моделировании систем в различных областях деятельности. В представляемой работе использовалась бесплатно распространяемая студенческая версия GPSSW.

Движение автомобилей имитируется перемещением транзактов по модели. Реализована следующая схема имитации прохождения автомобилей в каждом из 12 направлений.

1. Генерация транзакта, имитирующего автомобиль.
2. Постановка автомобиля в очередь перед перекрестком.
3. Ожидание освобождения первого места перед перекрестком (T1).
4. Захват автомобилем первого места перед перекрестком.
5. Ожидание автомобилем сигнала светофора, разрешающего движение (T3).
6. Проезд первого места перед светофором (T2).
7. Освобождение первого места перед светофором.
8. Выход из очереди перед перекрестком.
9. Пересечение перекрестка (T4).

На первом этапе для направления задается закон распределения времени между поступлением автомобилей на перекресток (используется блок GENERATE). В каждом конкретном случае интенсивность поступления автомобилей зависит от многих факторов: тип предыдущей задержки (светофор, переход, перекресток), время суток, время года и др. Подъезжающие к перекрестку автомобили становятся в очередь (используется блок QUEUE). В модели считается, что в каждом направлении образуется очередь автомобилей, независимая от других направлений (перестроения между полосами движения не учитываются). Если первое место перед перекрестком занято, автомобиль ожидает его освобождения в течение времени T1 (этап 3). При освобождении первого места ближайший к нему автомобиль становится первым перед перекрестком (этап 4, блок SEIZE). Если нет разрешающего сигнала светофора, автомобиль ожидает его в течение времени T3 (этап 5). При получении разрешающего сигнала светофора автомобиль проезжает (блок 6) и осво-

бождает первое место для следующего за ним автомобиля (этап 7, блок RELEASE). Считается, что после этого автомобиль выходит из очереди перед перекрестком (этап 8, блок DEPART). Таким образом, время нахождения автомобиля в очереди перед перекрестком T_q равно сумме T_1 , T_2 и T_3 . Затем автомобиль пересекает перекресток за время T_4 (блок 9) и транзакт, имитирующий его движение, удаляется из модели. Общее время задержки при прохождении автомобилем перекрестка равно сумме T_q и T_4 .

СУС может находиться в одном из пяти состояний.

В состоянии 1 разрешается движение в направлениях ab , ah , af потока a и в направлении gd (поворот направо) потока g .

В состоянии 2 разрешается движение в направлениях cf , cd , ch потока c и в направлении eb (поворот направо) потока e .

В состоянии 3 разрешается движение в направлениях ed , ef , eb потока e и в направлении ah (поворот направо) потока a .

В состоянии 4 разрешается движение в направлениях gd , gh , gb потока g и в направлении cf (поворот направо) потока c .

Состояния 1–4 являются переходными между состояниями 1–4, что соответствует желтому свету светофора.

К моменту написания статьи реализованы алгоритмы четырех СУС. В модели различаются неадаптивные (детерминированные) и адаптивные СУС. К неадаптивному относится алгоритм СУС1. К адаптивным относятся алгоритмы СУС2, СУС3, СУС4.

В алгоритме СУС1 состояния СУС переключаются в заданной последовательности через заданное время. Во всех алгоритмах перед выбором нового состояния СУС в течение заданного времени находится в состоянии 5, соответствующее желтому свету светофора.

Основное отличие алгоритма СУС2 от СУС1 состоит в том, что переход СУС в следующее состояние может произойти не только через заданное время, но и при отсутствии автомобилей в открытом для движения потоке. Для СУС2 задается минимальное время нахождения СУС в выбранном состоянии. Таким образом, в СУС2 сведено до минимума время нахождения перекрестка в ситуации, когда в открытом потоке нет автомобилей.

Алгоритм СУС3 разрешает движение автомобилям того потока, для которого суммарная длина очередей направлений максимальна. Исключение составляют потоки, для которых время ожидания превышает заданную величину. Такие направления имеют высший приоритет. В СУС3 при отсутствии автомобилей в открытом потоке или при превышении допустимого времени нахождения СУС в определенном состоянии, выбирается для открытия другой поток.

Алгоритм СУС4 использует эталонные состояния для принятия решения о выборе потока для открытия. Суть алгоритма состоит в следующем.

Считаем, что текущее состояние перекрестка определяется вектором $S = \{Q_a, Q_c, Q_e, Q_g\}$, где Q_a, Q_c, Q_e, Q_g — суммарные длины очередей автомобилей в направлениях соответствующих потоков (a, c, e, g). Тогда каждому потоку можно задать эталон $ST = \{Q_a, Q_c, Q_e, Q_g\}$, при котором ему отдается предпочтение при принятии решения о разрешении движения. Эталонное состояние можно считать точкой в четырехмерном пространстве. Таким образом, имеем четыре точки, соответствующих потокам

a, c, e, g . Вычисляются евклидовы расстояния текущего состояния S до каждой из четырех точек. СУС разрешает движение тому потоку, расстояние до точки эталона которого минимально. В остальном СУС4 аналогичен СУС3.

2. Результаты имитационных экспериментов. Проведение экспериментов имело две цели.

1. Проверка работоспособности и адекватности модели в рамках поставленных целей моделирования.

2. Сравнительный анализ алгоритмов управления потоками автомобилей (СУС1, СУС2, СУС3, СУС4) с точки зрения задержек на перекрестке.

Во всех экспериментах неизменными оставались следующие основные параметры:

— в любом состоянии СУС не может находиться более 20 с;

— время блокировки потока не может превышать более 60 с;

— время между поступлением автомобилей на перекресток T_{in} (время между генерацией транзактов) — случайная величина, распределенная экспоненциально. Интенсивность поступления автомобилей IN — величина, обратная T_{in} .

В результате предварительных экспериментов для СУС4 был выбран следующий набор эталонных состояний для потоков a, c, e, g : $ST_a = \{8, 0, 0, 0\}$, $ST_c = \{0, 8, 0, 0\}$, $ST_e = \{0, 0, 8, 0\}$, $ST_g = \{0, 0, 0, 8\}$.

Серия экспериментов 1. Цель экспериментов состояла в сравнении СУС1 с адаптивными СУС2, СУС3, СУС4 при малой, средней и высокой IN . На рис. 2 для четырех СУС показаны зависимости времени ожидания автомобилей в очереди T_q (в секундах) от IN (число автомобилей, поступающих за 1 с). Во всех направлениях автомобили поступали с одинаковой интенсивностью. Для каждой СУС значение T_q определялось как среднее по всем направлениям. Анализ зависимостей позволяет сделать следующие выводы.

1. Адаптивные СУС при любой IN эффективнее СУС1. Правда, при средних IN разница меньше, чем при малых и высоких значениях.

2. Перегрузка для СУС1 наступает раньше, чем для адаптивных СУС.

3. При равных для направлений IN значение T_q для трех адаптивных СУС примерно одинаково.

Результаты, показанные на рис. 3, получены при разных IN для потоков. Для направлений потоков a, c интенсивность равна 0,07. Для направлений потоков e, g интенсивность равна 0,11. На диаграмме показаны средние значения T_q для всех направлений, а также усредненное по направлениям T_q (помечено буквой m). Анализ данных диаграммы позволяет сделать следующие выводы, относящиеся к условиям разной IN для потоков.

1. Для потоков слабой интенсивности СУС1 не хуже, а иногда и лучше адаптивных алгоритмов, отдающих предпочтение тем потокам, очереди которых больше. Для потоков высокой интенсивности преимущества адаптивных алгоритмов возрастают.

2. Для потоков с поворотом направо (ah, cb, eb, gd) значение T_q значительно меньше, чем для других потоков. Это происходит благодаря тому, что им разрешено движение при двух состояниях СУС, а другим потокам только при одном.

3. Адаптивные СУС выравнивают T_q для потоков слабой и высокой интенсивности.

4. Алгоритм СУС4 и в условиях разной интенсивности не хуже других адаптивных алгоритмов.

Дальнейшие эксперименты показали, что при повышении интенсивности потоков алгоритмы СУС3 и СУС4 значительно превосходят алгоритм СУС2.

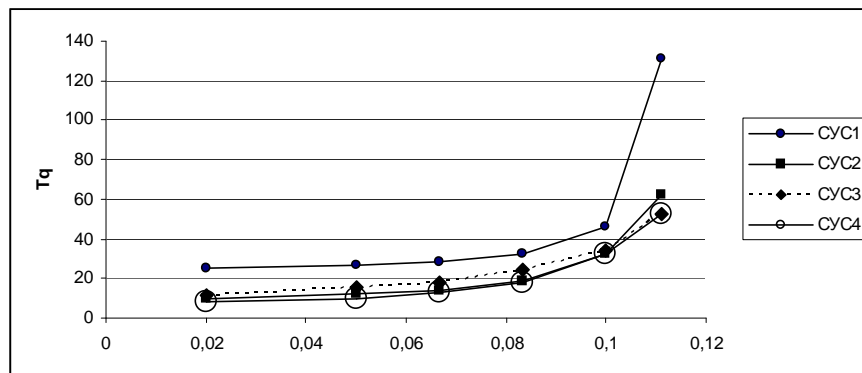
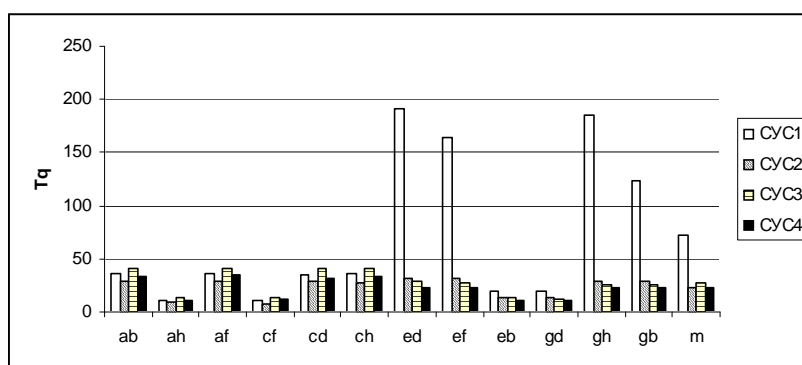
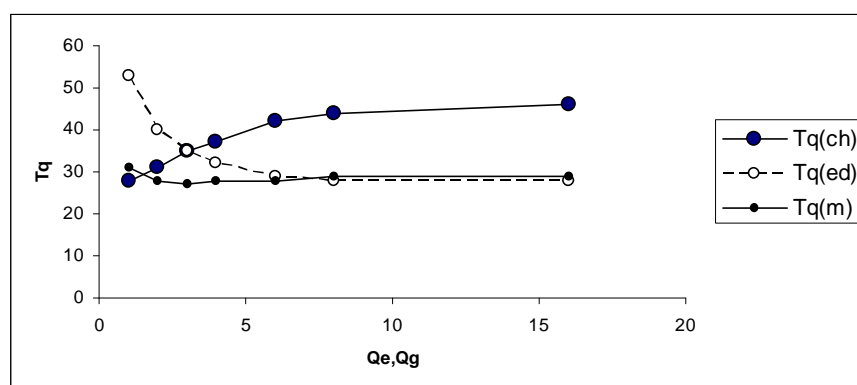
Рис. 2. Зависимость Tq от интенсивности

Рис. 3. Сравнение СУС при разных интенсивностях

Рис. 4. Зависимость Tq от разницы эталонов

Серия экспериментов 2. Цель экспериментов заключалась в получении данных о том, как разница эталонов в СУС4 влияет на Tq при разных интенсивностях потоков. Для направлений потоков а, с интенсивность равна 0,07. Для направлений потоков е, г интенсивность равна 0,125. Для потоков а, с эталонные состояния не изменялись ($STa = \{8, 0, 0, 0\}$, $STc = \{0, 8, 0, 0\}$). Для потоков е, г ненулевые значения эталонов принимали значения 1, 2, 3, 4, 6, 8, 16. График зависимости Tq от перечисленных значений показан на рис. 4. На графике представлены Tq для направления ch (направление низкой интенсивности), для направления ed (направление высокой интенсивности), а также $Tq(m)$, усредненное по всем направлениям. Анализ графика позволяет сделать следующие выводы.

1. Соотношение значений эталонов может сильно влиять на Tq для отдельных направлений.

2. Усредненное значение $Tq(m)$ слабо зависит от соотношения эталонов. Это связано с перераспределением времени между направлениями.

3. Алгоритм СУС4 с помощью подбора эталонных состояний можно настраивать на различные режимы обслуживания направлений разной интенсивности.

Считается, что основной недостаток имитационных моделей заключается в том, что получаемые результаты имеют меньшую общность, чем аналитические модели. Например, в проведенных экспериментах можно так подобрать параметры (законы распределения времени поступления автомобилей, циклы светофоров), чтобы разница между адаптивными и неадаптивными алгоритмами была минимальной. Но для этого придется создавать специальные условия для неадаптивного алгоритма. В целом имитационные модели при правильном подборе параметров могут показывать общие тенденции, отвечать

на вопросы о последствиях каких-то изменений. Главное — правильно относиться к полученным результатам. Необоснованный выход за рамки любой модели приводит к неверным выводам.

Заключение. Разработаны имитационные модели для анализа временных задержек при проезде перекрестка автомобилями под управлением СУС, функционирующих по детерминированным и адаптивным алгоритмам. Разработан новый адаптивный алгоритм для СУС, основанный на использовании эталонных состояний перекрестка. Модель апробирована при проведении имитационных экспериментов с целью сравнительного анализа детерминированной и трех адаптивных СУС. Эксперименты показали, что алгоритм на основе эталонных состояний не хуже известных алгоритмов с точки зрения временных задержек. Показаны большие возможности настройки алгоритма на решение задач распределения ресурса времени между направлениями.

Разработанные модели и алгоритмы в совокупности с описанными выше методами и средствами системы GisAUTO могут использоваться для анализа и модернизации автотранспортных сетей крупных городов.

Продолжение работы может быть связано с расширением области применения эталонов в системах управления и принятия решений.

Библиографический список

1. Пуртов, А. М. Методы и средства анализа автотранспортных сетей в GisAuto / А. М. Пуртов // Омский научный вест-

ник. Сер. Приборы, машины и технологии. — 2015. — № 3 (143). — С. 282–286.

2. Пуртов, А. М. Разработка и анализ имитационной модели перекрестка для системы GisAuto / А. М. Пуртов // Омский научный вестник. Сер. Приборы, машины и технологии. — 2013. — № 1 (117). — С. 225–229.

3. Загоруйко, Н. Г. Прикладные методы анализа данных и знаний / Н. Г. Загоруйко. — Новосибирск : Изд-во Ин-та математики, 1999. — 270 с.

4. Загоруйко, Н. Г. Когнитивный анализ данных / Н. Г. Загоруйко. — Новосибирск : Гео, 2013. — 186 с.

5. Задорожный, В. Н. Аналитико-имитационные методы решения актуальных задач системного анализа больших сетей : моногр. / В. Н. Задорожный, Д. Ю. Долгушин, Е. Б. Юдин ; под ред. В. Н. Задорожного. — Омск : Изд-во ОмГТУ, 2013. — 324 с.

ПУРТОВ Андрей Михайлович, кандидат технических наук, доцент (Россия), старший научный сотрудник лаборатории методов проблем преобразования информации.

Адрес для переписки: andr.purtov@yandex.ru

Статья поступила в редакцию 17.02.2016 г.

© А. М. Пуртов

Книжная полка

Информационно-коммуникационные технологии в управлении : моногр. / А. А. Косолапов [и др.]. — Одесса : Куприенко СВ, 2015. — 244 с.

Монография содержит научные исследования авторов в области информационных технологий. Может быть полезна для руководителей, экономистов, менеджеров и других работников предприятий и организаций, представителей органов государственной власти и местного самоуправления, преподавателей, соискателей, аспирантов, магистрантов и студентов высших учебных заведений.

Цехановский, В. В. Управление данными : учеб. для вузов по направлению подготовки бакалавра «Информационные системы и технологии» / В. В. Цехановский, В. Д. Чертовской. — СПб. : Лань, 2015. — 432 с.

В учебнике излагаются вопросы построения и использования технологии баз данных в процессе выработки и принятия решений. Учебник отличается системным рассмотрением теоретических вопросов, которые сопровождаются компьютерной реализацией. Это позволяет лучше понять процедуры построения, работы и использования баз данных. Рассмотрены устоявшиеся теоретические вопросы и новые аспекты, относящиеся как к локальным, так и к распределенным базам данным, объектно-ориентированным базам данных, хранилищам данных. Подробно рассмотрен режим «клиент – сервер», в том числе в удаленном варианте и при облачных вычислениях. Учебник предназначен для бакалавров, обучающихся по направлениям подготовки, входящим в УГС «Информатика и вычислительная техника», изучающих дисциплину «Управление данными», и может быть использован в других родственных дисциплинах («Интеллектуальные информационные системы и технологии», «Информационные технологии», «Корпоративные информационные системы», «Системы поддержки принятия решений»).

ОПТИМИЗАЦИЯ ВЗАИМОДЕЙСТВИЯ ПРЕДПРИЯТИЙ ДЛЯ МУНИЦИПАЛЬНОГО ПЛАНИРОВАНИЯ И УПРАВЛЕНИЯ

Работа посвящена решению задачи оптимального взаимодействия предприятий близкой отраслевой специализации в целях муниципального планирования и управления. Построена и исследована модель целочисленного линейного программирования, приведены некоторые ее свойства. Предложены подходы к решению рассматриваемой задачи, основанные на применении декомпозиционных алгоритмов.

Ключевые слова: математическое моделирование, дискретная оптимизация, целочисленное программирование, взаимодействие предприятий, маршрутизация транспорта, муниципальное управление.

Особенностью административно-территориального устройства ряда регионов Российской Федерации является доминирование сельских поселений в структуре расселения. Поэтому актуальной проблемой регионального управления является поиск точек роста за пределами городской агломерации и повышение самоорганизации и саморазвития муниципальных образований. К примеру, на территории Омской области расположено 424 муниципальных образования, в том числе один городской округ (г. Омск), 32 муниципальных района, включающих 26 городских и 365 сельских поселений [1]. По данным на 1 января 2016 года в муниципальных районах Омской области проживают 797,6 тыс. чел., или 40,3 % населения региона [2].

Социально-экономическое благополучие населения муниципальных районов в значительной степени определяется эффективностью базовой отрасли — сельского хозяйства. Так, в 2015 году объем произведенной продукции сельского хозяйства в Омской области составил 96 236,7 млн руб., что соответствует второму месту в Сибирском федеральном округе (далее — СФО) после Алтайского края [3]. Одной из важнейших задач муниципального управления, закрепленной Федеральным законом от 6 октября 2003 г. № 131-ФЗ «Об общих принципах организации местного самоуправления в Российской Федерации», является «создание условий для развития сельскохозяйственного производства, содействие развитию малого и среднего предпринимательства» [4]. В настоящее время, учитывая сокращение внешних источников стимулирования развития, актуальным направлением активизации потенциала самоорганизации локальных социально-экономических систем может служить оптимизация взаимодействия предприятий, расположенных на их территории.

Как отмечают специалисты, развитие партнерских отношений между хозяйствующими субъектами может гармонично сочетаться с отношениями

конкуренции, характерными для рыночной среды [5, с. 102]. Мировая и отечественная практика показывает, что сильные локальные взаимосвязи в экономике регионов мобилизуют их эндогенный потенциал и обеспечивают устойчивое развитие.

Необходимо принимать во внимание, что планирование, организация и управление взаимодействием предприятий предполагает решение комплекса сложных задач, таких как определение сфер кооперации, формирование базы данных о ресурсах возможных партнеров, отбор приоритетных форм сотрудничества, обеспечение баланса интересов, оценка результатов, корректировка системы взаимодействий по результатам оценивания и других. Высокий уровень сложности принятия управленческих решений обуславливает необходимость разработки соответствующего математического аппарата. Опыт применения различных экономико-математических моделей в качестве поддержки принятия решений в сфере организации сотрудничества предприятий отражен в российской и зарубежной литературе [5–8]. При этом анализ источников показал, что для решения задачи оптимизации взаимодействия предприятий в рамках локальных территорий, учитывающей особенности сельскохозяйственной отрасли, требуется разработка специальных экономико-математических моделей.

В современных условиях рыночной организации хозяйства каждое предприятие в значительной степени экономически изолированно преследует свои коммерческие интересы, рассматривает другие субъекты отрасли как конкурентов и не стремится к сотрудничеству. При этом общий эффект от производственно-экономической деятельности системы (муниципального образования) складывается из положительных результатов работы всех ее элементов. Поэтому для достижения общесистемных целей органы муниципального управления должны организовывать взаимовыгодное сотрудничество хозяйствующих субъектов. Необходимо также учитывать, что

для сельскохозяйственной отрасли характерна повышенная степень риска и неопределенности, поскольку результативность деятельности во многом зависит от природных условий. Данное обстоятельство обуславливает жесткие требования к выполнению сроков запланированных работ.

На наш взгляд, взаимодействие сельскохозяйственных предприятий позволит решить следующие проблемы. Во-первых, обеспечит предприятиям, не обладающим требуемым объемом производственных ресурсов, совершить запланированную деятельность в срок. Во-вторых, предприятия, имеющие избыточные ресурсы, получают дополнительную выгоду от предоставления их во временное пользование. Муниципальное образование в целом получит интегральный эффект от повышения объемов сельскохозяйственного производства.

Наиболее целесообразной экономической формой взаимодействия, по нашему мнению, является оказание производственных услуг. Для организации подобного взаимодействия на уровне муниципального района необходимо сформировать базу данных, содержащую информацию об автономных возможностях каждого субъекта хозяйственной деятельности с целью выявления недостающих и избыточных производственных ресурсов с учетом запланированных объемов работ. Созданная база данных позволит разработать систему поддержки принятия решений для органов муниципального планирования и управления с использованием экономико-математического моделирования и современных информационных технологий.

В данной работе представлена математическая модель для решения задачи оптимизации взаимодействия предприятий сельскохозяйственной отрасли. В качестве объекта моделирования рассмотрены сельскохозяйственные предприятия близкой отраслевой специализации, имеющие схожие наборы производственно-технических средств (сельскохозяйственной техники) и выполняющие различные виды работ (вспашка, боронование, прикатывание и другие) в рамках локальной территории (муниципального района).

Перейдем к описанию математической модели оптимизации системы производственных связей взаимодействующих предприятий. Пусть имеется конечное множество предприятий, каждому из которых необходимо выполнить некоторые виды работ в заданные сроки. Будем называть *поставщиками производственных услуг* (далее — *поставщиками*) предприятия, которые имеют избыточные, неиспользованные производственно-технические средства. Предприятия, которые испытывают потребность в необходимых технических средствах для выполнения работ в установленные сроки, определим как *потребителей производственных услуг* (далее — *потребители*).

Для каждого из поставщиков известен набор имеющихся у него свободных технических средств, для которых определены значения следующих параметров: производительность и стоимость перемещения технического средства, а также затраты на выполнение работ. Для всех потребителей заданы необходимые объемы работ, которые должны быть завершены в плановый период. Кроме этого, определена транспортная сеть, связывающая предприятия, известны кратчайшие расстояния между каждой парой из них. *Маршрутом технического средства* будем называть последовательность предприятий, состоящую из поставщика и упорядоченного

набора потребителей. В задаче требуется найти маршруты технических средств и объемы производимых ими работ, при которых суммарные затраты на выполнение работ и перемещение технических средств для потребителей будут минимальными, при следующих условиях:

— потребители производственных услуг должны выполнить необходимые объемы работ в установленные сроки;

— поставщики производственных услуг должны предоставить технические средства с учетом установленных потребителями сроков выполнения работ;

— потребитель может обслуживаться произвольным числом технических средств любых поставщиков;

— каждое техническое средство принадлежит одному из поставщиков;

— маршрут должен начинаться и заканчиваться на предприятии-поставщике.

Маршрут технического средства называется *допустимым*, если суммарное время выполнения работ для потребителей не превосходит плановый период.

Для построения математической модели воспользуемся следующими обозначениями:

$G=(V, E)$ — граф, вершины которого соответствуют предприятиям, а ребра — путям между ними, причем $V = I \cup J : I \cap J = \emptyset$, где I — множество поставщиков, J — множество потребителей и каждому ребру приписан вес, равный расстоянию между соответствующими предприятиями;

T — множество технических средств;

p_t — производительность технического средства t , $t \in T$;

d — длительность планового периода;

$A = (a_{it}), i \in I, t \in T$ — матрица принадлежности технических средств, где

$$a_{it} = \begin{cases} 1, & \text{если техническое средство } t \\ & \text{принадлежит } i \\ 0, & \text{в противном случае} \end{cases};$$

S_j — объем работ, которые должен выполнить потребитель j , $j \in J$;

c_t — стоимость выполнения единицы объема работы техническим средством t , $t \in T$;

w_{ij} — вес ребра, соединяющего поставщика производственной услуги i и потребителя j , $i, j \in V$;

b_t — стоимость перемещения технического средства t на единицу расстояния.

Переменные модели:

$$z_{ijt} = \begin{cases} 1, & \text{если техническое средство } t \\ & \text{перемещалось из } i \text{ в } j \\ 0, & \text{в противном случае} \end{cases}, i, j \in V, t \in T;$$

x_{jt} — объем работ, производимых техническим средством t в пункте j , $j \in J, t \in T$.

Полученная модель частично целочисленного линейного программирования принимает следующий вид:

$$f = \sum_{t \in T} \sum_{j \in J} c_t x_{jt} + \sum_{i \in V} \sum_{j \in V} \sum_{t \in T} w_{ij} b_t z_{ijt} \rightarrow \min \quad (1)$$

при условиях:

$$\sum_{t \in T} x_{jt} = S_j, j \in J; \quad (2)$$

$$\sum_{j \in J} x_{jt} \leq dp_t, t \in T; \quad (3)$$

$$\sum_{j \in J} z_{ijt} \leq a_{it}, i \in I, t \in T; \quad (4)$$

$$\sum_{i \in V} \sum_{j \in V} z_{ijt} \leq M \sum_{i \in I} \sum_{j \in V} z_{ijt}, t \in T; \quad (5)$$

$$\sum_{i \in V} z_{ijt} = \sum_{k \in V} z_{jkt}, j \in V, t \in T; \quad (6)$$

$$\sum_{i \in R} \sum_{j \in R} z_{ijt} \leq |R| - 1, \forall R \subseteq J, t \in T; \quad (7)$$

$$x_{jt} < M \sum_{i \in V} z_{ijt}, j \in J, t \in T, M \gg 1; \quad (8)$$

$$x_{jt} \geq 0, j \in J, t \in T; \quad (9)$$

$$z_{ijt} \in \{0, 1\}, i, j \in V, t \in T. \quad (10)$$

Целевая функция (1) состоит в минимизации общих затрат всех предприятий на оказание производственных услуг и транспортировку технических средств к потребителям. Условия (2) гарантируют оказание производственных услуг всем нуждающимся в них потребителям в полном объеме. Ограничения (3) означают, что все работы должны быть выполнены до конца планового периода. Неравенства (4) и (5) задают условия, при которых каждое техническое средство может выезжать только из предприятия, которому принадлежит, и в его маршруте должен содержаться хотя бы один поставщик. Балансовые условия (6) означают, что число прибывших на каждое предприятие транспортных средств должно быть равно числу выехавших. Неравенства (7) исключают подциклы в маршрутах [9]. Ограничения (8) допускают использование в данном предприятии только тех технических средств, которые прибыли в него. $M > 0$ — достаточно большое число, предназначенное для настройки системы ограничений.

Необходимо отметить следующие свойства полученной модели. Если зафиксировать значения переменных, отвечающих за маршрут (z_{ijt}), то мы получим задачу транспортного типа [10], если же будут известны значения объемов работ, которые необходимо произвести, задача превращается в задачу маршрутизации с несколькими депо (Multiple Depot VRP, MDVRP) [11]. Указанные свойства могут быть использованы для разработки декомпозиционных алгоритмов решения задачи, основанных на ее разбиении на подзадачи (транспортную и MDVRP) и их последовательном решении.

Следует заметить, что современные пакеты прикладных программ, предназначенные для нахождения точных решений задач ЦЛП, не позволяют получать за приемлемое время оптимальные решения MDVRP даже относительно небольшой размерности. Поэтому перспективным является использование альтернативных методов решения. К ним относится, например, алгоритм последовательной (лексикографической) оптимизации [12]. В настоящий момент нами ведется разработка программного комплекса, основанного на использовании данного метода.

В работе [6] была рассмотрена более простая математическая модель с аналогичными свойствами. Построенная модель была апробирована на реальных данных сельскохозяйственных предприятий Большереченского района Омской области. В качестве решателя использовался пакет GAMS. Были получены оптимальные решения, которые указывают на целесообразность применения данных моделей и деком-

позиционных алгоритмов к описанным задачам.

Проведенные исследования позволяют сделать вывод о том, что предложенные математические модели и алгоритмы могут служить основой для систем поддержки принятия решений в области муниципального планирования и управления.

Библиографический список

1. Регионы России. Основные характеристики субъектов Российской Федерации. 2015 : стат. сб. / Росстат. — М., 2015. — 672 с.
2. Оценка численности постоянного населения на 1 января 2016 года и в среднем за 2015 год (опубликовано 09.03.2016 г.) // Официальный сайт Федер. службы гос. стат. — Режим доступа : http://www.gks.ru/wps/wcm/connect/rosstat_main/rosstat/ru/statistics/population/demography/ (дата обращения: 13.03.2016).
3. Социально-экономическое положение Сибирского федерального округа в 2015 году // Официальный сайт Федер. службы гос. стат. — Режим доступа : http://www.gks.ru/wps/wcm/connect/rosstat_main/rosstat/ru/statistics/publications/catalog/doc_1140086420641 (дата обращения: 13.03.2016).
4. РФ. Законы. Об общих принципах организации местного самоуправления в Российской Федерации : Федер. закон [от 06.10.2003, № 131-ФЗ] // Доступ из справ.-правовой системы «Консультант плюс».
5. Никитаева, А. Ю. Модели межфирменного взаимодействия: использование потенциала партнерства для модернизации основных сегментов хозяйственного комплекса юга России / А. Ю. Никитаева, А. В. Алешин // TERRA ECONOMICUS. — 2013. — Т. 11, № 3, ч. 2. — С. 101–106.
6. Михаль, В. А. О решении некоторых производственно-распределительных задач с логическими и ресурсными ограничениями / В. А. Михаль // Проблемы оптимизации сложных систем : тр. X Междунар. Азиатской шк.-семинара. — Кыргызская Республика, 2014. — С. 502–505.
7. Колоколов, А. А. Оптимизация системы производственных услуг в условиях межфирменного взаимодействия / А. А. Колоколов, Е. Я. Семерханова // Вестник СибГУТИ. — 2014. — № 3. — С. 13–22.
8. Колоколов, А. А. О решении одной задачи производственного обслуживания АПК / А. А. Колоколов, Т. Ю. Степанова, Е. Я. Семерханова, В. А. Михаль // Сб. науч. тр. V Междунар. школы-симпозиума АМУР–2011. — Симферополь : ТНУ им. В. И. Вернадского, 2011. — 411 с.
9. Dantzig G., Fulkerson R., Johnson S. Solution of a large-scale traveling-salesman problem // Operations Research. 1954. V. 2. № 6. P. 393–410.
10. Юдин, Д. Б. Задачи и методы линейного программирования: задачи транспортного типа / Д. Б. Юдин, Е. Г. Гольштейн. — Изд. 3-е. — М. : Либроком, 2010. — 184 с.
11. Wren A., Holliday A. Computer scheduling of vehicles from one or more depots to a number of delivery points // Journal of the Operational Research Society. 1972. 23:333–344.
12. Математические методы в экономике / И. И. Еремин [и др.]. — Екатеринбург : У-Фактория Пресс, 2002. — 303 с.

МИХАЛЬ Виктор Александрович, аспирант лаборатории дискретной оптимизации Института математики им. С. Л. Соболева Сибирского отделения Российской академии наук, Омский филиал.

Адрес для переписки: msheul.san@gmail.com

ЮДИНА Марина Александровна, старший преподаватель кафедры «Финансы, кредит, бухгалтерский учет и аудит» Омского государственного университета путей сообщения.

Адрес для переписки: myum@inbox.ru

Статья поступила в редакцию 14.03.2016 г.

© В. А. Михаль, М. А. Юдина

МЕТОД МАШИННОГО ОБУЧЕНИЯ РАНЖИРОВАНИЮ НА ОСНОВЕ МОДИФИЦИРОВАННОГО ГЕНЕТИЧЕСКОГО АЛГОРИТМА ДЛЯ МЕТРИКИ NDCG

Рассмотрены задача ранжирования документов на странице результатов информационного поиска и вопросы машинного обучения ранжированию. Предложен подход к оптимизации функции ранжирования с использованием метрики качества $NDCG$ на основе модифицированного генетического алгоритма. Проведены исследования разработанных алгоритмов (на тестовых коллекциях LETOR) и показана их эффективность для машинного обучения ранжированию.

Ключевые слова: информационный поиск, машинное обучение ранжированию, релевантность, оптимизация, генетические алгоритмы.

1. Введение. В современных информационно-поисковых системах (ИПС) объемы данных, которыми оперирует система, настолько велики, что ключевой задачей становится ранжирование релевантных документов в ответ на поисковый запрос пользователя. На данном этапе развития ИПС наибольший интерес представляет машинное обучение (МО) ранжированию. Существующие подходы к МО, основанные на численных методах (в частности, на градиентных методах) или на аналитических расчетах [1], имеют ряд недостатков, которые существенно влияют на качество информационного поиска и временные затраты, необходимые для ранжирования релевантных документов.

В начале проведенных исследований были рассмотрены списочные подходы к машинному обучению ранжированию [1, 2], в большинстве из которых используется метод градиентного спуска [2]. В рассмотренных работах МО сводится к оптимизации метрик качества поиска (МКП), однако используются только метрики, представленные непрерывными функциями. Это ограничение зачастую приводит к тому, что в результате оптимизации функция ранжирования имеет менее высокие оценки по многим важным принятым показателям (DCG , $nDCG$, $Graded Mean Reciprocal Rank$ и т.д.), являющимся дискретными функциями. В работе [3] предложено применение генетических алгоритмов (ГА) [4, 5] при обучении ранжированию для минимизации функции потерь Хубера [6] с использованием экспертных оценок релевантности в качестве эталонных значений. Также был предложен подход к МО на основе оптимизации дискретных метрик качества информационного поиска [7].

2. Постановка задачи машинного обучения ранжированию. В большинстве современных информационно-поисковых систем функция ранжирования строится на основе n простых ранжирующих функций (ПРФ) и может быть записана в виде:

$$RF(q, d) = \sum_{i=1}^n WC_i \cdot SRF_i, \quad (1)$$

где SRF_i — i -ая простая ранжирующая функция для документа d и запроса q , WC_i — весовой коэффициент i -ой простой ранжирующей функции, n — количество ПРФ в системе ранжирования.

В ходе машинного обучения ранжированию использован набор поисковых документов D и запросов Q из тестовой коллекции LETOR [8]. Для всех запросов $q \in Q$ сформирована пара с каждым документом $d \in D$. Для каждой такой пары ИПС определяет значения релевантности, которые используются для ранжирования поисковой выдачи. Для того чтобы произвести оценку качества ранжирования, системе требуются эталонные значения релевантности E для каждой пары документ-запрос (d, q) . Для этих целей используются экспертные оценки релевантности.

Для проведения исследования использована ИПС, в которой ранжирование производится на основе $N=5$ простых ранжирующих функций $SRF_i(WC)$, $i = \overline{1, N}$, которые образуют векторный критерий оптимальности:

$$RF(WC) \in RF,$$

где $WC \in \{WC\}$ — вектор варьируемых параметров; $\{WC\}$, $\{RF\}$ — пространства параметров и векторных критериев соответственно.

Применение генетических алгоритмов для МО ранжированию делает возможным максимизацию дискретных метрик качества, таких как $nDCG$. Метрика $nDCG$ ранжирования документов в поисковой системе определяется в соответствии с выражением [9]:

$$DCG @ n = \sum_{p=1}^n \frac{2^{grade(p)} - 1}{\log_2(2 + p)}, \quad (2)$$

где $grade(p)$ — средняя оценка релевантности, выставленная экспертами документу, расположенному на позиции p в списке результатов, $grade \in [0, 3]$; $1/\log_2(2 + p)$ — коэффициент, зависящий от позиций документа (первые документы имеют больший вес).

Тогда нормализованная версия $NDCG$ запишется в виде

$$NDCG @ n = DCG @ n / Z,$$

где Z — фактор нормализации, который равен максимально возможному значению $DCG@n$ для данного запроса (т.е. равен DCG идеального ранжирования).

Таким образом, для оптимизации (максимизации) метрики $nDCG$, целевая функция (RM) запишется в следующем виде

$$RM = \left(\sum_{p=1}^n \frac{2^{grade(p)}}{\log_2(2+p)} \right) / Z \rightarrow \max, \quad (3)$$

3. Метрики качества ранжирования поисковой выдачи. При ранжировании документов в поисковой выдаче в роли критериев выступают метрики качества. Из списка общепринятых метрик оценки качества ИПС выбраны три основные, оценивающие точность, релевантность и полноту информационного поиска.

1. Критерий точности информационного поиска

$$p = a / (a + b),$$

где a — количество найденных релевантных документов, b — количество документов, ошибочно принятых за релевантные.

2. Критерий V_{pref} , оценивающий релевантность информационного поиска, используется для обработки задания с R релевантными документами и вычисляется по формуле

$$V_{pref} = \frac{1}{R} \sum_r (1 - Non\ Rel\ Before(r) / R). \quad (4)$$

Здесь символом r обозначен известный релевантный документ, а $NonRelBefore(r)$ — число известных нерелевантных документов, ранжированных выше, чем r (при вычислении учитываются только первые R оцененных нерелевантных документов из прогона).

3. Критерий полноты поисковой выдачи

$$r = a / (a + c),$$

где a — количество найденных релевантных документов, c — количество ненайденных релевантных документов.

4. Тестовые коллекции. В задаче машинного обучения ранжированию необходим набор документов и запросов с соответствующими оценками релевантности, определенными экспертами. Эти данные используются для машинного обучения ранжирующей функции, а также для оценки качества

ранжирования поисковой выдачи системой. В процессе МО тестовые коллекции используются в качестве обучающей выборки и, следовательно, оказывают значительное влияние на результаты. Для проведения исследований использована тестовая коллекция документов и запросов LETOR. Эта коллекция используется в исследованиях в области поиска информации подразделениями Microsoft Research. В табл. 1 приведены характеристики тестовых коллекций LETOR.

5. Модифицированный генетический алгоритм.

Для использования генетических алгоритмов в машинном обучении ранжированию задача должна быть поставлена таким образом, чтобы решение было закодировано в виде вектора (генотипа), где каждый ген может быть битом, числом или другим объектом. В данном случае генотип представлен вектором весовых коэффициентов для соответствующих факторов ранжирования. Условием остановки выполнения генетического алгоритма является нахождение оптимального решения, исчерпание числа поколений или времени, отведенного на эволюцию.

Следует отметить, что ГА наиболее эффективны в поисках области глобального экстремума, однако они могут медленно работать, когда необходимо найти локальный минимум в этой области. Предлагаемый способ избежать этого недостатка — создание модифицированного генетического алгоритма (МГА), который будет переключаться на локальный (быстродействующий) алгоритм оптимизации после нахождения области глобального оптимума с помощью базового ГА. Предлагаемый в работе МГА представляет собой гибридный метод на основе классического ГА и метода Нелдера — Мида (симплекс-алгоритма). Метод Нелдера — Мида, часто используемый алгоритм нелинейной оптимизации, является численным методом для нахождения минимума целевой функции в многомерном пространстве [10]. Предлагаемый в данной работе гибридный алгоритм MGA переключается на метод Нелдера — Мида после выполнения условий остановки ГА. Блок-схема алгоритма MGA показана на рис. 1.

При выполнении исследований принято ограничение на количество вычислений целевой функции ($N_{RF} = 16\ 000$) при поиске области глобального экстремума и условие перехода на алгоритм локальной оптимизации на основе метода Нелдера — Мида (после того как базовый генетический алгоритм выполнит 75 % N_{RF} операций).

6. Результаты. В результате проведенных исследований с помощью алгоритма машинного обучения

Таблица 1
Количество документов и запросов в тестовых коллекциях

Название тестовой коллекции	Название подсистемы	Количество запросов	Количество документов
LETOR 4.0	MQ2007	1692	69623
LETOR 4.0	MQ2008	784	15211
LETOR 3.0	OHSUMED	106	16140
LETOR 3.0	Gov03td	50	49058
LETOR 3.0	Gov03np	150	148657
LETOR 3.0	Gov03hp	150	147606
LETOR 3.0	Gov04td	75	74146
LETOR 3.0	Gov04np	75	73834
LETOR 3.0	Gov04hp	75	74409

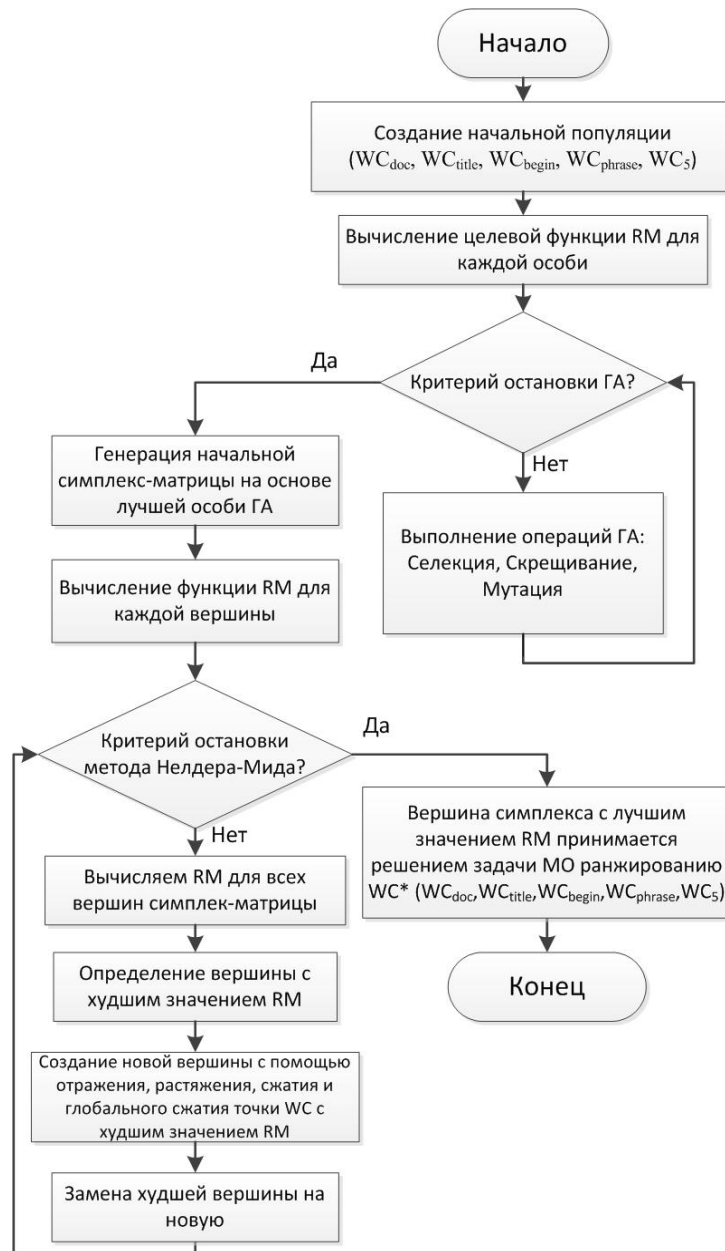


Рис. 1. Блок-схема гибридного алгоритма MGA на основе генетических алгоритмов и метода Нелдера-Мида

ранжированию *LTR-MGA* получен вектор весовых коэффициентов WC^* для функции ранжирования. Далее на основе данных из тестовой коллекции LETOR произведена оценка качества ранжирования, для чего вычислены метрики качества. Дискретная метрика качества ранжирования $NDCG@n$ оценивает качество первых n документов ответа системы. Общепринятыми метриками для оценки качества ранжирования являются $NDCG@1$, $NDCG@5$ и $NDCG@10$. Однако для более детального рассмотрения изменений метрики в зависимости были рассмотрены значения $NDCG@n$ для всех n от 1 до 10. Для сравнения эффективности разработанного алгоритма с существующими решениями проведен сравнительный анализ с использованием ранжирующих алгоритмов, предоставленных в коллекциях LETOR 3.0. Результаты выполнения алгоритмов для тестовых коллекций TD2003 и TD2004 для метрики $NDCG$ представлены на рис. 2. Результаты показывают, что алгоритм *LTR-MGA* превосходит тестовые алгоритмы, причем наиболее высокие значения име-

ются для $NDCG@1$ (на уровне первого документа). Превосходство алгоритма *LTR-MGA* вызвано тем, что, в отличие от рассмотренных в экспериментах тестовых ранжирующих функций, в предлагаемом подходе для оптимизации функции ранжирования именно метрика $NDCG$ используется в качестве целевой функции.

Для того, чтобы оценить качество ранжирования при использовании предлагаемого алгоритма *LTR-MGA* вычислены значения метрик качества ранжирования документов в поисковой выдаче (рис. 3). Сравнение результатов ранжирования (табл. 2) при использовании базовой ранжирующей функции, базового алгоритма *LTR-GA* и модифицированного алгоритма *LTR-MGA* свидетельствует о преимуществе последнего.

Кроме того, при исследовании выполнена оценка времени, требуемого для МО ранжированию. Это необходимо для подтверждения того, что предложенный метод *LTR-MGA* превосходит в этом показателе подход, основанный на использовании традици-

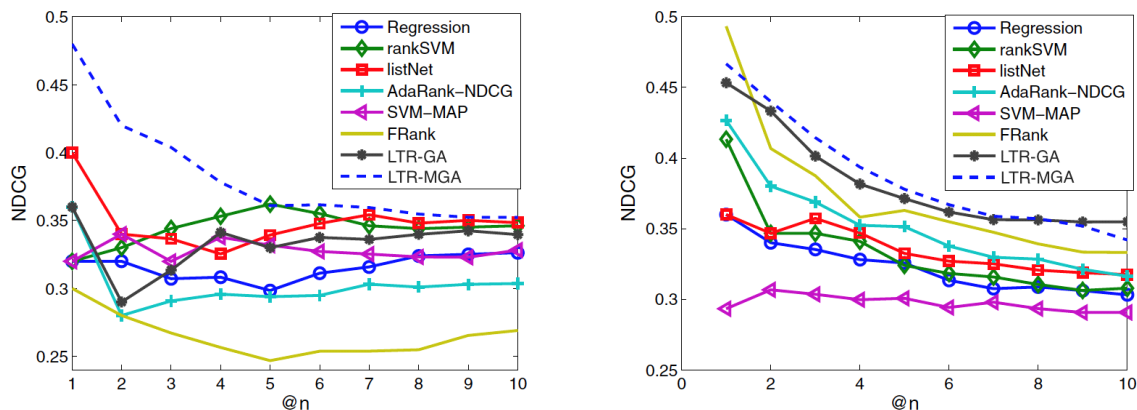


Рис. 2. Сравнение алгоритмов машинного обучения ранжированию по метрике NDCG для тестовых коллекций: слева — набор данных Gov03td, справа — набор данных Gov04td

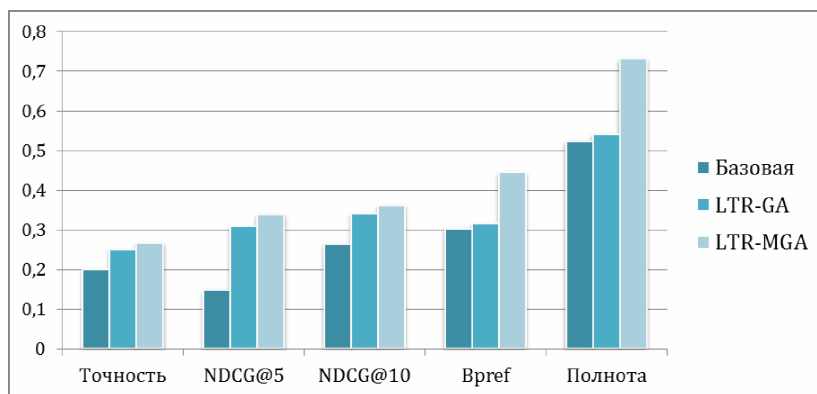


Рис. 3. Оценка метрик качества ранжирования для базовой ранжирующей формулы и алгоритмов обучения LTR-GA и LTR-MGA

Метрики качества ранжирования для разных алгоритмов машинного обучения ранжированию

Таблица 2

Метрика качества ранжирования	Базовая ранжирующая функция	LTR-GA	LTR-MGA	Повышение значения метрики, %
Точность	0,201	0,251	0,267	26,81
NDCG@5 (первые 5 документов)	0,149	0,31	0,339	90,47
NDCG@10 (первые 10 документов)	0,265	0,342	0,362	29,14
Vpref	0,303	0,316	0,446	51,49
Полнота	0,524	0,542	0,732	39,03

*Серым выделены лучшие значения для соответствующей метрики

онного генетического алгоритма (*LTR-GA*). Результаты сравнения временных затрат на выполнение алгоритмов *LTR-GA* и *LTR-MGA* приведены в табл. 3.

7. Заключение. Таким образом, проведенные исследования показали, что при использовании предлагаемого подхода значения рассмотренных метрик ранжирования в ИПС увеличиваются (в среднем на 19,55 % по сравнению с алгоритмом *LTR-GA*). Это подтверждает, что *LTR-MGA* работает корректно и значительно улучшает функцию ранжирования, другими словами — успешно решает задачу оптимизации. С помощью модифицированного алгоритма

за счет применения метода локальной оптимизации и введенных ограничений на количество вычислений целевой функции время машинного обучения снизилось (в среднем на 17,71 % по сравнению с использованием традиционного генетического алгоритма *LTR-GA*).

Разработанный алгоритм машинного обучения ранжированию *LTR-MGA* может быть использован в ИПС, использующих модель ранжирования на основе комбинации простых ранжирующих функций. Однако следует учесть некоторые ограничения по применению предлагаемого подхода. На основе

Оценка времени выполнения машинного обучения ранжированию
в зависимости от размера обучающей выборки

Размер текстовой коллекции документов	Время выполнения LTR-GA	Время выполнения LTR-MGA	Уменьшение времени выполнения, %
1000	19,17	15,42	19,56
5000	54,06	43,96	18,68
10000	84,07	72,18	15,01
50000	261,99	218,32	16,76
100000	652	531,25	18,52
Среднее значение			17,71
*Серым цветом выделены лучшие значения для соответствующего размера тестовой коллекции			

полученных результатов выявлено, что после МО наибольший прирост имеет та метрика качества ранжирования, значение которой принималось в качестве целевой функции. В то же время остальные метрики могут не иметь существенного улучшения, а в некоторых случаях даже ухудшить свои значения. В качестве одного из возможных подходов к устранению этого недостатка предполагается решать задачу оптимизации как многокритериальную: равномерно улучшать несколько основных метрик ранжирования поисковой выдачи, вместо того чтобы оптимизировать одну. Кроме того, в дальнейших исследованиях планируется разработать методику построения целевой функции на основе линейной свертки основных метрик качества ранжирования для улучшения процесса информационного поиска.

Библиографический список

1. Tie-Yan Liu. Learning to Rank for Information Retrieval // Journal Foundations and Trends in Information Retrieval. Vol. 3, issue 3. March 2009. P. 225–331.
2. Christopher J. C. Burges, Tal Shaked, Erin Renshaw. Learning to Rank using Gradient Descent // Proceeding ICML '05 Proceedings of the 22nd international conference on Machine learning. 2005. P. 89–96.
3. Семенихин, С. В. Исследование подходов к машинному обучению ранжированию документов поисковой системой на базе генетических алгоритмов / С. В. Семенихин // Россия молодая: передовые технологии – в промышленность. – 2013. – № 2. – С. 82–85.
4. Многокритериальная оптимизация на основе генетических алгоритмов при синтезе систем управления : моногр. / Л. А. Денисова. – Омск : Изд-во ОмГТУ, 2014. – 170 с. – ISBN 978-5-8149-1822-2.

5. Денисова, Л. А. Автоматизация параметрического синтеза системы регулирования с использованием генетического алгоритма / Л. А. Денисова, В. А. Мещеряков // Автоматизация в промышленности. – 2012. – № 7. – С. 34–38.

6. Huber, Peter J. Robust Estimation of a Location Parameter // Annals of Statistics. – 1964. – № 53. – P. 73–101.

7. Семенихин, С. В. Автоматизация информационного поиска на базе многокритериальной оптимизации и генетических алгоритмов / С. В. Семенихин, Л. А. Денисова // Динамика систем, механизмов и машин. – 2014. – № 3. – С. 224–227.

8. Tie-Yan Liu, Jun Xu, Tao Qin, Wenying Xiong and Hang Li. LETOR: Benchmark Dataset for Research on Learning to Rank for Information Retrieval // SIGIR 2007 Workshop on Learning to Rank for Information Retrieval. – 2007. – С. 3–10.

9. Агеев, М. С. Официальные метрики РОМИП'2004 / М. С. Агеев, И. Е. Кураленок // II Российский семинар по оценке методов информационного поиска (РОМИП 2004), Пущино, 2004 : тр. ; под ред. И. С. Некрестьянова. – СПб. : НИИ химии СПбГУ. – С. 142–150.

10. J. A. Nelder, R. Mead, A simplex method for function minimization, The Computer Journal 7 (1965). 308–313.

СЕМЕНИХИН Святослав Витальевич, аспирант кафедры «Автоматизированные системы обработки информации и управления».

Адрес для переписки: svsemenihin@gmail.com

ДЕНИСОВА Людмила Альбертовна, доктор технических наук, доцент кафедры «Автоматизированные системы обработки информации и управления».

Адрес для переписки: denisova@asoju.com

Статья поступила в редакцию 18.03.2016 г.

© С. В. Семенихин, Л. А. Денисова