

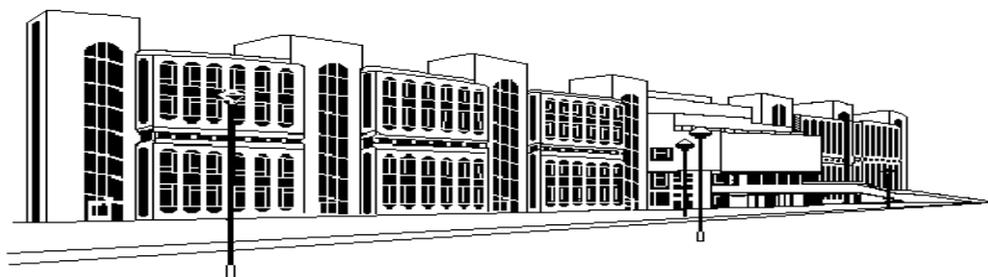


ISSN 2079-0619

МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ
ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ
ГРАЖДАНСКОЙ АВИАЦИИ

НАУЧНЫЙ ВЕСТНИК МГТУ ГА

№ 169



Москва
2011

**ФЕДЕРАЛЬНОЕ АГЕНТСТВО ВОЗДУШНОГО ТРАНСПОРТА
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ
«МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ
ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ
ГРАЖДАНСКОЙ АВИАЦИИ»**

НАУЧНЫЙ ВЕСТНИК МГТУ ГА

№ 169 (7)

Издается с 1998 г.

**Москва
2011**

Научный Вестник МГТУ ГА решением Президиума ВАК Министерства образования и науки РФ включен в перечень ведущих рецензируемых научных журналов и изданий, в которых должны быть опубликованы основные научные результаты диссертаций на соискание ученой степени доктора и кандидата наук.

Главная редакция

Главный редактор - заслуженный юрист РФ, д-р юрид. наук,
проф. Б.П. Елисеев (МГТУ ГА)

Зам. главного редактора - д-р техн. наук, проф. Е.Е. Нечаев (МГТУ ГА)

Ответственный секретарь главной редакции - д-р техн. наук, доц. О.Г. Феоктистова
(МГТУ ГА)

Члены главной редакции - д-р экон. наук, проф. Б.В. Артамонов (МГТУ ГА);
д-р филос. наук, проф. О.Д. Гаранина (МГТУ ГА);
д-р экон. наук, проф. В.А. Казаков (МГУ им. М.В. Ломоносова);
заслуженный деятель науки РФ, д-р техн. наук,
проф. В.Т. Калугин (МГТУ им. Н.Э. Баумана);
заслуженный деятель науки и техники РФ,
д-р физ.-мат. наук, проф. А.И. Козлов (МГТУ ГА);
д-р техн. наук, проф. В.Л. Кузнецов (МГТУ ГА);
д-р техн. наук, проф. С.В. Кузнецов (МГТУ ГА);
заслуженный деятель науки и техники РФ,
д-р физ.-мат. наук, проф. Д.С. Лукин (МФТИ);
д-р техн. наук, проф. В.В. Соломенцев (НТЦ «Промтехаэро»);
заслуженный деятель науки РФ, д-р техн. наук,
проф. В.Г. Ципенко (МГТУ ГА);
д-р техн. наук, проф. В.С. Шапкин (ГосНИИ ГА)

Редакционная коллегия выпуска

Ответственный редактор выпуска - д-р техн. наук, проф. В.Л. Кузнецов

Ответственный секретарь выпуска - канд. физ.-мат. наук, доц. П.В. Филонов

Члены редакционной коллегии - д-р техн. наук, проф. А.А. Егорова;
д-р техн. наук, проф. Л.Е. Рудельсон;
д-р физ.-мат. наук, проф. А.В. Пантелеев;
д-р техн. наук, проф. В.В. Соломенцев

E-mail: kuznetsov@mstuca.ru

тел. +7(495) 458-84-16

ISBN 978-5-86311-787-4

Плата за публикацию в Научном Вестнике МГТУ ГА с аспирантов не взимается

**НАУЧНЫЙ ВЕСТНИК
МОСКОВСКОГО ГОСУДАРСТВЕННОГО
ТЕХНИЧЕСКОГО УНИВЕРСИТЕТА
ГРАЖДАНСКОЙ АВИАЦИИ**

№ 169

2011

СОДЕРЖАНИЕ

Пантелеев А.В., Дмитраков И.Ф. Применение метода дифференциальной эволюции и его модификаций в задаче поиска оптимального управления дискретными детерминированными системами.....	5
Пантелеев А.В., Метлицкая Д.В. Приближенный синтез оптимального управления дискретными системами с помощью генетических алгоритмов с вещественным кодированием.....	13
Кузнецов В.Л., Рудковский А.С. Аномальные продольные резонансы в фотонных кристаллах типа “woodpile”.....	20
Костоглотов А.А., Андрашитов Д.С., Оруджева А.А., Дерябкин И.В. Оценка индекса трендового прогнозирования на основе итерационной регуляризации.....	26
Бадалова А.Г., Пантелеев П.А. Применение нечеткого вывода для определения типа кризиса в адаптивной системе стратегического управления риском.....	33
Зырянов Ю.Т., Коновалов О.А., Малыков К.А. Система управления рациональным распределением ресурсов организационно-технической системы.....	41
Быкова В.В., Моисеев О.Н., Рудельсон Л.Е., Черникова М.А. Анализ сезонного расписания на соответствие нормативам пропускной способности.....	48
Фридман Г.М., Зенкова Н.А. Решение задачи расстановки парка воздушных судов по рейсам полетного расписания с помощью цепочек рейсов.....	56
Кубицкий В.И. Восстановление стёртых пакетов в компьютерных сетях.....	65
Терентьев А.И. Принципы оценки состояния системы информационной безопасности авиапредприятия.....	73
Ерзакова Н.А. О системе нелинейных интегральных уравнений типа Гаммерштейна.....	78
Фонарев А.А. О минимизации выпуклых функционалов.....	85

КРАТКИЕ СООБЩЕНИЯ

Бояркин С.В. Математическая модель отражения электромагнитного поля от гофрированной поверхности с увлажненным верхним слоем.....	90
Кузнецов В.Л., Чепурина А.А. Об одной аналитической модели пропускной способности аэропорта.....	94
Касимов Ю.Ф., Лузгина А.С. Двухпериодная модель перекрывающихся поколений с экзогенным предложением труда: эффекты замещения и дохода в условиях солидарной пенсионной системы.....	99
Дворак А.В., Ивенина Е.М., Филимонов С.В. Модифицированный метод дискретных вихрей для решения сингулярных интегральных уравнений на отрезке.....	103
Бахиркин М.В., Канадин В.Н., Христофоров Г.Ю. Использование специализированного комплекса SCADe для разработки программного обеспечения бортовой функции поддержки принятия решений.....	107
Соломенцев В.В., Романов Н.В. Вопросы построения комплексной системы обработки аэронавигационных данных.....	111

Этот выпуск Научного Вестника посвящен вопросам, объединяемым лаконичным и, в то же время, очень емким направлением – «Прикладная математика и информатика». В публикуемых здесь статьях рассматриваются различные аспекты математического моделирования, являющегося научной основой любого познавательного процесса, исследуется функционирование сложных технических систем, изучаются возможности их оптимизации. Много внимания уделяется разработке и развитию новых технологических решений, лежащие в основе исследования математических моделей. Мы публикуем здесь работы, где рассматриваются вопросы, связанные с продвижением и развитием новых информационных технологий в решении разнообразных научных задач, стоящих перед отраслью.

Широкий спектр проблем, поднимаемых в публикациях этого выпуска, является естественным отображением глубокого проникновения как самой идеологии математического моделирования, так и современных методов и средств вычислительной техники в современные прикладные научные исследования.

В представленном выпуске Научного Вестника традиционно публикуются статьи, посвященные различным аспектам планирования и использования воздушного пространства, обработки полетной информации. Обсуждаются новые подходы к проектированию программного обеспечения автоматизированных систем управления воздушным движением. Оригинальным видится подход к задаче о пропускной способности аэропорта, записанный в формализме ферми-систем.

В работах, посвященных задачам оптимального управления, читатели могут познакомиться с интересными возможностями генетических алгоритмов с вещественным кодированием, а также с применением метода дифференциальной эволюции и его модификацией в задаче поиска оптимального управления.

Близкие по духу, но не математической постановке, вопросы рассматриваются в ряде других статей, направленных на поиск рациональных решений для организационно-технических систем.

Используемому методу инвариантного погружения, сводящему краевые задачи к задаче Коши, посвящены две публикуемые работы в области прикладной электродинамики и дистанционного зондирования.

В этом выпуске Научного Вестника также публикуются работы, посвященные исследованию математических проблем, возникающих в задачах прикладной математики.

В Вестнике представлены работы преподавателей, научных сотрудников и аспирантов МГТУ ГА, МАИ, МФТИ, ВВИА им. Н.Е. Жуковского, ГосНИИ «Аэронавигация», ГосНИИАС, РУДН, ТТГУ, СПбГМТУ и других вузов и научных центров России.

Надеемся, что специалисты, занимающиеся вопросами, затронутыми в публикациях Вестника, найдут здесь интересующую их информацию. Представленные материалы будут полезны также аспирантам и студентам, специализирующимся в области информационных технологий и математического моделирования.

Ответственный редактор –

*доктор технических наук, профессор,
зав. кафедрой прикладной математики МГТУ ГА*

В.Л. Кузнецов

УДК 519.8

ПРИМЕНЕНИЕ МЕТОДА ДИФФЕРЕНЦИАЛЬНОЙ ЭВОЛЮЦИИ И ЕГО МОДИФИКАЦИЙ В ЗАДАЧЕ ПОИСКА ОПТИМАЛЬНОГО УПРАВЛЕНИЯ ДИСКРЕТНЫМИ ДЕТЕРМИНИРОВАННЫМИ СИСТЕМАМИ

А.В. ПАНТЕЛЕЕВ, И.Ф. ДМИТРАКОВ

Сформирован алгоритм применения метода дифференциальной эволюции и его модификаций в задаче поиска оптимального управления дискретными детерминированными системами, создана программная среда визуализации его работы, позволяющая оценить его сравнительную эффективность.

Ключевые слова: метод дифференциальной эволюции, оптимальное управление, дискретные системы, метаэвристические методы.

Введение

Отличительной особенностью решения актуальных задач анализа и синтеза сложных детерминированных систем управления является высокая трудоемкость вычислений при использовании прямых методов оптимизации. Поэтому в настоящее время большое внимание уделяется метаэвристическим методам глобальной оптимизации [1-5], позволяющим получить решение «высокого качества» за приемлемое (с практической точки зрения) время. К метаэвристическим методам относятся и алгоритмы дифференциальной эволюции [1-5], основанные на анализе эволюционных процессов. Особенностью алгоритмов дифференциальной эволюции является использование различий между индивидами (значениями аргумента целевой функции), реализованное линейным оператором, называемым «дифференциация».

Целью данной работы является создание алгоритмического и программного обеспечения применения метода дифференциальной эволюции и его модификаций в задаче оптимального управления дискретными детерминированными системами, разработка комплекса программных средств, позволяющего получить характеристики тестируемого метода оптимизации и делающего процесс сравнения его с другими методами удобным и наглядным.

Методы решения поставленной задачи основываются на алгоритме, изложенном в [1], и модификациях, изложенных в [5].

1. Постановка задачи

Поведение модели объекта управления описывается разностным уравнением

$$x(k+1) = f(k, x(k), u(k)), \quad k = 0, 1, \dots, N-1, \quad (1)$$

где x – вектор состояния системы, $x \in \mathbb{R}^n$; u – вектор управления, $u(k) \in U(k) \subseteq \mathbb{R}^q$; $U(k) = [a_{1,k}, b_{1,k}] \times [a_{2,k}, b_{2,k}] \times \dots \times [a_{q,k}, b_{q,k}]$ – множество допустимых значений управления; k – дискретное время, $k \in T = [0, 1, \dots, N-1]$, число шагов N задано; $f(k, x, u)$ – непрерывная вектор-функция, $f(k, x, u) = (f_1(k, x, u), \dots, f_n(k, x, u))^T$.

Начальное состояние системы (1) задано $x(0) = x_0$. Правый конец траектории свободен. Предполагается, что при управлении используется информация только о дискретном времени k , т.е. применяется так называемое программное управление.

Множество допустимых процессов $\mathcal{D}(0, x_0)$ – это множество пар $d = (x(\cdot), u(\cdot))$, включаю-

щих траекторию $x(\cdot) = \{x_0, x(1), \dots, x(N)\}$ и управление $u(\cdot) = \{u(0), u(1), \dots, u(N-1)\}$, удовлетворяющих уравнению (1) с начальным условием $x(0) = x_0$.

На множестве $\mathcal{D}(0, x_0)$ определен функционал качества управления

$$I(d) = \sum_{k=0}^{N-1} f^0(k, x(k), u(k)) + F(x(N)), \quad (2)$$

где $f^0(k, x, u)$, $F(x)$ – заданные непрерывные функции.

Требуется найти такую пару $d^* = (x^*(\cdot), u^*(\cdot)) \in \mathcal{D}(0, x_0)$, что

$$I(d^*) = \min_{d \in \mathcal{D}(0, x_0)} I(d). \quad (3)$$

2. Стратегия поиска решения

При решении задачи (3) будем искать оптимальное управление $u^*(\cdot)$ и, как следствие, оптимальную траекторию $x^*(\cdot)$. Пусть вектор V – блочный столбец, состоящий из значений управления на каждом шаге $V = (V_0, \dots, V_k, \dots, V_{N-1})^T = (u(0), \dots, u(k), \dots, u(N-1))^T$, где блок $V_k = (V_{1,k}, V_{2,k}, \dots, V_{q,k})^T = u(k) = (u_1(k), u_2(k), \dots, u_q(k))^T$ – управление на k -м шаге, $k = 0, 1, \dots, N-1$.

Будем обозначать через $I(V)$ значение критерия (2), соответствующее V . Для его вычисления сначала ищется траектория системы $x(\cdot) = \{x(0), x(1), \dots, x(N)\}$ из уравнения состояния (1) с известными значениями управления – компонентами вектора V и начальным условием. Далее вычисляется значение критерия (2).

Поясним основную стратегию метода. Сначала на множестве допустимых управлений $\mathcal{U}_0 = U(0) \times U(1) \times \dots \times U(N-1)$ случайным образом генерируется конечный набор $C_0 = \{V^m, m = 1, 2, \dots, NP\}$ векторов, называемый начальной популяцией, где NP – размер популяции. Далее идет циклический процесс замены текущей популяции новой. Он заканчивается, когда количество сформированных популяций оказывается равным заданному максимальному числу популяций.

Для формирования новой популяции последовательно выбирается каждый элемент текущей популяции (он называется вектором-мишенью V^t), и принимается решение, остается ли он в новой популяции или его заменяет сгенерированный специальным образом вектор-образец.

3. Алгоритм

Шаг 1. Задать размер популяции NP ; весовой коэффициент F ; параметр операции скрещивания CR ; максимальное количество популяций M .

Шаг 2. Сгенерировать с помощью датчика равномерного распределения на множестве $\mathcal{U}_0 = U(0) \times U(1) \times \dots \times U(N-1)$ начальную популяцию – NP векторов V^1, V^2, \dots, V^{NP} . Положить $m = 0$.

Шаг 3. Положить $h = 1$ (вектором-мишенью выбран вектор $V^t = V^1$).

Шаг 4. Из числа оставшихся членов популяции случайным образом выбрать три различных члена V^a, V^b, V^c , отличающихся от вектора-мишени V^t .

Шаг 5. Сформировать вектор V^{c1} : $V^{c1} = V^c + F(V^a - V^b)$. Если в полученном векторе V^{c1} координата $V_{s,k}^{c1} \notin [a_{s,k}, b_{s,k}]$, то сгенерировать с помощью равномерного распределения на отрезке $[a_{s,k}, b_{s,k}]$ случайную величину и $V_{s,k}^{c1}$ приравнять ей.

Шаг 6. Сформировать вектор–образец V^β .

Для этого положить $i = 1, j = 0$ и выполнить следующие действия:

а) сгенерировать при помощи равномерного распределения на $[0;1]$ случайное число $\delta_{i,j}$;

б) если $i = q$ и $j = N - 1$, то $V_{i,j}^\beta = V_{i,j}^{c1}$; если $\delta_{i,j} \leq CR$, то $V_{i,j}^\beta = V_{i,j}^{c1}$; если $\delta_{i,j} > CR$, то $V_{i,j}^\beta = V_{i,j}^t$;

в) проверить выполнение неравенства $i \geq q$:

– если оно выполнено, то положить $i = 1$ и перейти к шагу 6 г;

– если нет, положить $i = i + 1$ и перейти к шагу 6 а;

г) проверить выполнение неравенства $j \geq N - 1$:

– если оно выполнено, то процедуру формирования вектора V^β завершить;

– если нет, положить $j = j + 1$ и перейти к шагу 6 а.

Шаг 7. Формирование новой популяции. Подсчитать значения функционала для вектора–образца V^β и вектора–мишени V^t . Сравнить полученные величины $I(V^\beta)$ и $I(V^t)$:

а) если $I(V^\beta) < I(V^t)$, то поместить вектор–образец V^β в новую популяцию;

б) если $I(V^\beta) \geq I(V^t)$, то поместить вектор–мишень V^t в новую популяцию.

Шаг 8. Проверить выполнение неравенства $h \geq NP$:

а) если оно выполнено, то перейти к шагу 9;

б) если нет, то положить $h = h + 1$, $V^t = V^h$ и перейти к шагу 4.

Шаг 9. Проверить условие окончания:

а) если число сформированных популяций $m < M$, то положить $m = m + 1$ и перейти к шагу 3;

б) если $m = M$, то процесс закончить. Выбрать в полученной популяции вектор, которому соответствует наименьшее значение целевой функции, и считать его приближенным решением задачи.

Замечания

1. Обычно полагают $NP = (5 \div 15) \cdot N \cdot q$; $0,2 \leq F \leq 1$; $0 \leq CR \leq 1$. При $N \cdot q \geq 10$ для быстрого получения приблизительного решения рекомендуется брать $0,5 \leq F \leq 0,8$, для улучшения точности – увеличить M совместно с понижением F : $0,2 \leq F \leq 0,4$. Также рекомендуется начинать с $CR = 0,9$ или $CR = 1$ перед тем как попробовать $CR = 0,1$ или $CR = 0$.

2. Если при использовании рекомендованных параметров полученные результаты не обеспечивают достаточной точности, рекомендуется применить метод еще раз с теми же параметрами или задать большее значение NP или (и) M .

3. В пунктах 3–5 осуществляется операция мутации, а в пункте 6 – скрещивания.

4. На шаге 5 алгоритма реализуется оператор «дифференциация».

5. Изложенное в [5] обобщение схем метода дифференциальной эволюции также можно распространить на решение задачи (3). На шаге 5 предлагается использовать формулу

$$V^{c1} = \beta + F \delta,$$

где β – базисный вектор; δ – вектор различий (направление поиска); F – коэффициент влияния.

Вектор β получается двумя способами:

а) в популяции случайным образом выбирается элемент без учета информации о значении критерия (2);

б) в популяции находится элемент V_{best} с наилучшим значением критерия (2).

Вектор δ определяется также двумя способами:

а) различиями между случайно выбранными элементами (двумя, четырьмя и т.д.);

б) направлением (оно принимается за «хорошее» направление поиска, имитирующее градиентное, поскольку характеризуется различиями между наилучшим элементом V_{best} и случайно выбранным из популяции).

В результате получены четыре группы методов.

Первая группа. Базисный вектор и направление поиска определяются по случайно выбранным элементам текущей популяции без учета значений критерия (2):

а) в популяции выделяются два класса C' и C'' , содержащих n' и n'' элементов соответственно;

б) находятся барицентры каждого класса: $BR_{C'} = \frac{1}{n'} \sum_{i=1}^{n'} V^i$, $BR_{C''} = \frac{1}{n''} \sum_{i=1}^{n''} V^i$, где V^i – элементы популяции, входящие в данные классы;

в) базисный вектор β находится одним из двух способов: $\beta \in C' \cup C''$ или $\beta \notin C' \cup C''$;

г) сгенерированный вектор находится по формуле

$$V^{c1} = \beta + F (BR_{C''} - BR_{C'}),$$

где F выбирается из множества $(-1, 0) \cup (0, 1)$.

Вторая группа. По сравнению с формулой п. «г» первой группы методов в качестве базисного вектора выбирается наилучший элемент $V_{best C'}$ из класса C'

$$V^{c1} = V_{best C'} + F (BR_{C''} - BR_{C'}).$$

Третья группа. Вектор различий формируется с учетом информации о величине критерия (2):

а) в популяции выбираются два класса C_+ и C_- с числом элементов n_+ и n_- соответственно так, что значение критерия для каждого элемента из класса C_+ меньше значения критерия для любого элемента из класса C_- : $I(V^i) \leq I(V^j)$, $i = 1, \dots, n_+$; $j = 1, \dots, n_-$;

б) в каждом классе находятся «минимальный» и «максимальный» элементы: $V_{C_+}^{\min}$, $V_{C_+}^{\max}$, $V_{C_-}^{\min}$, $V_{C_-}^{\max}$, для которых выполняются условия

$$I(V_{C_+}^{\min}) \leq I(V^i) \leq I(V_{C_+}^{\max}) \quad \forall V^i \in C_+; \quad I(V_{C_-}^{\min}) \leq I(V^i) \leq I(V_{C_-}^{\max}) \quad \forall V^i \in C_-;$$

в) в каждом классе находятся приращения $BR_S^+ = V_{C_+}^{\min} - V_{C_+}^{\max}$, $BR_S^- = V_{C_-}^{\min} - V_{C_-}^{\max}$ и среднее приращение $BR_S = \frac{BR_S^+ + BR_S^-}{2}$;

г) находятся барицентры классов C_+ и C_- : $BR_{C_+} = \frac{1}{n_+} \sum_{i=1}^{n_+} V^i$, $BR_{C_-} = \frac{1}{n_-} \sum_{i=1}^{n_-} V^i$;

д) сгенерированный вектор находится по формуле

$$V^{c1} = BR_{C_+} + F (BR_{C_+} - BR_{C_-} + BR_S).$$

Заметим, что по сравнению с классической формулой параметр F может принимать как положительные, так и отрицательные значения.

Четвертая группа. В формуле п. «д» третьей группы в качестве базисного вектора используется наилучший элемент $V_{best C_+}$ в классе C_+

$$V^{c1} = V_{best C_+} + F (BR_{C_+} - BR_{C_-} + BR_S).$$

Выделение описанных четырех групп позволяет в рамках каждой группы выделять еще несколько различных стратегий поиска.

4. Программное обеспечение

Разработана программа поиска оптимального управления, реализующая вышеописанный метод и основные его модификации. Среда разработки – Microsoft Visual Studio, язык программирования – C#.

Разработан интерфейс пользователя, позволяющий: вводить параметры постановки задачи; легко управлять параметрами тестируемого метода; анализировать эффективность его работы на модельных примерах; «видеть» процесс его работы; анализировать сходимость метода; просмотреть промежуточные результаты счета; удобно проводить сравнение выбранного метода с другими методами глобальной оптимизации.

Уравнение состояния оптимизируемых систем в большинстве случаев полагалось линейным

$$x(k+1) = Ax(k) + Bu(k), \quad k = 0, 1, \dots, N-1, \quad (4)$$

где A – матрица размером $n \times n$; B – матрица размером $n \times q$.

Функционал качества управления считался квадратичным или линейно-квадратичным (для случая $n = 1$):

$$I_1(d) = \sum_{k=0}^{N-1} \left(x^T(k)Sx(k) + u^T(k)Qu(k) \right) + Lx(N) \quad (n = 1), \quad (5)$$

$$I_2(d) = \sum_{k=0}^{N-1} \left(x^T(k)Sx(k) + u^T(k)Qu(k) \right) + x^T(N)Lx(N), \quad (6)$$

где S – матрица размером $n \times n$; Q – матрица размером $q \times q$; L – матрица размером $n \times n$ или число при $n = 1$.

5. Тестовые примеры

Пример 1. Даны модель объекта управления

$$\begin{aligned} x_1(k+1) &= x_1(k) + u(k), \\ x_2(k+1) &= 2x_1(k) + x_2(k), \quad k = 0, 1, \end{aligned}$$

где $x \in \mathbb{R}^2$, $|u(k)| \leq 10^5$, начальные условия $x_1(0) = 2$, $x_2(0) = 1$ и функционал

$$I = \sum_{k=0}^1 \left[x_1^2(k) + x_2^2(k) + u^2(k) \right] \rightarrow \min.$$

Требуется найти минимальное значение функционала и оптимальный процесс $(x^*(\cdot), u^*(\cdot))$, на котором это значение достигается.

Аналитическое решение задачи: $u^*(0) = -1$, $u^*(1) = 0$, $x_1^*(0) = 2$, $x_1^*(1) = x_2^*(1) = 1$, $x_2^*(0) = 1$, $x_2^*(1) = 5$, $x_2^*(2) = 7$, $\min I = 32$.

Результат работы программы представлен на рис. 1: $u^*(0) = -1$, $u^*(1) = 0$, $\min I = 32$. В данном случае найденное с помощью программного обеспечения оптимальное управление совпадает с точным значением.

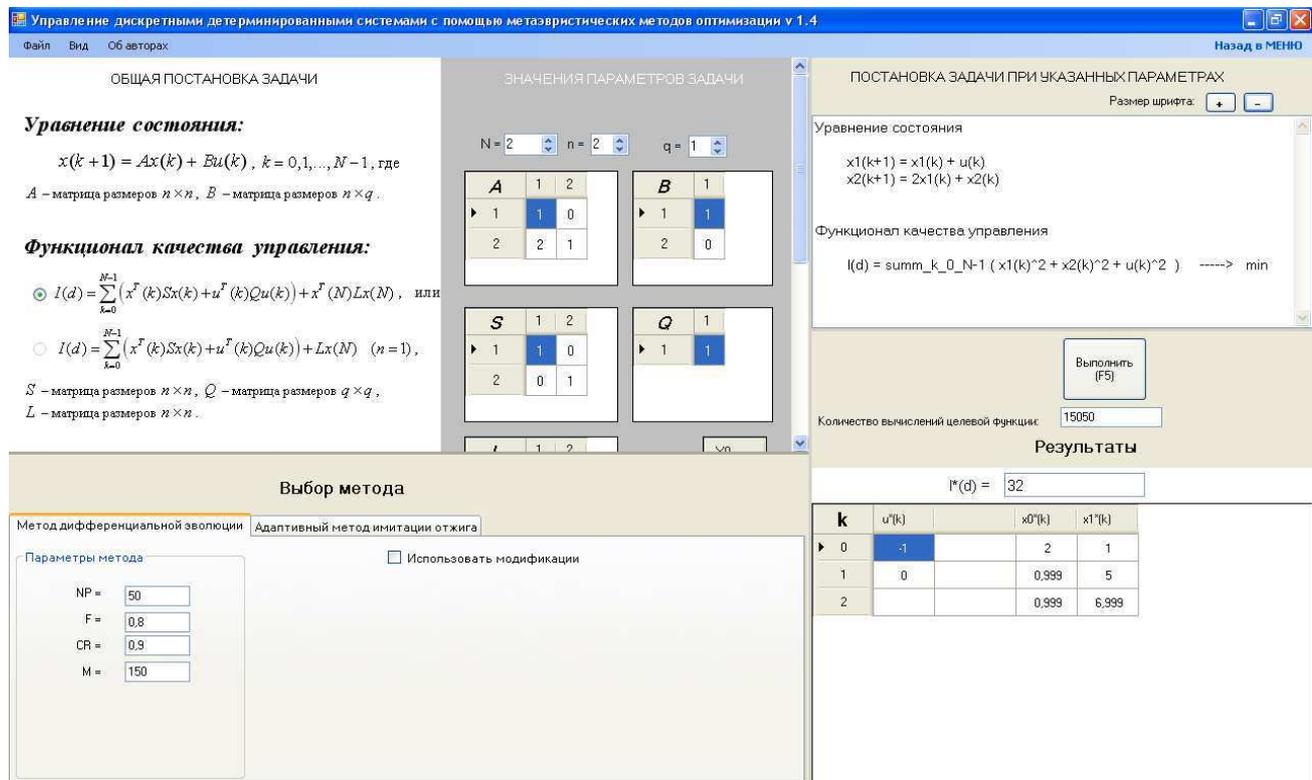


Рис. 1. Ввод параметров постановки задачи

Пример 2. Даны модель объекта управления $x(k+1) = x(k) + u(k)$, начальное состояние $x(0) = 3$, $k = 0, 1, \dots, N-1$, где $x \in \mathbb{R}$, $|u(k)| \leq 10^2$, и функционал

$$I = \sum_{k=0}^{N-1} [x^2(k) + u^2(k)] \rightarrow \min.$$

Требуется найти минимальное значение функционала и оптимальный процесс $(x^*(\cdot), u^*(\cdot))$, на котором это значение достигается.

Будем искать численное решение, совпадающее с точным до 3-го знака после запятой, с помощью классического метода дифференциальной эволюции и четырех его модификаций, описанных в п. 5 замечаний к алгоритму. Приведем табл. 1, в которой покажем, сколько раз в процессе решения избранным методом понадобилось вычислять функционал качества управления для достижения требуемой точности.

Пусть $N = 10$. Аналитическое решение задачи $\min I = 60885/4181 \cong 14,562$.

Всеми методами (табл. 1) получено точное решение с требуемой точностью. Высокое, по сравнению с обычным методом, количество вычислений критерия при использовании групп модификаций 3 и 4 обусловлено тем, что каждый раз при формировании вектора V^{c1} требуется упорядочивать элементы популяции по возрастанию значения критерия, что вызывает многократное вычисление критерия. При использовании модификаций группы 2 для формирования вектора V^{c1} требуется находить элемент класса C' , соответствующий наименьшему значению критерия, что также увеличивает количество вычислений критерия.

Таблица 1

Метод	Параметры метода					Количество вычислений функционала
	<i>NP</i>	<i>F</i>	<i>CR</i>	<i>M</i>	дополнительные	
Классический	45	0,4	0,9	190		$1,7 \cdot 10^4$
1-я группа	70	-0,9	0,5	300	$n' = 10, n'' = 10$	$4,0 \cdot 10^4$
2-я группа	70	-0,9	0,5	200	$n' = 10, n'' = 10$	$17,0 \cdot 10^4$
3-я группа	70	-0,9	0,5	300	$n_+ = 10, n_- = 10$	$46,0 \cdot 10^4$
4-я группа	50	0,5	0,5	500	$n_+ = 20, n_- = 20$	$100,0 \cdot 10^4$

Можно заключить, что использование групп модификаций 1-4 связано с большими вычислительными затратами. Поэтому, если условия задачи не требуют высокой точности численного решения, можно рекомендовать использовать классический метод.

Пример 3. Даны модель объекта управления

$$x_1(k+1) = \frac{x_1(k)}{1+0,01 \cdot u_1(k)(3+u_2(k))}, \quad x_2(k+1) = \frac{x_2(k) + \frac{u_1(k)x_1(k)}{1+0,01 \cdot u_1(k)(3+u_2(k))}}{1+u_1(k)(1+u_2(k))},$$

$$x_3(k+1) = \frac{x_3(k)}{1+0,01 \cdot u_2(k)(1+u_3(k))},$$

начальное состояние $x(0) = [2; 5; 7]^T$, $k = 0, 1, \dots, N-1$, где $x \in \mathbb{R}^3$; ограничения на управление: $0 \leq u_1(k) \leq 4$, $0 \leq u_2(k) \leq 4$, $0 \leq u_3(k) \leq 0,5$; и функционал

$$I = x_1^2(N) + x_2^2(N) + x_3^2(N) + \left[\left(\sum_{k=0}^{N-1} x_1^2(k) + x_2^2(k) + 2u_3^2(k) \right) \left(\sum_{k=1}^N x_3^2(k) + 2u_1^2(k) + 2u_2^2(k) \right) \right]^{\frac{1}{2}} \rightarrow \min.$$

Требуется найти минимальное значение функционала и оптимальный процесс $(x^*(\cdot), u^*(\cdot))$, на котором это значение достигается. Для $N = 100$ аналитическое решение задачи, найденное в [8] $\min I \cong 258,339$; численные решения, полученные с помощью разработанного программного обеспечения, представлены в табл. 2.

Таблица 2

№ п/п	Параметры метода				min <i>I</i>
	<i>NP</i>	<i>F</i>	<i>CR</i>	<i>M</i>	
1	2000	0,7	0,90	2000	358,787
2	2000	0,3	0,30	4000	258,341
3	2000	0,3	0,20	4000	258,340
4	2000	0,3	0,15	4000	258,340

Для выбранного значения N с помощью разработанного программного обеспечения получены решения, близкие к точному. В случае 1 задано большое значение параметра F , что привело к увеличению погрешности.

Заклучение

В данной работе создан алгоритм и программное обеспечение применения метода дифференциальной эволюции и его модификаций в задаче поиска оптимального управления дискретными детерминированными системами. Эффективность разработанного программного обеспечения продемонстрирована на модельных примерах.

ЛИТЕРАТУРА

1. **Пантелеев А. В., Дмитраков И. Ф.** Применение метода дифференциальной эволюции для оптимизации параметров аэрокосмических систем // Электронный журнал "Труды МАИ" - 2010. - № 37.
2. **Пантелеев А.В.** Метаэвристические алгоритмы поиска глобального экстремума. - М.: МАИ-Принт, 2009.
3. **Дмитраков И.Ф.** Анализ эффективности метаэвристических методов поиска глобального экстремума функций многих переменных // Актуальные проблемы авиационных и аэрокосмических систем: процессы, модели, эксперимент. - 2009. - Т. 14. - Вып. 2 (29).
4. **Price K., Storn R., Lampinen J.** Differential evolution - a practical approach to global optimization, Springer, 2005.
5. **Feoktistov V.** Differential Evolution In Search of Solutions, Springer, 2006.
6. **Пантелеев А.В., Легова Т.А.** Методы оптимизации в примерах и задачах. - М.: Высшая школа, 2008.
7. **Пантелеев А.В., Бортакровский А.С.** Теория управления в примерах и задачах. - М.: Высшая школа, 2003.
8. **Luus R.** Iterative Dynamic Programming, London, UK: Chapman & Hall/CRC, 2000.

APPLICATION OF DIFFERENTIAL EVOLUTION METHOD TO THE DISCRETE TIME DETERMINISTIC SYSTEMS OPTIMAL CONTROL PROBLEM

Panteleyev A.V., Dmitrakov I.F.

Detailed algorithms of application of the method of differential evolution and its modifications to the problem of optimal control of the discrete time deterministic system are suggested; the software environment for visualization of algorithms, allowing to estimate its comparative efficiency, is developed.

Key words: the differential evolution method, optimal control, the discrete time systems, the metaheuristic methods.

Сведения об авторах

Пантелеев Андрей Владимирович, 1955 г.р., окончил МГТУ им. Н.Э. Баумана (1978), доктор физико-математических наук, профессор, заведующий кафедрой математической кибернетики МАИ, автор более 130 научных работ, область научных интересов - методы синтеза оптимальных нелинейных систем управления, методы оптимизации.

Дмитраков Игнатий Филиппович, 1988 г.р., студент МАИ, автор 5 научных работ, область научных интересов – анализ и оценка эффективности сложных систем, методы оптимизации.

УДК 517.97

ПРИБЛИЖЕННЫЙ СИНТЕЗ ОПТИМАЛЬНОГО УПРАВЛЕНИЯ ДИСКРЕТНЫМИ СИСТЕМАМИ С ПОМОЩЬЮ ГЕНЕТИЧЕСКИХ АЛГОРИТМОВ С ВЕЩЕСТВЕННЫМ КОДИРОВАНИЕМ

А.В. ПАНТЕЛЕЕВ, Д.В. МЕТЛИЦКАЯ

Предложено решение задачи оптимального управления дискретными детерминированными системами с помощью генетических алгоритмов поиска условного глобального экстремума с вещественным кодированием. Сформирован алгоритм решения поставленной задачи, на основе которого создано программное обеспечение. Приведены примеры, иллюстрирующие эффективность сформированного алгоритма.

Ключевые слова: генетический алгоритм, вещественное кодирование, оптимальное управление.

Введение

В данной работе рассмотрено применение генетических алгоритмов (ГеА) поиска условно-глобального экстремума с вещественным кодированием [1-5] к задаче нахождения оптимального программного управления дискретными детерминированными системами [6]. Предложен алгоритм решения поставленной задачи. На его основе сформирован комплекс программных средств, предназначенный для различных моделей систем управления и различных функционалов качества. Комплекс имеет удобный интерфейс, с помощью которого можно легко менять параметры задачи и алгоритма, а также анализировать эффективность работы алгоритма. Решено несколько тестовых примеров, на которых продемонстрировано, что предложенный алгоритм позволяет подобрать параметры так, чтобы полученное приближенное решение с достаточно высокой точностью совпадало с оптимальным.

1. Постановка задачи

Поведение модели объекта управления описывается разностным уравнением $x(t+1) = f(t, x(t), u(t))$, $t = 0, 1, \dots, N-1$, где x – вектор состояния системы, $x \in R^n$; u – вектор управления, $u \in U(t) \subseteq R^q$, $U(t)$ – некоторое замкнутое выпуклое множество допустимых значений управления; t – дискретное время, $t \in T = [0, 1, \dots, N-1]$, число шагов N задано; $f(t, x, u)$ – непрерывная вектор-функция, $f(t, x, u) = (f_1(t, x, u), \dots, f_n(t, x, u))^T$.

Начальное состояние системы задано $x(0) = x_0$. Правый конец траектории свободен. При управлении используется информация только о дискретном времени t , т.е. применяется так называемое программное управление.

Множество допустимых процессов $D(0, x_0)$ – это множество пар $d = (x(\cdot), u(\cdot))$, включающих траекторию $x(\cdot) = \{x_0, x(1), \dots, x(N)\}$ и управление $u(\cdot) = \{u(0), u(1), \dots, u(N-1)\}$, $u(t) \in U(t)$, удовлетворяющих уравнению состояния и начальному условию.

На множестве допустимых процессов $D(0, x_0)$ определен функционал качества управления $I(d) = \sum_{t=0}^{N-1} f^0(t, x(t), u(t)) + F(x(N))$, где $f^0(t, x, u)$, $F(x)$ – заданные непрерывные функции.

Требуется найти такую пару $d^* = (x^*(\cdot), u^*(\cdot)) \in D(0, x_0)$, что $I(d^*) = \min_{d \in D(0, x_0)} I(d)$.

2. Стратегия поиска решения

Основная идея ГеА состоит в том, что они имитируют в своей работе природные способы оптимизации: генетическое наследование и естественный отбор. Суть механизма естественного отбора заключается в том, что в процессе эволюции выживают и размножаются наиболее приспособленные особи. Благодаря механизму генетического наследования их потомки сохраняют основные качества родителей и, подвергаясь случайным мутациям, приобретают ряд новых свойств. Если новые свойства полезны, то они сохраняются и наследуются.

ГеА рассматривают целевую функцию $f(x) = f(x_1, x_2, \dots, x_n)$, определенную на множестве допустимых решений $D \subseteq R^n$ и находят ее условный глобальный максимум на заданном множестве, т.е. такую точку $x^* \in D$, что $f(x^*) = \max_{x \in D} f(x)$. Задача поиска минимума функции $f(x)$ сводится к задаче поиска максимума путем замены знака перед функцией на противоположный: $f(x^*) = \min_{x \in D} f(x) = -\max_{x \in D} [-f(x)]$.

Рассматриваемая целевая функция $f(x)$ эквивалентна понятию приспособленности живого организма. Будем называть ее функцией приспособленности. Вектор параметров $x = (x_1, x_2, \dots, x_n)^T$ целевой функции называется хромосомой, а отдельные его параметры x_i – генами, $i = 1, 2, \dots, n$.

При решении задачи используются наборы $Y = \{x^k = (x_1^k, x_2^k, \dots, x_n^k)^T, k = 1, 2, \dots, m\} \subset D$ возможных решений, называемые популяциями, где x^k – хромосома с номером k ; m – размер популяции; x_i^k – ген с номером i . Применение ГеА сводится к исследованию множества D . Чем больше значение $f(x^k)$, тем более подходит хромосома x^k в качестве решения.

ГеА имитирует эволюцию начальной популяции как циклический итерационный процесс, во время которого к популяции применяются основные биологические операторы: селекция, скрещивание, мутация. В результате происходит ее смена на новую популяцию, к которой, если решение не найдено, опять применяются биологические операторы, и так далее до выполнения условия окончания. Если хотя бы одно из условий окончания выполнено, то в качестве приближенного решения задачи $f(x^*) = \max_{x \in D} f(x)$ из последней популяции выбирается особь с наибольшим значением функции приспособленности.

При использовании ГеА для решения поставленной задачи будем оптимизировать управление $u(\cdot)$. Так как необходимо найти пару $d^* = (x^*, u^*)$, соответствующую минимальному значению функционала, то будет решаться задача $I(d^*) = \min_{d \in D(0, x_0)} I(d) = -\max_{d \in D(0, x_0)} [-I(d)]$. Функцией приспособленности будет являться функционал $-I(d) = -I(x(\cdot), u(\cdot))$. Хромосома k -й особи в популяции будет представлять собой вектор-строку $u^k = (u^k(0), u^k(1), \dots, u^k(N-1))$. Для того чтобы найти значение функционала $-I(d^k)$, соответствующее паре $d^k = (x^k, u^k)$, необходимо вычислить траекторию системы $x^k = (x_0^k, x^k(1), \dots, x^k(N))$, соответствующую управлению u^k , из уравнения состояния с учетом начального условия.

3. Алгоритм решения задачи

Шаг 1. Создание начальной популяции.

1. Задается номер популяции $p = 0$, максимальное количество популяций p_{\max} , номер итерации $j = 1$, число шагов N , размер популяции m .

2. Случайным образом, используя равномерное распределение на множестве допустимых

значений управления, формируется начальная популяция Y_0 . В результате получаем m начальных вектор-строк $u^k = (u_{1,1}^k, \dots, u_{1,q}^k, u_{2,1}^k, \dots, u_{2,q}^k, \dots, u_{N,1}^k, \dots, u_{N,q}^k) = (u_1^k, u_2^k, \dots, u_n^k)$, где $n = N \cdot q$, $k = 1, 2, \dots, m$ (заметим, что здесь $u_{i,j}^k = u_j^k(i-1)$).

3. Вычисляется траектория x^k , а затем значение функции приспособленности $I_k = -I(x^k, u^k)$, для каждой особи $u^k \in Y_0$, $k = 1, 2, \dots, m$. Вычисляется приспособленность популяции Y_0 в целом $\bar{I} = \sum_{k=1}^m I_k$.

Шаг 2. Селекция – отбор особей (хромосом) u^k в соответствии со значениями функции приспособленности I_k для последующего их скрещивания. Вероятность участия в скрещивании обычно пропорциональна относительной приспособленности $\frac{I_k}{I}$ особи u^k . В ГеА с вещественным кодированием используются следующие операторы селекции: 1) «панмиксия» - случайный равновероятностный отбор, вероятность участия особи в селекции $p_k = \frac{1}{m}$; 2) рулетка - отбирает особей с помощью двух «запусков» рулетки, вероятность участия особи в селекции пропорциональна относительной приспособленности особи $p_k = \frac{I_k}{I}$; 3) турнирный отбор - реализует два турнира, в каждом из которых случайным образом выбираются два элемента из популяции, затем выбирается лучшая особь среди них.

Результатом шага 2 являются две родительские особи u^1 и u^2 .

В ГеА может применяться стратегия элитизма. Ее суть в том, что небольшое количество особей u^k переходит в следующее поколение без изменений, не участвуя в селекции и последующем скрещивании.

Шаг 3. Скрещивание – операция, при которой из двух хромосом (особей), называемых родителями, порождается одна или несколько новых, называемых потомками, путем обмена частями родительских хромосом. Применяются следующие операторы скрещивания: 1) плоский кроссовер - создается один потомок $u^q = (u_1^q, \dots, u_n^q)^T$, где $u_i^q, i = 1, \dots, n$ – случайное число из промежутка $[u_i^1, u_i^2]$, $u_i^2 \geq u_i^1$ или из $[u_i^2, u_i^1]$, $u_i^2 < u_i^1$; 2) простейший кроссовер - из множества $\{1, 2, \dots, n-1\}$ выбирается случайное число p и генерируются два потомка: $u^{q1} = (u_1^1, \dots, u_p^1, u_{p+1}^2, \dots, u_n^2)^T$ и $u^{q2} = (u_1^2, \dots, u_p^2, u_{p+1}^1, \dots, u_n^1)^T$; 3) арифметический кроссовер - создаются два потомка $u^{q1} = (u_1^{q1}, \dots, u_n^{q1})^T$, $u^{q2} = (u_1^{q2}, \dots, u_n^{q2})^T$, где $u_i^{q1} = \eta u_i^1 + (1-\eta) u_i^2$, $u_i^{q2} = \eta u_i^2 + (1-\eta) u_i^1$, $i = 1, 2, \dots, n$, где $\eta \in (0; 1)$; 4) дискретный кроссовер - создается один потомок $u^q = (u_1^q, \dots, u_n^q)^T$, где u_i^q выбирается из двух значений u_i^1 , u_i^2 случайно с вероятностью 0,5 для всех $i = 1, \dots, n$; 5) линейный кроссовер - создаются три потомка u^{q1} , u^{q2} , u^{q3} , где $u_i^{q1} = \frac{u_i^1 + u_i^2}{2}$, $u_i^{q2} = \frac{3u_i^1 - u_i^2}{2}$, $u_i^{q3} = \frac{-u_i^1 + 3u_i^2}{2}$, $i = 1, \dots, n$.

Результатом шага 3 являются потомки u^{chi} , $i = 1, 2, 3$.

Шаг 4. Мутация – преобразование хромосомы, случайно изменяющее обычно один (реже несколько) из ее генов. Оператор мутации предназначен для того, чтобы поддерживать разнообразие особей в популяции. В ГеА с вещественным кодированием применяется случайная мутация. Поочередно рассматривается каждый потомок, полученный в результате скрещивания.

Среди генов u_1, \dots, u_n случайно (с вероятностью $\frac{1}{n}$) выбирается один с номером $p \in \{1, \dots, n\}$, подлежащий замене. Его новое значение u_p^M случайным образом выбирается из промежутка $[a_p(t), b_p(t)]$ изменения выбранной координаты u_p при заданном времени t .

Результатом шага 4 являются потомки - мутанты $u^{M_i}, i = 1, 2, 3$.

Шаг 5. Формирование новой популяции. С равной вероятностью выбирается один из потомков-мутантов, сформированных на предыдущем шаге. Пусть это будет u^{M_1} . Выбранный потомок добавляется в популяцию вместо хромосомы, которой соответствует наименьшее среди элементов популяции значение функции приспособленности, вычисляется траектория x^{M_1} , значение функции приспособленности добавленного потомка $I_{M_1} = -I(x^{M_1}, u^{M_1})$ и приспособленность новой популяции в целом $\bar{I} = \sum_{k=1}^m I_k$.

Проверяются условия: а) если $j < m$, то положить $j = j + 1$ и перейти к шагу 2; б) если $j = m$, то положить $p = p + 1$ и перейти к шагу 6.

Шаг 6. Проверка условия окончания работы ГеА (сформировано заданное число популяций $p = p_{\max}$): а) если условие не выполнено, то полагаем $j = 1$ и переходим к шагу 2; б) если условие окончания работы выполнено, то в качестве решения (приближенного) задачи выбирается особь u^* с наибольшим значением функции приспособленности из текущей популяции. Минимальное значение функционала при этом будет равно $\min I = I(d^*) = I(x^*, u^*)$.

Более подробное описание типов операторов селекции, скрещивания и мутации в ГеА с вещественным кодированием приведено в [3].

4. Программное обеспечение

На основе изложенного алгоритма сформирована программа поиска оптимального управления дискретной детерминированной системой при помощи генетических алгоритмов с вещественным кодированием. Среда разработки Microsoft Visual Studio 2005, язык программирования C#. Программа работает в режиме диалога с пользователем. Пользовательский интерфейс включает основное окно с формами для ввода данных, а также позволяет увидеть графическое изображение получаемых результатов и сохранять их в памяти компьютера. Работа программы начинается с ввода начальных данных: параметров оптимизируемой системы, начального состояния, диапазонов изменения переменных, а также задания параметров алгоритма: характеристик популяции (размера, максимального количества особей), типа операторов селекции, скрещивания, мутации, количества элитных особей. Для удобства реализации на компьютере в постановке задачи были сделаны следующие упрощения.

Уравнение состояния полагалось линейным: $x(t+1) = Ax(t) + Bu(t)$, $t = 0, 1, \dots, N-1$, где A – матрица размером $n \times n$; B – матрица размером $n \times q$. Функционал качества управления считался квадратичным или линейно-квадратичным (для случая $n = 1$)

$$I_1(d) = \sum_{t=0}^{N-1} (x^T(t)S(t)x(t) + u^T(t)Q(t)u(t)) + Lx(N), \quad (n = 1);$$

$$I_2(d) = \sum_{t=0}^{N-1} (x^T(t)S(t)x(t) + u^T(t)Q(t)u(t)) + x^T(N)Lx(N),$$

где $S(t)$ – неотрицательно определенная матрица размером $n \times n$; $Q(t)$ – положительно определенная матрица размером $q \times q$; L – неотрицательно определенная матрица размером $n \times n$ (L – число при $n = 1$).

Множество допустимых значений управления $U(t)$ полагалось равным прямому произведению отрезков $[a_{t,i}, b_{t,i}]$, $i = \overline{1, q}$.

Требуется найти минимальное значение функционала и оптимальный процесс $(x^*(\cdot), u^*(\cdot))$, на котором достигается это значение.

5. Примеры

Пример 1

Заданы модель объекта управления $x(t+1) = x(t) + u(t)$, где $x \in \mathbb{R}$, $|u| < 30$, $t = 0, 1, \dots, N-1$, $N = 4$; начальное состояние $x(0) = 3$ и функционал $I = \sum_{t=0}^{N-1} \frac{1}{t+1} u^2(t) + 2x(N)$.

Требуется найти минимальное значение функционала и оптимальный процесс $(x^*(\cdot), u^*(\cdot))$, на котором это значение достигается.

Параметры постановки задачи: $n = 1$, $q = 1$, $A = 1$, $B = 1$, $S(t) = 0$, $Q(t) = \frac{1}{t+1}$, $L = 2$.

Будем использовать следующие параметры алгоритма: максимальное количество популяций $p_{\max} = 500$; количество особей в популяции $m = 20$; количество элитных особей – 1; оператор селекции – рулетка; оператор скрещивания – арифметический кроссовер; оператор мутации – случайная мутация.

Результаты работы алгоритма представлены в табл. 1 и на рис. 1.

Точное решение задачи имеет вид

$$u^*(t) = -t - 1, \quad x^*(t) = x_0 - \frac{1}{2}t(t+1),$$

$$\min I = 2x_0 - \frac{1}{2}N(N+1).$$

Для $N = 4$ и $x_0 = 3$ решение будет выглядеть

$$\text{так: } u^*(0) = -1, \quad u^*(1) = -2, \quad u^*(2) = -3,$$

$$u^*(3) = -4, \quad x^*(0) = x_0 = 3, \quad x^*(1) = 2,$$

$$x^*(2) = 0, \quad x^*(3) = -3, \quad x^*(4) = -7, \quad \min I = -4.$$

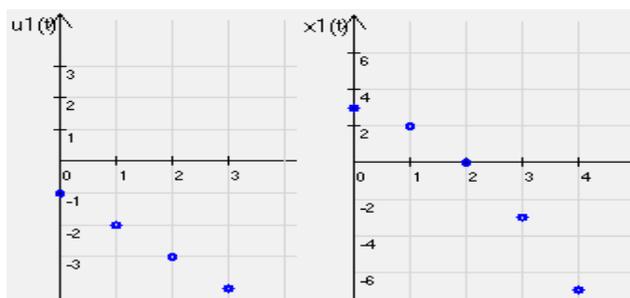


Рис. 1. Графическое изображение результатов

Таблица 1

Сравнивая результаты, полученные при помощи ГеА с бинарным кодированием, с точным решением, можно оценить эффективность алгоритма.

t	0	1	2	3	4
$u^*(t)$	-0,99	-2,0003	-3,0465	-4,0196	
$x^*(t)$	3	2,01	0,0097	-3,0368	-7,0564
$\min I = -3,999$					

Пример 2

Даны модель объекта управления $x_1(t+1) = x_2(t)$, $x_2(t+1) = x_1(t) - u(t)$, где $x \in \mathbb{R}^2$, $|u| \leq 1$, $t = 0, 1$; $x_1(0) = 2$, $x_2(0) = -3$ и функционал $I = x_1^2(2) + x_2^2(2)$.

Требуется найти минимальное значение функционала и оптимальный процесс $(x^*(\cdot), u^*(\cdot))$, на котором это значение достигается. Параметры системы: $N = 2$, $n = 2$, $q = 1$, $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$,

$$B = \begin{pmatrix} 0 \\ -1 \end{pmatrix}, \quad S = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad Q = 0, \quad L = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Пусть параметры алгоритма следующие: количество популяций $p_{\max} = 50$; количество особей в популяции $m = 20$; количество элитных особей – 1; оператор селекции – рулетка; оператор скрещивания – арифметический кроссовер; оператор мутации – случайная мутация.

Результаты работы алгоритма представлены в табл. 2 и на рис. 2.

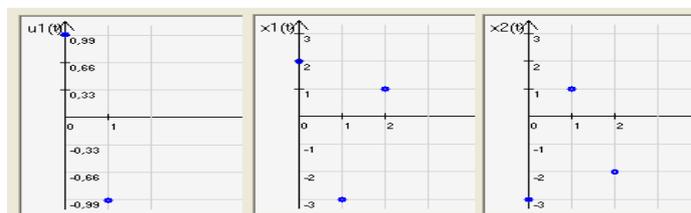


Рис. 2. Результаты работы алгоритма в примере 2

Таблица 2

t	0	1	2
$u^*(t)$	0,989	-0,985	-
$x_1^*(t)$	2	-3	1,011
$x_2^*(t)$	-3	1,011	-2,015
$\min I = 5,081$			

Точное решение задачи для $N = 2$, $x_1(0) = 2, x_2(0) = -3$ будет выглядеть так: $u^*(0) = 1$, $u^*(1) = -1$, $x_1^*(0) = 2, x_2^*(0) = -3, x_1^*(1) = -3, x_2^*(1) = 1, x_1^*(2) = 1, x_2^*(2) = -2$, $\min I = 5$. Сравнивая полученные результаты с точным решением, можно сделать вывод, что ГеА с вещественным кодированием дает результат, близкий к точному. При этом, изменяя параметры алгоритма, можно добиться более высокой точности.

Пример 3

Воспользуемся примером 1, но пусть теперь $N = 20$. Параметры алгоритма прежние. В результате работы получено значение критерия $\min I = -203,846$. Точное минимальное значение критерия для $N = 20$ и $x_0 = 3$ равно -204 . Сравнивая полученные результаты с точным решением, можно сделать вывод, что эффективность работы ГеА с вещественным кодированием падает с увеличением времени функционирования системы. Изменим параметры алгоритма, увеличив количество популяций. Результаты работы алгоритма при $p_{\max} = 3000$ ($\min I = -203,996$) представлены в табл. 3 и на рис. 3. Как видно из примера, при изменении параметров алгоритма точность найденного решения можно увеличить.

Таблица 3

t	0	1	2	3	4	5	6	
$u^*(t)$	-1,0107	-1,9992	-3,0664	-3,9599	-5,0355	-5,9829	-6,9736	
$x^*(t)$	3,0000	1,9893	-0,0098	-3,0762	-7,0362	-12,0717	-18,0546	
t	7	8	9	10	11	12	13	
$u^*(t)$	-8,0820	-9,0732	-10,0065	-10,9995	-12,0226	-13,0346	-13,9602	
$x^*(t)$	-25,0282	-33,1102	-42,1835	-52,1900	-63,1895	-75,2122	-88,2467	
t	14	15	16	17	18	19	20	
$u^*(t)$	-15,0263	-16,0349	-16,9868	-18,0639	-19,0285	-20,0086	--	
$x^*(t)$	-102,2069	-117,2332	-133,2681	-150,2549	-168,3188	-187,3473	-207,3559	
$\min I = -203,995521426833$								

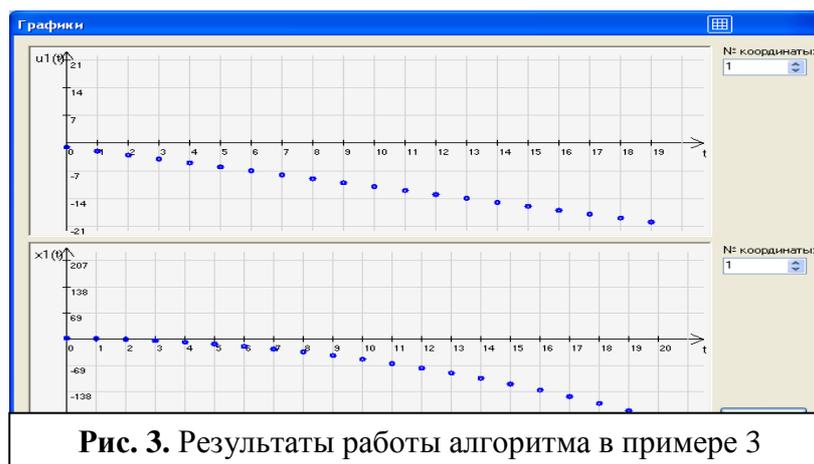


Рис. 3. Результаты работы алгоритма в примере 3

Заключение

Предложен алгоритм решения задачи нахождения оптимального программного управления дискретными детерминированными нелинейными системами с помощью генетических алгоритмов поиска глобального экстремума с вещественным кодированием, на основе которых было разработано соответствующее программное обеспечение. Эффективность алгоритмов продемонстрирована на конкретных примерах. Было выявлено, что точность решения с увеличением времени функционирования системы ухудшается. Однако широкие возможности выбора параметров алгоритма позволяют обеспечить желаемую точность.

ЛИТЕРАТУРА

1. **Herrera F., Lozano M., Verdegay J.L.** Tackling real-coded genetic algorithms: operators and tools for the behavior analysis // *Artificial Intelligence Review*, V. 12, 1998. - № 4.
2. **Wright A.** Genetic algorithms for real parameter optimization // *Foundations of Genetic Algorithms*, V. 1, 1991.
3. **Пантелеев А.В.** Метаэвристические алгоритмы поиска глобального экстремума. – М.: МАИ-ПРИНТ, 2009.
4. **Пантелеев А.В., Метлицкая Д.В.** Применение генетических алгоритмов с бинарным и вещественным кодированием к задаче поиска условного экстремума функций // *Теоретические вопросы вычислительной техники и программного обеспечения: межвуз. сб. науч. тр., МИРЭА*, 2010.
5. **Пантелеев А.В., Метлицкая Д.В.** Комплекс программных средств «Генетические алгоритмы условной оптимизации с бинарным и вещественным кодированием» // *Электронный журнал «Труды МАИ»*. - 2010. - № 37. <http://www.mai.ru/science/trudy/published.php?ID=13421>.
6. **Пантелеев А.В., Бортакровский А.С.** Теория управления в примерах и задачах. - М.: Высшая школа, 2003.

THE GENETIC ALGORITHM WITH REAL CODING IN THE OPTIMUM CONTROL PROBLEMS FOR DISCRETE SYSTEMS

Panteleyev A.V., Metlitskaya D.V.

This paper presents the decision of an optimum control problem for the discrete deterministic systems by means of genetic algorithm of a conditional global extremum search with real coding. Was generated the algorithm of the decision of a task in view on which basis the software is created. The examples illustrating efficiency of generated algorithm are submitted.

Key words: genetic algorithm, real coding, optimal control.

Сведения об авторах

Пантелеев Андрей Владимирович, 1955 г.р., окончил МГТУ им. Н.Э. Баумана (1978), доктор физико-математических наук, профессор, заведующий кафедрой математической кибернетики МАИ, автор более 130 научных работ, область научных интересов – методы синтеза оптимальных нелинейных систем управления, методы оптимизации.

Метлицкая Дарья Вадимовна, студентка МАИ, автор 3 научных работ, область научных интересов – методы оптимизации, численные методы, математическая статистика.

УДК 537.874

АНОМАЛЬНЫЕ ПРОДОЛЬНЫЕ РЕЗОНАНСЫ В ФОТОННЫХ КРИСТАЛЛАХ ТИПА “WOODPILE”

В.Л. КУЗНЕЦОВ, А.С. РУДКОВСКИЙ

При численном эксперименте на модели фотонного кристалла типа "woodpile" обнаружены аномальные продольные резонансы, связанные с наличием изолированных групп неоднородных мод. На примере частного случая плоскопараллельной пластины показан механизм усиления поля. Рассмотрены межмодовые взаимодействия и выявлены изолированные группы неоднородных мод, играющие роль аккумулятора волнового поля.

Ключевые слова: фотонный кристалл, метод погружения, резонансы.

Введение

Фотонный кристалл (ФК) представляет собой пространственно-периодическую структуру, благодаря периодичности которой распространяющиеся в нем электромагнитные волны приобретают зонный спектр. При определенных условиях в зонной структуре ФК образуются щели, подобные запрещенным зонам для электронов в естественных кристаллах. Структура энергетического спектра бесконечного ФК может быть достаточно просто вычислена, например, с помощью метода разложения по плоским волнам [1, 2]. Для модели кристалла конечной толщины наиболее адекватным аппаратом описания взаимодействия поля со средой видится метод инвариантного погружения, где в качестве параметра погружения используется переменная толщина кристалла [3-5]. При таком подходе резонансные эффекты, сопровождающие распространение волн в периодических структурах, проявляются в особенностях поведения коэффициентов отражения и прозрачности как функций таких параметров кристалла, как нормированный вектор обратной решетки и безразмерная толщина образца.

По природе возникновения такие резонансы принято делить на продольные и поперечные. Поперечный резонанс возникает в случае, когда волновой вектор одной из дифрагированных волн ложится на границу ФК. Возникающие вблизи этого состояния аномалии в поведении коэффициентов отражения и прозрачности (резонансные зависимости) принято называть аномалиями Вуда [6, 7]. Другой вид резонансов, называемых продольными, связан с возможными преотражениями волн от неоднородностей, параллельных границам кристалла конечной толщины.

В силу конечности толщины самого кристалла резонансные эффекты могут быть существенны лишь в случае, когда локальные коэффициенты отражения на неоднородностях будут близки к единице, а сами неоднородности удалены друг от друга на расстояние, соответствующее брэгговскому резонансу. Из сказанного может сложиться впечатление, что для ФК с конечными значениями диэлектрической проницаемости неоднородностей продольные резонансы не могут быть сколь-либо значимыми. В этой работе мы покажем, что это не так.

1. Модель ФК типа “Woodpile” и некоторые особенности результатов численного эксперимента

Как было показано в работах [4,5] ФК типа "woodpile", с учетом векторного характера электромагнитного поля может быть описан следующей математической моделью

$$\frac{d\left(\alpha\beta\hat{R}_{nm}^{sP+}\right)}{dz} = \alpha\beta\hat{R}_{nk}^{st+} \cdot (\alpha\beta\hat{\rho}_{kl}^{tr-}) \cdot \alpha\beta\hat{R}_{lm}^{rP+} + \alpha\beta\hat{R}_{nk}^{st+} \cdot (\alpha\beta\hat{\tau}_{km}^{tP-}) + (\alpha\beta\hat{\tau}_{nk}^{st+}) \cdot \alpha\beta\hat{R}_{km}^{tP+} + \alpha\beta\hat{\rho}_{nm}^{sP+}, \quad (1)$$

$$\frac{d\left(\alpha\beta\hat{T}_{nm}^{sp-}\right)}{dz} = \alpha\beta\hat{T}_{nk}^{st-} \cdot (\alpha\beta\hat{\rho}_{kl}^{tr-}) \cdot \alpha\beta\hat{R}_{lm}^{rp+} + \alpha\beta\hat{T}_{nk}^{st-} \cdot \alpha\beta\hat{t}_{km}^{tp-}, \quad (2)$$

$$\alpha\beta\hat{R}_{nm}^{sp+}(0) = 0, \quad \alpha\beta\hat{T}_{nm}^{sp-}(0) = \hat{I}. \quad (3)$$

Здесь $\alpha\beta\hat{R}_{nm}^{sp-} = \alpha\beta\hat{R}_{nm}^{sp-}(z)$ и $\alpha\beta\hat{T}_{nm}^{sp+} = \alpha\beta\hat{T}_{nm}^{sp+}(z)$ – матричные коэффициенты отражения и прозрачности ФК как функции переменной толщины кристалла – z , которые описывают связь между прошедшим и отраженным полями в базисе мод Флоке

$$E^+(z) = \hat{R}(z) \cdot E^-(z), \quad E^+(z) = \hat{T}(z) \cdot E^-(z).$$

Индексы α, β принимают значения h или v в зависимости от того, рассматривается вертикальная или горизонтальная поляризации иницирующего и дифрагированного полей. 6-индексные матрицы $\alpha\beta\hat{\rho}_{nm}^{sp\pm}$ и $\alpha\beta\hat{t}_{nm}^{sp\pm}$ описывают характеристики отражения и прозрачности элементарного слоя ФК и имеют следующий вид

$$\alpha\beta\hat{t}_{nm}^{sp\pm} = \hat{I} \cdot ik_z(n, s) \cdot \delta_{nm} \cdot \delta_{sp} + \alpha\beta\hat{G}_{nm}^{sp\pm} \cdot \frac{i \cdot (\varepsilon - 1)}{2\pi(n-m) \cdot k_z(n, s)} \sin\left(\frac{\pi d(n-m)}{\Lambda_x}\right), \quad (4)$$

$$\alpha\beta\hat{\rho}_{nm}^{sp} = \alpha\beta\hat{G}_{nm}^{sp\pm} \cdot \frac{i \cdot (\varepsilon - 1)}{2\pi(n-m) \cdot k_z(n, s)} \sin\left(\frac{\pi d(n-m)}{\Lambda_x}\right). \quad (5)$$

Здесь ε – диэлектрическая проницаемость вставок; Λ_x – период структуры; d – ширина диэлектрических вставок (рис. 1); $\alpha\beta\hat{G}_{nm}^{sp\pm}$ – матричная функция Грина.

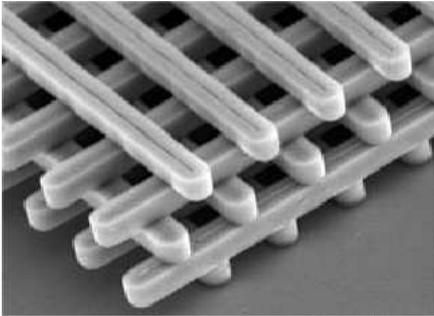


Рис. 1. Общий вид кристалла типа "woodpile", образованного диэлектрическими брусьями шириной d

Матрицы $\alpha\beta\hat{R}_{nm}^{sp-}$ и $\alpha\beta\hat{T}_{nm}^{sp+}$, как, впрочем, и другие, фигурирующие в приведенной модели, имеют бесконечную размерность, т.е. описывают распространение как однородных, так и неоднородных мод. Здесь следует пояснить, что волновые векторы \vec{k} дифрагированных волн могут отличаться от волнового вектора иницирующей волны \vec{k}_0 на произвольное целое число векторов обратной решетки кристалла – $\vec{\kappa}_1, \vec{\kappa}_2$, т.е. $\vec{k} = \vec{k}_0 + n_1 \cdot \vec{\kappa}_1 + n_2 \cdot \vec{\kappa}_2$. Если ориентировать ось Z нормально

верхней грани кристалла, то проекция \vec{k} некоторой дифрагированной моды на это направление будет равна $k_z = \sqrt{k_0^2 - (q_{0x} + n_1 \cdot \kappa_1)^2 - (q_{0y} + n_2 \cdot \kappa_2)^2}$. Моды, для которых

k_z оказывается мнимым, называются неоднородными или эванесцентными. Учет последних существенен, поскольку, благодаря межмодовым взаимодействиям, они дают весомый вклад в формирование наблюдаемых вне кристалла однородных мод с вещественными k_z . Из сказанного следует,

что при численных расчетах необходимо учитывать значения матричных коэффициентов $\alpha\beta\hat{R}_{nm}^{sp-}$ и $\alpha\beta\hat{T}_{nm}^{sp+}$, соответствующих, в частности, преобразованиям неоднородных мод в неоднородные.

Численный эксперимент показал, что для некоторых неоднородных мод при некоторых толщинах образцов ФК начинают наблюдаться аномалии – соответствующие элементы матричных коэффициентов отражения и прозрачности устремляются к бесконечности, в их поведении наблюдается сингулярность. Кажущаяся нефизичность полученного результата требует интерпретации.

Нетрудно убедиться, что приведенная математическая модель (1)-(5) допускает предельный переход $d \rightarrow \Lambda$. В этом случае рассматриваемый ФК вырождается в плоскопараллельную од-

нородную пластину, межмодовые взаимодействия исчезают, и система допускает аналитическое исследование.

2. Анализ продольных резонансов на примере однородной пластины

В случае однородной пластины задача значительно упрощается, т.к. здесь мы можем рассматривать только одну произвольную моду электромагнитного поля (межмодового взаимодействия нет). Коэффициенты ρ и τ , приведенные в уравнениях (4), (5), в случае однородной пластины могут быть представлены в следующем виде

$$\tau = i \cdot k_z + \frac{i \cdot (\varepsilon - 1)}{2 \cdot k_z}; \quad \rho = \frac{i \cdot (\varepsilon - 1)}{2 \cdot k_z},$$

для горизонтальной поляризации волны и для вертикальной поляризации

$$\tau = i \cdot k_z + \frac{i \cdot (\varepsilon - 1)}{2 \cdot k_z} \cdot \left(k_z^2 + \frac{q^2}{\varepsilon}\right); \quad \rho = -\frac{i \cdot (\varepsilon - 1)}{2 \cdot k_z} \cdot \left(k_z^2 + \frac{q^2}{\varepsilon}\right),$$

Кроме того, поскольку рассматривается только одна мода электромагнитного поля, то можно найти аналитическое решение уравнения (1), которое при $d = \Lambda$ имеет вид

$$R = \frac{\rho \cdot (\exp(2 \cdot h \cdot \sqrt{\tau^2 - \rho^2}) - 1)}{\tau + \sqrt{\tau^2 - \rho^2} - (\tau - \sqrt{\tau^2 - \rho^2}) \cdot \exp(2 \cdot h \cdot \sqrt{\tau^2 - \rho^2})} \quad \text{при } q^2 - \varepsilon > 0$$

$$R = \frac{\tau^2 \cdot h}{\rho \cdot (1 - \tau \cdot h)} \quad \text{при } q^2 - \varepsilon = 0$$

$$R = \frac{\rho \cdot \sin(h \cdot \sqrt{\tau^2 - \rho^2})}{\sqrt{\tau^2 - \rho^2} \cdot \cos(h \cdot \sqrt{\tau^2 - \rho^2}) - \tau \cdot \sin(h \cdot \sqrt{\tau^2 - \rho^2})} \quad \text{при } q^2 - \varepsilon < 0$$

Это решение полностью совпадает с известными результатами, полученными для плоской пластины другими методами [8].

В дальнейшем, чтобы проанализировать поведение коэффициента отражения, были построены трехмерные графики зависимости R от безразмерной толщины пластины $h' = h \cdot k$, где k – волновое число, в диапазоне $[0, 16]$, и величины $q' = \frac{q}{k}$ – нормированной проекции волнового вектора \vec{k} на верхнюю грань пластины с диапазоном значений $[0, 4]$. Соответствующий вид зависимостей $R = R(q', z)$ для значения диэлектрической проницаемости $\varepsilon = 9$ в случаях горизонтальной и вертикальной поляризации поля приведен на рис.2.

Из графиков видно, что вся область распадается на 3 части. Первая, соответствующая случаю падения однородной волны $0 \leq q' < 1$, дает понятные результаты: коэффициент отражения не превышает единицы и периодически меняется по мере увеличения толщины пластины. Не вызывает удивления и поведение коэффициента отражения в третьей области, где $q' > \sqrt{\varepsilon} = 3$. Здесь присутствуют только неоднородные моды как вне, так и внутри пластины, поведение коэффициента отражения монотонное.

Наибольший интерес представляет центральная область, в которой отчетливо прослеживаются резонансные эффекты. Специфика этой области заключается в том, что на пластину извне действует неоднородное поле, а внутри пластины это поле становится обычным однородным

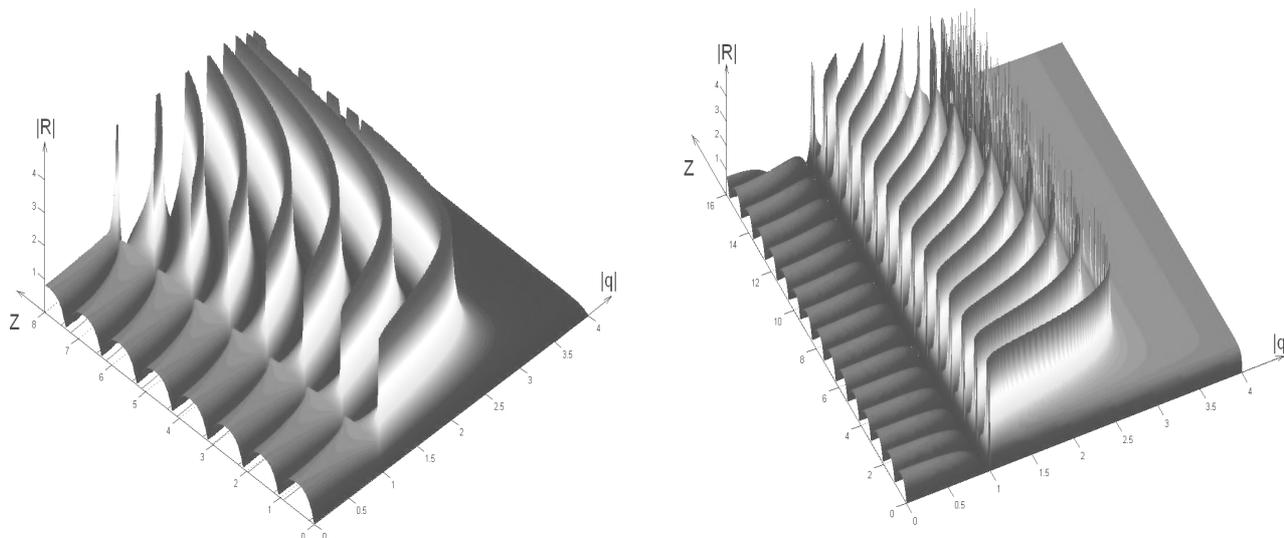


Рис. 2. Зависимость коэффициента отражения R от толщины пластины и величины проекции падающего поля для случаев горизонтальной и вертикальной поляризации поля

Это поле не может покинуть пластины, испытывая на ее границах полное внутреннее отражение. Вне пластины поле становится неоднородным и энергию выводить из пластины не может. При определенной толщине пластины z наступает резонанс, и амплитуда поля стремится к бесконечности. Периодичность таких резонансов определяется по известной формуле [8]

$$z = \frac{\pi \cdot n - \arg(R(q))}{\sqrt{\epsilon - q^2}},$$

где $R(q)$ – коэффициент отражения, определяемый формулами Френеля.

С физической точки зрения возникновение резонансов – неограниченного роста поля внутри пластины – связано с тем, что инициирующая неоднородная мода, непрерывно воздействуя на среду пластины, закачивает в нее энергию. Поскольку выходящие неоднородные моды в свободное пространство энергию не переносят, то она аккумулируется в этой локализованной области (рис. 3).

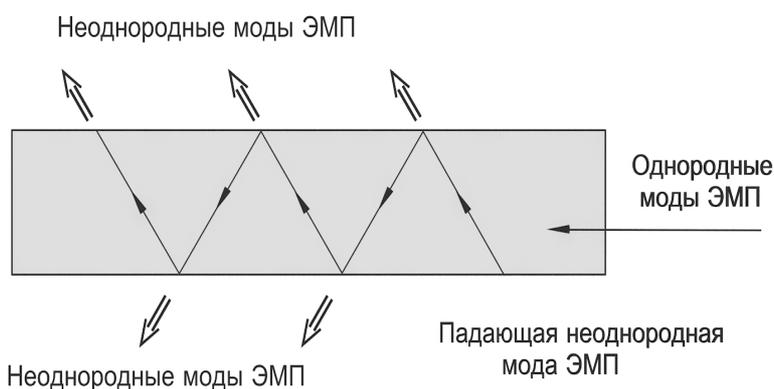


Рис. 3. Преобразование инициирующей неоднородной моды в плоскопараллельной пластине

Рассматриваемая нами модель взаимодействия излучения с пластиной предполагает стационарный режим и поэтому выдает сингулярные результаты, предсказывающие неограниченный рост поля при анализе переходных режимов.

Здесь уместно провести аналогию с резонансом вынужденных колебаний гармонического осциллятора. В стационарном режиме амплитуда вынужденных колебаний определяется по формуле $x(t) = \frac{\Phi_0}{\omega_0^2 - \Omega^2} \cos(\Omega t)$, описывающей резонансную зависимость, в то время как анализ динамики развития колебаний дает просто неограниченный линейный рост амплитуды со временем – $x(t) = \frac{\Phi_0}{2\Omega} \cdot t \cdot \sin(\Omega t)$.

3. Интерпретация аномальных резонансов в модели ФК типа "woodpile"

Приведенная выше интерпретация аномалий основана на том, что отвод энергии электромагнитного поля от пластины невозможен, поскольку единственной модой поля, излучаемой системой, является неоднородная мода, не способная переносить энергию. В случае ФК излучение содержит весь возможный спектр как однородных, так и неоднородных мод, и описанный механизм аномального резонанса для ФК необходимо обсудить отдельно.

ФК типа "woodpile" состоит из отдельных слоев, каждый из которых характеризуется лишь одним вектором обратной решетки, перпендикулярным к направляющей брусьев слоя.

Напомним, что уравнения модели (1)-(5) интегрируются, начиная с нижней границы. При этом, при расчете, например, матричного коэффициента отражения ${}_{\alpha\beta}\hat{R}_{nm}^{sP-}$ необходимо учитывать все моды, которые могут порождаться при дифракции поля в более высоких слоях ФК. Другими словами, если на весь кристалл сверху падает одна плоская волна с волновым вектором $\vec{k}_0 = (\vec{q}, k_z)$, то при расчете первого, нижнего слоя, мы должны предполагать падение всего углового спектра поля, характеризующегося векторами $\vec{k} = \vec{k}_0 + n_1 \cdot \vec{k}_1 + n_2 \cdot \vec{k}_2$, где $n_1, n_2 \in \mathbf{Z}$. Покажем, что при взаимодействии с нижним слоем возникает группа неоднородных мод, изолированных от однородных, т.е. не взаимодействующих с ними. Энергия этой группы мод не может уменьшаться за счет излучения поля (рис. 4).

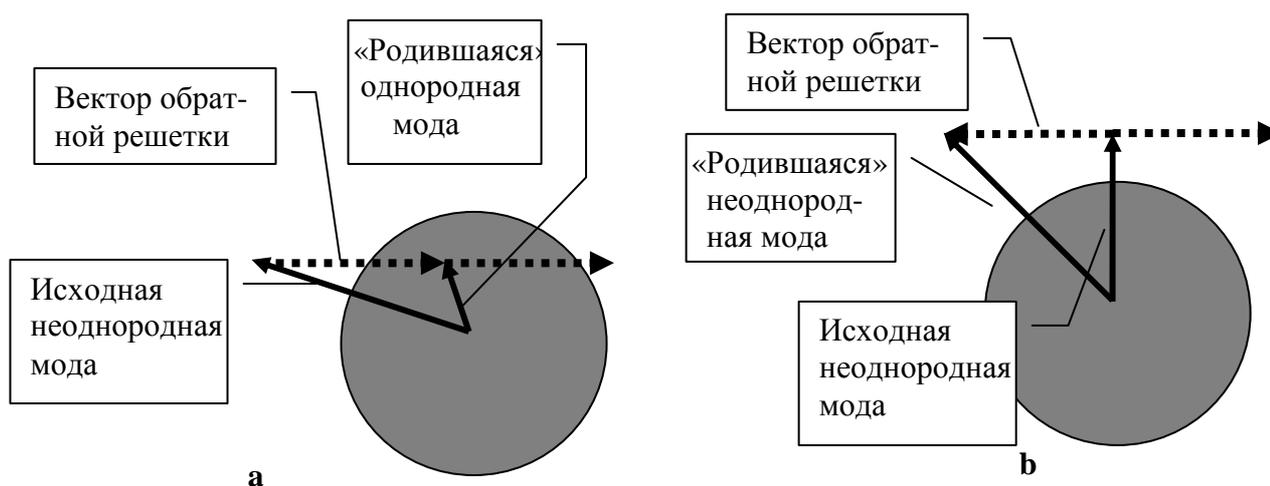


Рис. 4. Схемы межмодовых взаимодействий в нижнем слое ФК типа "woodpile".

Затемненный круг – область однородных мод: а - порождение неоднородной модой однородной – канал отвода энергии; б - замкнутость внутри класса неоднородных мод

На рис. 4 представлены некоторые схемы, иллюстрирующие межмодовые взаимодействия в первом слое. На рис. 4а проиллюстрирован «путь», по которому энергия из некоторой исходной неоднородной моды может перекачиваться в однородные моды и покидать нижний слой. На рис. 4б показана группа изолированных неоднородных мод, играющих роль той одной, что фигурировала в случае плоскопараллельной пластины.

Заключение

Исследования электродинамических характеристик ФК типа "woodpile" показали наличие групп параметров, при которых внутри кристалла может накапливаться энергия электромагнитного поля. Внешним проявлением этого эффекта является неограниченный рост напряженности поля в приповерхностном слое. Близкие, но не эквивалентные по своей сути эффекты, из-

вестны уже достаточно давно [7]. Это существенное усиление поля вблизи периодически возмущенной границе раздела двух сред, наблюдаемое при ее облучении электромагнитным полем. Здесь, в отличие от случая, рассмотренного в статье, аномалии связаны с проявлением поперечного резонанса и известны как аномалии Вуда.

Выявление аномальных продольных резонансов тесно связано с использованием метода инвариантного погружения, согласно идеологии которого, при решении задачи последовательно вычисляются значения матричных коэффициентов ${}_{\alpha\beta}\hat{R}_{nm}^{sp-}$ и ${}_{\alpha\beta}\hat{T}_{nm}^{sp+}$ при всех значениях толщины кристалла z .

ЛИТЕРАТУРА

1. Лозовик Ю.А., Эйдерман С.Л. Зонная структура сверхпроводящих фотонных кристаллов // Физика твердого тела. - 2008. - Т. 50. - Вып. 11.
2. Lifeng Li. Use of Fourier series in the analysis of discontinuous periodic structure.//J. Opt. Soc. Am. A. v.13, No.9, 1996.
3. Барабаненков Ю.Н., Барабаненков М.Ю. Метод соотношений переноса в теории резонансного многократного рассеяния волн с применением к дифракционным решеткам и фотонным кристаллам // ЖЭТФ. - 2003. - Т. 123. - Вып. 4.
4. Kuznetsov V.L., Rudkovskiy A.S. Invariant embedding method in the problem of 3D photonic crystal modeling PIERS Proceedings, 2009.
5. Кузнецов В.Л., Рудковский А.С. Трехмерная модель взаимодействия электромагнитного с фотонным кристаллом конечной толщины // Научный Вестник МГТУ ГА. Информатика. - 2009. - № 145.
6. Wood R.W. Anomalous diffraction grating //Phys. Rev., vol. 48, 928-936, 1935.
7. Ахманов С.А., Семиногов В.Н., Соколов В.Н. Дифракция света на шероховатой поверхности с «глубоким» профилем: взаимодействие дифрагированных волн, аномальное поглощение, максимально возможные локальные поля // ЖЭТФ. - 1987. - Т. 93.
8. Борн М., Вольф Э. Основы оптики. - М.: Наука, 1973.

ABNORMAL LONGITUDINAL RESONANCES IN THE PHOTON CRYSTALS WOODPILE TYPE

Kuznetsov V.L., Rudkovskiy A.S.

Abnormal longitudinal resonances for the model of a photon crystal woodpile type, by numerical experiment, are found out. This effect is connected with the presence of isolated groups of evanescent modes. The mechanism of strengthening of the field in the case of plane-parallel plate is shown. Cross-modes interactions are analyzed and the isolated groups of evanescent modes as an accumulator of wave field are noted.

Key words: photon crystal, imbedding method, resonances.

Сведения об авторах

Кузнецов Валерий Леонидович, 1949 г.р., окончил МГУ им. М.В.Ломоносова (1972), доктор технических наук, профессор, заведующий кафедрой прикладной математики МГТУ ГА, автор более 100 научных работ, область научных интересов – методы математического моделирования в задачах распространения излучения в пространственно неоднородных случайных и периодических средах, безопасность полетов.

Рудковский Антон Сергеевич, 1984 г.р., окончил МГТУ ГА (2006), аспирант МГТУ ГА, автор 6 научных работ, область научных интересов – математическое моделирование взаимодействия поля с периодическими структурами, фотонные кристаллы.

УДК 330.42

ОЦЕНКА ИНДЕКСА ТРЕНДОВОГО ПРОГНОЗИРОВАНИЯ НА ОСНОВЕ ИТЕРАЦИОННОЙ РЕГУЛЯРИЗАЦИИ

А.А. КОСТОГЛОТОВ, Д.С. АНДРАШИТОВ, А.А. ОРУДЖЕВА, И.В. ДЕРЯБКИН

Проблема прогнозирования динамики российского рынка акций активно обсуждается в научной литературе, поскольку существующие методы не достаточно эффективны. Основой решения задачи прогнозирования ценовых изменений является задача текущего оценивания. В работе предлагается новый алгоритм последовательного во времени оценивания вектора опережающих показателей на основе итерационной регуляризации.

Ключевые слова: индекс трендового прогнозирования, итерационная регуляризация.

Введение

Прогнозирование рыночных тенденций является актуальной научно-практической задачей. Ее решение позволит сформировать рациональную для инвесторов стратегию управления инвестиционным процессом в условиях изменяющихся во времени случайных факторов на основе определения моментов разворота тренда. Традиционно оценка момента разворота связана с использованием технических индикаторов, построенных на временных рядах цены, объема или различных индексов фондовых и товарных рынков.

В последнее время кроме классических индексов (RTSI, DJIA, S&P 500) для прогнозирования все чаще используют так называемые индексы трендового прогнозирования – сводный и диффузный [1], которые строятся на основе временных рядов вектора опережающих показателей, среди которых индекс, подлежащий анализу, цена нефти, значение пары EUR/USD и т.д. Существует множество способов построения сводного и диффузного индекса, при этом они характеризуются разным уровнем прогностической способности, что следует из материалов исследований различных авторов, которые дают субъективные оценки [1, 2]. Однако объединяет все эти способы то, что в любом случае при таких построениях производится сглаживание рядов данных с помощью методов построения скользящих средних, метода наименьших квадратов, динамической линейной и нелинейной фильтрации и другие.

Практической конструктивностью при достаточно высокой точности обладает алгоритм фильтрации Калмана, который дает оптимальное решение только для линейных моделей динамики, и винеровском формирующем шуме, причем параметры модели и шума должны быть известны точно. В реальных задачах обработки данных такое положение вещей является скорее экзотикой и приводит на практике к расходимости оценки фильтра Калмана с истинными значениями оцениваемых параметров, что ограничивает использование данного алгоритма при подготовке данных для построения прогнозных индексов. Исключить недостатки алгоритма Калмана позволяют итеративные регуляризованные алгоритмы [3-5], обеспечивающие получение устойчивых оценок, превосходящих по точности фильтр Калмана, за счет их последовательного уточнения с использованием метода итераций.

Настоящая статья посвящена разработке алгоритма последовательного во времени оценивания вектора влияющих показателей для индекса трендового прогнозирования (ИТП) на основе итерационной регуляризации.

1. Постановка задачи

Пусть динамика вектора влияющих показателей для ИТП описывается системой разностных уравнений в дискретном времени

$$\mathbf{x}(k+1) = \mathbf{F}(\mathbf{x}(k), k) + \boldsymbol{\eta}(k), \quad k = \overline{0, N-1}, \quad \mathbf{x}(0) = \hat{\mathbf{x}}(0), \quad (1)$$

где $\mathbf{x}(k) = [x^1(k), x^2(k), \dots, x^M(k)]^T \in E^M$ – вектор состояния в k -й момент времени $\mathbf{x}(t_k)$, E^M – M -мерное вещественное евклидово пространство, с нормой $\|\mathbf{x}\|_{E^M} = \sqrt{\sum_{l=1}^M (x^l)^2}$;

$\boldsymbol{\eta}(k) = [\eta^1(k), \eta^2(k), \dots, \eta^M(k)]^T \in E^M$ – вектор неслучайных неизвестных внешних воздействий с ограниченной нормой в евклидовом пространстве, характеризующий различие принятой модели и реальной динамики экономического показателя; k, N, M – натуральные числа;

$\mathbf{F}(\mathbf{x}(k), k) = [F^1(\mathbf{x}(k), k), F^2(\mathbf{x}(k), k), \dots, F^M(\mathbf{x}(k), k)]^T \in E^M$ – непрерывная вместе с частными производными вектор-функция своих аргументов.

Модель наблюдений описывается векторным уравнением

$$\mathbf{y}(k) = \mathbf{x}(k) + \mathbf{n}(k), \quad (2)$$

где $\mathbf{y}(k) = [y^1(k), y^2(k), \dots, y^M(k)]^T \in E^M$ – вектор наблюдения; $\mathbf{n}(k) = [n^1(k), n^2(k), \dots, n^M(k)]^T \in E^M$ – вектор дискретного белого гауссовского шума с известными локальными характеристиками

$$M[\mathbf{n}(k)] = 0, \quad M[\mathbf{n}(k)\mathbf{n}^T(l)] = \mathbf{W}\delta(k-l),$$

где \mathbf{W} – ковариационная матрица; $\delta(\cdot)$ – векторная дельта-функция.

Величина $\mathbf{n}(k)$ описывает случайные спекулятивные действия участников торгов.

Определим оценку $\hat{\mathbf{x}}(k+1)$, обеспечивающую минимум функционала невязки

$$J = \frac{1}{2} \sum_{k=0}^{N-1} [\mathbf{y}(k+1) - \hat{\mathbf{x}}(k+1)]^T \mathbf{W}^{-1} [\mathbf{y}(k+1) - \hat{\mathbf{x}}(k+1)]. \quad (3)$$

2. Итерационная регуляризация

Решение уравнения (1) непрерывно зависит от внешних воздействий $\boldsymbol{\eta}(k)$, поэтому функционал ошибки (3) непрерывно зависит от $\boldsymbol{\eta}(k)$. Таким образом, задача определения оценки $\hat{\mathbf{x}}(k+1)$, доставляющей минимум (3), равносильна задаче определения

$$\boldsymbol{\eta}^* = \arg \min_{\boldsymbol{\eta}} J[\hat{\mathbf{x}}, \boldsymbol{\eta}]. \quad (4)$$

Задача (1),(2),(4) является некорректно поставленной по Адамару [6].

Найти значения векторов $\hat{\mathbf{x}}(k+1)$, $\boldsymbol{\eta}^*(k)$ для $k = \overline{1, N-1}$ путём решения совокупности уравнений (1),(2),(4) в условиях некорректности исходной задачи достаточно сложно, в связи с этим широкое распространение получили итерационные градиентные методы. Однако использование таких методов может привести к расходящейся последовательности приближений. Поэтому применение любого итерационного метода для решения задачи (1),(2),(4) требует определения регуляризирующего семейства операторов, в котором параметром регуляризации является номер итерации.

В соответствии с общим определением регуляризирующего семейства операторов по А.Н. Тихонову [6] будем говорить, что итерационный метод

$$\boldsymbol{\eta}_{n+1} = \boldsymbol{\eta}_n - \alpha_n \cdot \text{grad} J[\mathbf{x}_n, \boldsymbol{\eta}_n], \quad n = 0, 1, 2, \dots, \quad (5)$$

в котором числовой параметр регуляризации α_n удовлетворяет условиям

$$\lim_{n \rightarrow \infty} \alpha_n = 0, \quad \lim_{n \rightarrow \infty} \sum_{k=1}^n \alpha_k = \infty, \quad \lim_{n \rightarrow \infty} \sum_{k=1}^n \alpha_k^2 < \infty \quad (6)$$

и порождает регуляризирующее семейство операторов, в котором параметром является номер итерации, если для любого начального приближения $\boldsymbol{\eta}_0$ и для любого значения погрешности

исходных данных σ , удовлетворяющего условию $0 < |\sigma| < \sigma_0$, $\sigma_0 = \text{const}$, существует номер $n^*(\sigma)$ такой, что $\lim_{\sigma \rightarrow 0} \eta_{n^*(\sigma)} = \eta^*$, т.е. полученные приближения сходятся к точному решению в норме пространства E^M при стремлении погрешности исходных данных к нулю.

Один из возможных подходов к выбору параметра регуляризации $n^*(\sigma)$ при использовании алгоритма простой итерации заключается в использовании критерия невязки [6], т.е. в выборе такого $n = n^*(\sigma)$, при котором $\Delta_n \rightarrow 0$

$$\Delta_n = \sum_{k=0}^{N-1} [\mathbf{y}(k) - \hat{\mathbf{x}}(k)]^T [\mathbf{y}(k) - \hat{\mathbf{x}}(k)].$$

В качестве приближенного решения будем выбирать элемент итерационной последовательности с номером n^* , удовлетворяющим условию [7]

$$n^* = \min\{n : \Delta_n \rightarrow 0\}.$$

Для реализации итерационного метода (5) требуется определить градиент функционала (3). Согласно [8], выражение для определения градиента в точке $\eta_n(k)$ имеет вид

$$\text{grad}J[\mathbf{x}_n(k), \eta_n(k)] = \Psi_n(k), \quad k = \overline{0, N-1}, \quad (7)$$

где $\mathbf{x}_n(k)$ – решение задачи (1) при $\eta_n(k)$, а вектор $\Psi_n(k)$ определяется из условий

$$\begin{aligned} \Psi_n(k) &= \frac{\partial \mathbf{F}^T}{\partial \mathbf{x}} \Psi_n(k+1) + \mathbf{W}^{-1}[\mathbf{y}(k+1) - \mathbf{x}_n(k+1)], \quad k = \overline{0, N-1}, \\ \Psi_n(N) &= 0. \end{aligned} \quad (8)$$

3. Рекуррентная итеративная оценка вектора влияющих показателей

Приняв за нулевое приближение $\eta_0(k) = 0$, $k = \overline{0, N-1}$, запишем итерационную последовательность в развернутой форме для $k = \overline{0, N-1}$

$$\begin{aligned} \eta_0(k) &= 0, \\ \eta_1(k) &= \eta_0(k) - \alpha_0 \Psi_0(k) = -\alpha_0 \Psi_0(k), \\ \eta_2(k) &= \eta_1(k) - \alpha_1 \Psi_1(k) = -\alpha_0 \Psi_0(k) - \alpha_1 \Psi_1(k), \\ &\dots \\ \eta_{n+1}(k) &= -\sum_{i=0}^n \alpha_i \Psi_i(k). \end{aligned}$$

В результате, с учетом (8), имеем последовательность дискретных двухточечных краевых задач (ДДТКЗ) вида

$$\begin{aligned} \mathbf{x}_n(k+1) &= \mathbf{F}(\mathbf{x}_n(k), k) - \sum_{i=0}^n \alpha_i \Psi_i(k), \\ \Psi_i(k) &= \frac{\partial \mathbf{F}^T}{\partial \mathbf{x}} \Psi_i(k+1) + \mathbf{W}^{-1}[\mathbf{y}(k+1) - \mathbf{x}_i(k+1)], \\ \Psi_i(N) &= 0, \mathbf{x}_i(0) = \hat{\mathbf{x}}(0), \quad i = \overline{0, n}, \quad k = \overline{0, N-1}. \end{aligned} \quad (9)$$

Используя процедуру инвариантного погружения, получим последовательность уравнений для оценки $\hat{\mathbf{x}}(k+1)$ в виде [4]

$$\begin{aligned}
\hat{\mathbf{x}}_n(k+1) &= \mathbf{F}(k) + \boldsymbol{\eta}(k) = \mathbf{F}(\hat{\mathbf{x}}_n(k), k) + \mathbf{W}^{-1} \mathbf{P}_n(k+1) \sum_{i=0}^n \alpha_i [\mathbf{y}(k+1) - \hat{\mathbf{x}}_i(k+1/k)], \\
\hat{\mathbf{x}}_i(k+1/k) &= \mathbf{F}(\mathbf{x}_i(k), k) \\
\mathbf{P}_n(k+1) &= \left\{ \mathbf{I} + \mathbf{W}^{-1} \left(\sum_{i=0}^n \alpha_i \right) \mathbf{P}_n(k+1/k) \right\}^{-1} \mathbf{P}_n(k+1/k), \\
\mathbf{P}_n(k+1/k) &= \mathbf{I} + \frac{\partial \mathbf{F}}{\partial \hat{\mathbf{x}}} \mathbf{P}_n(k) \frac{\partial \mathbf{F}^T}{\partial \hat{\mathbf{x}}} \\
\hat{\mathbf{x}}_i(0) &= \hat{\mathbf{x}}(0), \\
\mathbf{P}_i(0) &= \mathbf{P}(0), \\
i &= \overline{0, n}, \quad k = \overline{0, N-1}.
\end{aligned} \tag{10}$$

где \mathbf{I} – единичная матрица; $\mathbf{P}_n(k+1/k)$ – ковариационная матрица влияющих показателей размерности $M \times M$.

Таким образом, уравнения (10) определяют оценку $\hat{\mathbf{x}}(k+1)$ с учетом наблюдения $\mathbf{y}(k+1)$

$$\hat{\mathbf{x}}(k+1) = F(\hat{\mathbf{x}}(k), k) + \boldsymbol{\eta}(\mathbf{P}(k), \mathbf{y}(k+1), \hat{\mathbf{x}}, \alpha).$$

Совокупность уравнений (10) представляет собой последовательную во времени оценку вектора влияющих показателей ИТП. Как показывают результаты работ [3, 4], применение (10) позволяет повысить в сравнении с фильтром Калмана точность получаемых оценок за счет процесса последовательного уточнения на основе итераций по индексу n , что в свою очередь, позволит повысить точность последующего прогноза.

4. Пример

Оценку качества разработанного алгоритма проведем на основе численного моделирования эволюции вектора опережающих показателей для индекса трендового прогнозирования [1], в составе которого:

- индекс РТС;
- индекс S&P 500 – центральный показатель средней стоимости рынка США;
- пара EUR/USD – отношение официальных валют стран “еврозоны” и США;
- цена нефти марки Brent.

Примем модель состояния в виде [9]

$$\mathbf{x}(k+1) = \mathbf{x}(k)(1+S) + \boldsymbol{\eta}(k), \quad S = \text{const}, \quad k = \overline{0, N-1}, \tag{11}$$

где вектор $\mathbf{x}(k) = (x^1(k), x^2(k), x^3(k), x^4(k))^T$; $x^1(k)$ – индекс РТС; $x^2(k)$ – индекс S&P; $x^3(k)$ – значение пары EUR/USD; $x^4(k)$ – цена нефти марки Brent; S – величина, характеризующая динамику выбранного показателя с учетом масштаба.

Уравнение наблюдения в векторной форме имеет вид

$$\mathbf{y}(k) = \mathbf{x}(k) + \mathbf{n}(k), \tag{12}$$

где $\mathbf{n}(k) = (n^1(k), n^2(k), n^3(k), n^4(k))^T$, $\mathbf{y}(k) = (y^1(k), y^2(k), y^3(k), y^4(k))^T$, $M[\mathbf{n}(k)] = 0$,

$$M[\mathbf{n}(k)\mathbf{n}^T(l)] = \mathbf{W}\delta(k-l).$$

Уравнения алгоритма оценки (10) для модели (11) при $n = 1$ и параметрах регуляризации $\alpha_0 = 1$; $\alpha_1 = 0,5$ примут вид

$$\begin{aligned}
\hat{\mathbf{x}}_0(k+1) &= \mathbf{F}(\hat{\mathbf{x}}_0(k), k) + \mathbf{W}^{-1} \mathbf{P}_0(k+1) \alpha_0 [\mathbf{y}(k+1) - \hat{\mathbf{x}}_0(k+1/k)], \\
\hat{\mathbf{x}}_0(k+1/k) &= \mathbf{F}(\mathbf{x}_0(k), k), \\
\mathbf{P}_0(k+1) &= \left\{ \mathbf{I} + \alpha_0 \mathbf{W}^{-1} \mathbf{P}_0(k+1/k) \right\}^{-1} \mathbf{P}_0(k+1/k), \\
\mathbf{P}_0(k+1/k) &= \mathbf{I} + \frac{\partial \mathbf{F}}{\partial \hat{\mathbf{x}}} \mathbf{P}_0(k) \frac{\partial \mathbf{F}^T}{\partial \hat{\mathbf{x}}}, \\
\hat{\mathbf{x}}_0(0) &= \hat{\mathbf{x}}(0), \\
\mathbf{P}_0(0) &= \mathbf{P}(0),
\end{aligned} \tag{13}$$

$$\begin{aligned}
\hat{\mathbf{x}}_1(k+1) &= \mathbf{F}(\hat{\mathbf{x}}_1(k), k) + \mathbf{W}^{-1} \mathbf{P}_1(k+1) \{ \alpha_0 [\mathbf{y}(k+1) - \hat{\mathbf{x}}_0(k+1/k)] + \\
&\quad + \alpha_1 [\mathbf{y}(k+1) - \hat{\mathbf{x}}_1(k+1/k)] \}, \\
\hat{\mathbf{x}}_1(k+1/k) &= \mathbf{F}(\mathbf{x}_1(k), k), \\
\mathbf{P}_1(k+1) &= \left\{ \mathbf{I} + \mathbf{W}^{-1} (\alpha_0 + \alpha_1) \mathbf{P}_1(k+1/k) \right\}^{-1} \mathbf{P}_1(k+1/k), \\
\mathbf{P}_1(k+1/k) &= \mathbf{I} + \frac{\partial \mathbf{F}}{\partial \hat{\mathbf{x}}} \mathbf{P}_1(k) \frac{\partial \mathbf{F}^T}{\partial \hat{\mathbf{x}}}, \\
\hat{\mathbf{x}}_1(0) &= \hat{\mathbf{x}}(0), \\
\mathbf{P}_1(0) &= \mathbf{P}(0), \\
i &= \overline{0,1}, \quad k = \overline{0, N-1},
\end{aligned} \tag{14}$$

где $\mathbf{P}(k)$ – взаимная корреляционная матрица компонентов вектора состояния.

В качестве начальных условий для моделирования назначим следующие значения

$$\mathbf{x}(0) = \begin{bmatrix} 1357 \\ 1043 \\ 1,476 \\ 76,63 \end{bmatrix},$$

$$\mathbf{P}(0) = \begin{bmatrix} 1 & 0,532 & 0,77 & 0,74 \\ 0,532 & 1 & 0,709 & 0,565 \\ 0,77 & 0,709 & 1 & 0,553 \\ 0,74 & 0,565 & 0,553 & 1 \end{bmatrix},$$

$$\mathbf{W}^{-1} = \begin{bmatrix} 96 & 0 & 0 & 0 \\ 0 & 35 & 0 & 0 \\ 0 & 0 & 0,07 & 0 \\ 0 & 0 & 0 & 2,5 \end{bmatrix}.$$

Таким образом, уравнения (13) представляют собой уравнения дискретного фильтра Калмана, а уравнения (14) – его уточнение.

Данные о влияющих показателях использованы за период с 01.11.09 по 27.01.10 года с интервалом в один день. Результаты моделирования представлены на рис. 1, где показано, что сигнал традиционного индикатора, построенного на основе оценки индекса РТС по предлагаемому методу, как правило, происходит с упреждением относительно сигнала индикатора, построенного по значениям индекса РТС, что весьма важно при управлении инвестициями.

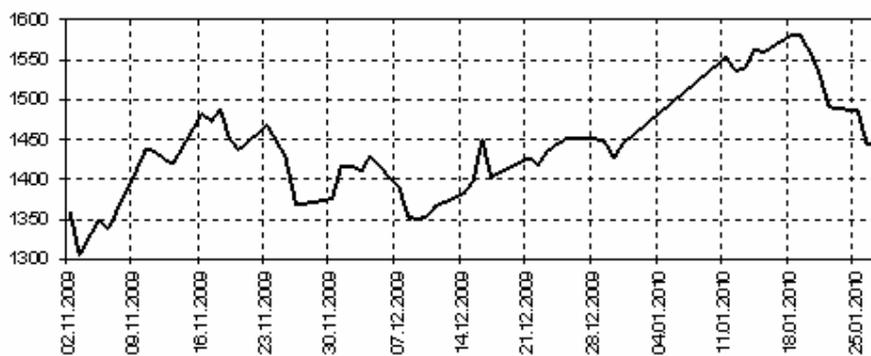


Рис. 1а. – PTC

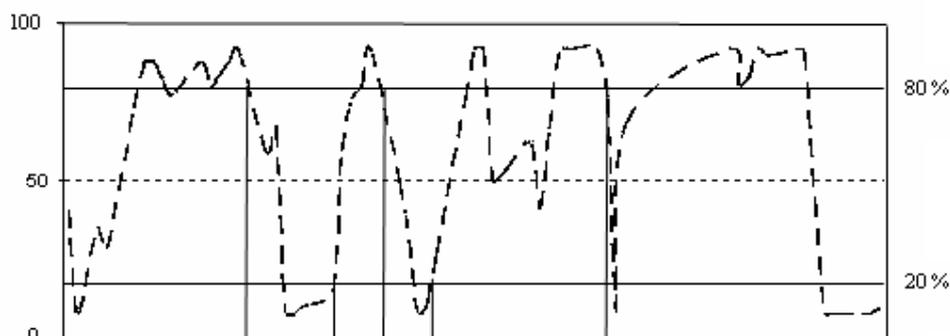


Рис. 1б. – Stochastic Oscillator PTC

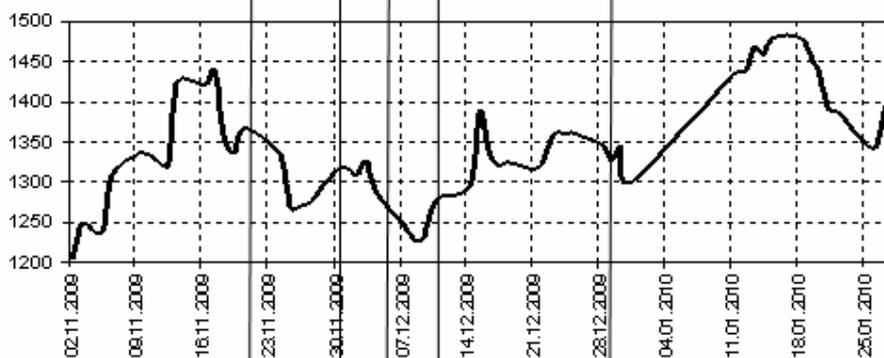


Рис. 1в. – Оценка \hat{x}^1

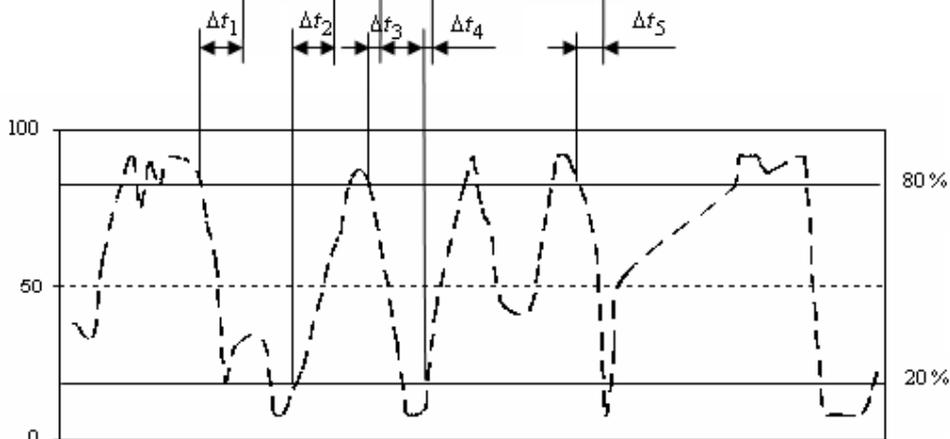


Рис. 1г. – Stochastic Oscillator оценки \hat{x}^1

Заключение

В статье разработан новый алгоритм последовательного во времени оценивания индекса трендового прогнозирования на основе итерационной регуляризации. Показано, что предлагаемый алгоритм позволяет получить сигналы о смене тренда раньше, чем использование текущих значений экономического показателя, что и определяет эффективность его использования.

ЛИТЕРАТУРА

1. **Рожков А.Г.** Построение прогнозного индекса для фондового рынка // Вопросы экономики. - 2005. - № 12.
2. **Kennedy J.** The Information in Diffusion Indexes for Forecasting Related Economic Aggregates.- Economics Letters. - 1994. vol. 44.
3. **Костоглотов А.А.** Синтез интеллектуальных измерительных процедур на основе принципа регуляризации А.Н.Тихонова // Измерительная техника. - 2001. - № 1.
4. **Костоглотов А.А.** Цифровая интеллектуальная измерительная процедура // Измерительная техника. - 2001. - № 1.
5. **Костоглотов А.А., Кузнецов А.А.** Итеративная измерительная процедура оценки параметров групповой меры // Научная мысль Кавказа. Приложение. - 2003. - №7 (48).
6. **Тихонов А.Н., Арсенин В.Я.** Методы решения некорректных задач. - М.: Наука, 1981.
7. **Алифанов О.М., Артюхин Е.А., Румянцев С.В.** Экстремальные методы решения некорректных задач. - М.: Наука, 1988.
8. **Васильев Ф.П.** Численные методы решения экстремальных задач. - М.: Наука, 1988.
9. **Лукашин Ю.П.** Адаптивные методы краткосрочного прогнозирования временных рядов. - М.: Финансы и статистика, 2003.

EVALUATION INDEX TREND FORECASTING BASED ON ITERATIVE REGULARIZATION

Kostoglotov A.A., Andrahitov D.S., Orudjeva A.A., Derabkin I.V.

The problem of quality of forecasting the stock market is actively discussed in scientific literature, as existing methods are currently not efficient enough. An algorithm for sequential estimation of the index of the trend forecasting based on the iterative regularization is given.

Key words: index of trend forecasting, iterative regularization.

Сведения об авторах

Костоглотов Андрей Александрович, 1964 г.р., окончил Ростовское высшее командно-инженерное училище ракетных войск (1986), доктор технических наук, доцент кафедры социально-экономических дисциплин Ростовского филиала МГТУ ГА, автор более 120 научных работ, область научных интересов – логистика, управленческие решения, фондовые рынки, российский рынок акций.

Андрашитов Дмитрий Сергеевич, 1985 г.р., окончил Ростовский военный институт ракетных войск им. Главного маршала артиллерии М.И. Неделина РВКИУ РВ (1988), адъюнкт кафедры 35 Ростовского военного института ракетных войск, автор 11 научных работ, область научных интересов – фондовые рынки.

Оруджева Анна Александровна, окончила Северо-Кавказскую академию государственной службы (2004), аспирантка ПИ ЮФУ, автор 5 научных работ, область научных интересов – инвестиции, человеческий капитал.

Дерябкин Игорь Владимирович, 1984 г.р., окончил ДГТУ (2006), аспирант Ростовской академии сервиса, автор 4 научных работ, область научных интересов – оценка, прогнозирование временных рядов.

УДК 658.15

ПРИМЕНЕНИЕ НЕЧЕТКОГО ВЫВОДА ДЛЯ ОПРЕДЕЛЕНИЯ ТИПА КРИЗИСА В АДАПТИВНОЙ СИСТЕМЕ СТРАТЕГИЧЕСКОГО УПРАВЛЕНИЯ РИСКОМ

А.Г. БАДАЛОВА, П.А. ПАНТЕЛЕЕВ

Рассмотрена задача определения типа кризиса на предприятии авиационно-промышленного комплекса, возникающая в случае отсутствия или неэффективного функционирования системы управления рисками, следствием чего, в свою очередь, является перерастание рисков в кризисы предприятия. Для решения указанной задачи авторами предложен алгоритм формирования нечеткой системы, служащей для нахождения типа кризиса по соответствующей шкале на основании лингвистических значений процента охваченных кризисом стратегических бизнес-единиц и общего числа реализовавшихся рисков.

Ключевые слова: риск, кризис, тип кризиса, нечеткие множества, нечеткая система, нечеткий вывод, лингвистические переменные, фаззификация, дефаззификация.

Введение

В [1,2] описан процесс создания полноценной теоретико-методологической базы управления рисками. Под стратегическим управлением рисками авторами понимается процесс управления рисками промышленного предприятия, ориентированного на реализацию выбранной стратегии развития и достижения целевых стратегических показателей предприятия. В [3] авторами предложена функциональная схема системы управления риском (СУР) предприятия авиационно-промышленного комплекса (АПК), разрабатывающего новые образцы авиационной техники, представляющая собой систему с четырьмя обратными связями, соответствующими четырем горизонтам управления. В процессе функционирования потенциальные риски предприятия АПК при отсутствии или неэффективном функционировании системы управления ими с течением времени постепенно перерастают в кризисы предприятия. Если в процессе функционирования СУР возникает ситуация, в которой, несмотря на применение различных методов управленческого воздействия, часть потенциальных рисков реализовалась и некоторое количество стратегических бизнес-единиц и/или подразделений предприятия охвачено кризисом, система стратегического управления рисками нуждается в дополнении новыми обратными связями, отражающими эффект адаптации к возникающим ситуациям. При этом актуальными задачами являются нахождение типа кризиса (локального, комплексного, глобального) и выработка соответствующих управленческих решений [2].

1. Постановка задачи нахождения типа кризиса

В [1, 2] предложена гипотеза перерастания реализовавшихся рисков в кризисы, причинами которого являются: резкое повышение уровня риска, увеличение числа реализовавшихся рисков, появление новых рискообразующих факторов. Как следствие, может проявиться неэффективность СУР, выражающаяся в неудовлетворении критериев качества функционирования всех или нескольких контуров управления. Правильное функционирование СУР безусловно направлено на предотвращение кризисов, однако при наихудшем для предприятия сочетании рискообразующих факторов возможно перерастание рисков в локальный кризис, дальнейший переход из локального кризиса в комплексный, а затем в глобальный [1, 2]. *Локальный* кризис, как правило, охватывает одну функциональную бизнес-единицу, *комплексный* кризис – более одной. *Глобальный* кризис охватывает все предприятие в целом (рис. 1).

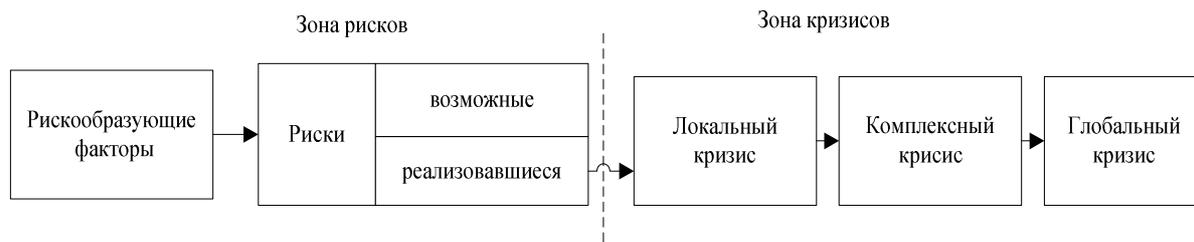


Рис. 1. Схема перерастания рисков в кризисы предприятия

В случае возникновения кризисов СУР [3] должна быть дополнена новыми контурами обратной связи (рис. 2). При возникновении локального (и отчасти комплексного) кризиса требуется адаптация управляющих воздействий в СУР на основе стоимостных показателей. При развитии комплексного кризиса применяются методы антикризисного управления [7], а при глобальном кризисе – методы управления банкротством [8].

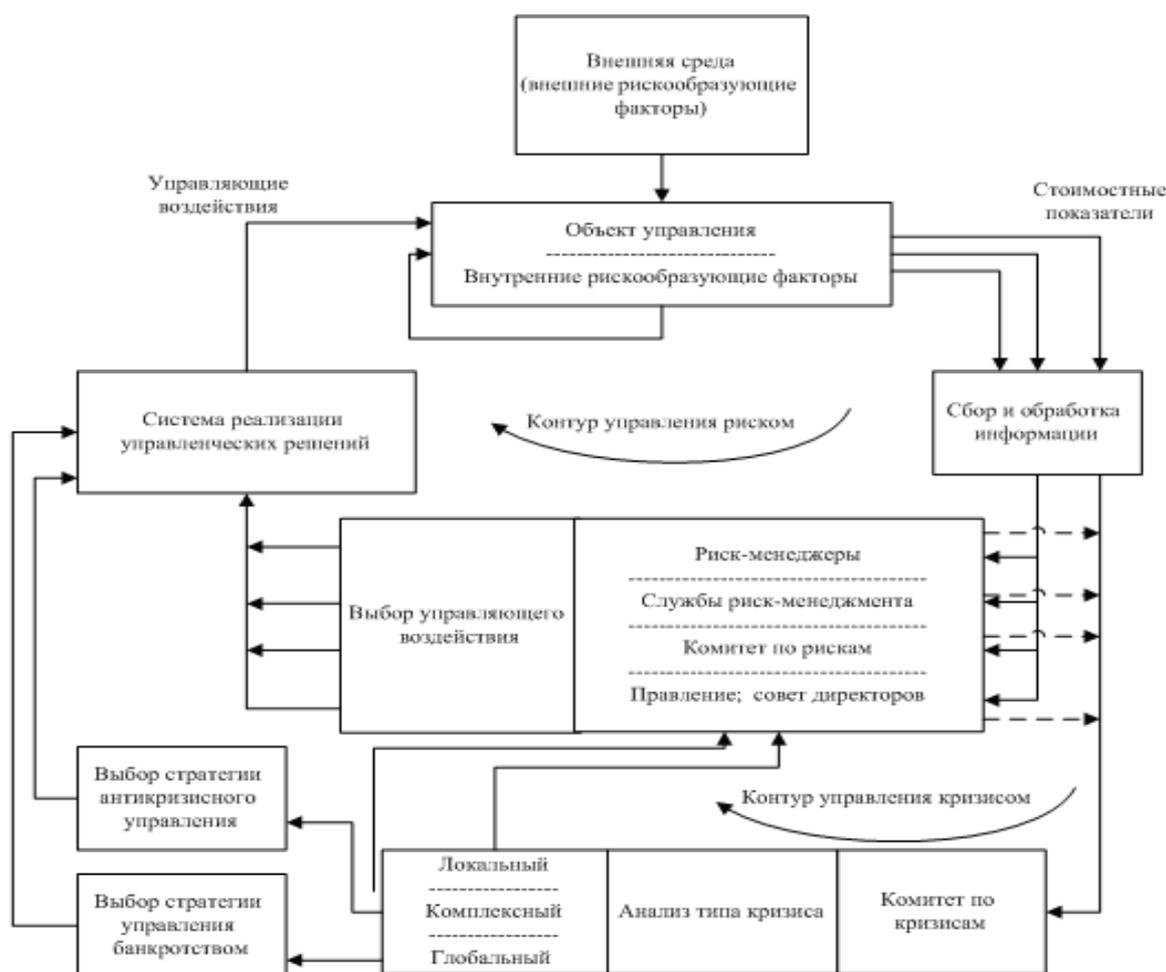


Рис. 2. Адаптивная система стратегического управления рисками

По решению правления или совета директоров (а также иных органов) создается *комитет по кризисам* в качестве субъекта управления. Комитет получает информацию об основных стоимостных показателях деятельности предприятия, а также о количестве реализовавшихся рисков и их уровне от субъектов управления в четырех контурах управления риском, в том числе: оперативных рисков от риск-менеджеров; тактических рисков от служб риск-менеджмента;

нормативно-стратегических рисков от правления; стратегических рисков от комитета по рискам. Таким образом, комитету по кризисам в качестве входной информации поступают данные в виде пары: u_1 – общее число реализовавшихся рисков, u_2 – процент охваченных кризисом стратегических бизнес-единиц в рамках предприятия (рис. 3).

Итоговой оценкой будем считать *тип кризиса* $T_k = f(u_1, u_2)$, где f – некоторая функция, характеризующая зависимость типа кризиса от двух входных переменных.

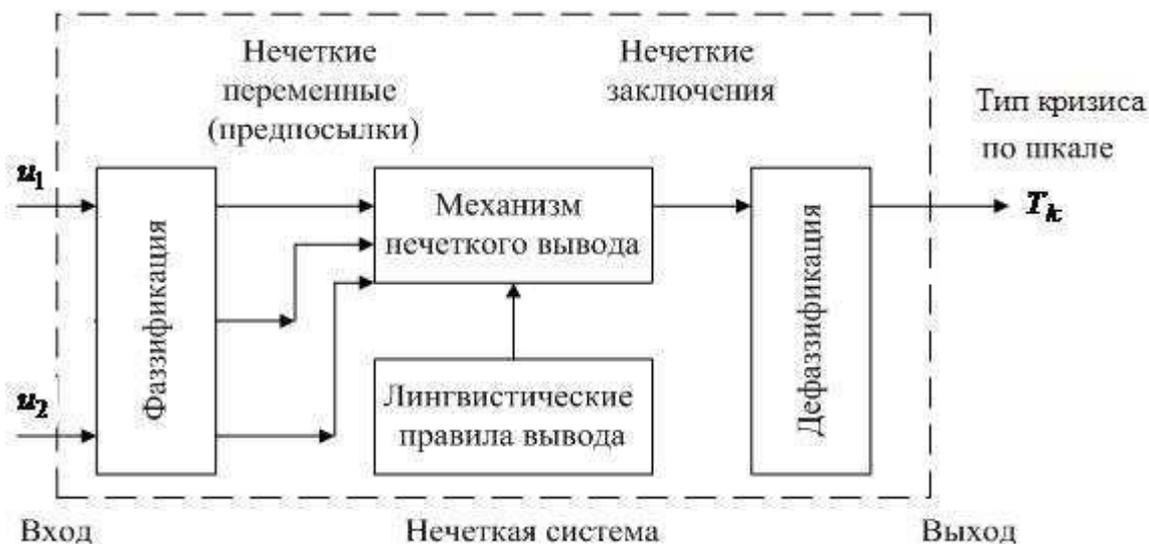


Рис. 3. Нечеткая система нахождения типа кризиса

Поскольку эксперты при принятии решений, как правило, используют лингвистические переменные, т.е. оперируют качественными категориями более чем количественными, применим для решения задачи определения типа кризиса теорию нечетких систем [5, 6]. Структура такой системы приведена на рис. 3.

В процессе функционирования нечеткой системы производятся:

1. *Фаззификация*. Входные четкие переменные (в данном случае u_1, u_2) преобразуются в нечеткие.

2. *Применение механизма нечеткого вывода*. Из базы нечетких правил вывода, связывающих входные нечеткие переменные (предпосылки) с выходными нечеткими переменными (заключениями), выбираются так называемые активные правила, применимые к данной ситуации. Результатом работы механизма нечеткого вывода являются нечеткие заключения.

3. *Дефаззификация*. Нечеткие заключения преобразуются в четкие (в данном случае целью является нахождение типа кризиса по некоторой заданной шкале).

Рассмотрим процедуру решения для типовой системы с n входами и одним выходом, схема которой изображена на рис. 4.

Допустим, что каждая входная переменная u_i принимает значения на своем универсальном множестве $U_i, i = 1, \dots, n$, а выходная переменная y – на универсальном множестве Y . Каждой переменной u_1, \dots, u_n поставим в соответствие лингвистическую переменную $\tilde{u}_1, \dots, \tilde{u}_n$, а переменной y – лингвистическую переменную \tilde{y} .

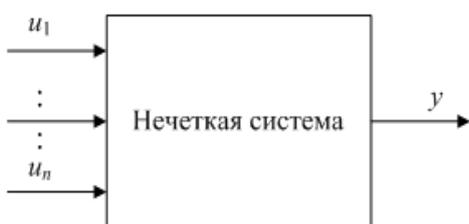


Рис. 4. Типовая система

В анализируемой задаче $n = 2$, а входные лингвистические переменные: \tilde{u}_1 – «общее число реализовавшихся рисков»; \tilde{u}_2 – «процент охваченных кризисом стратегических

бизнес-единиц». Выходная лингвистическая переменная \tilde{y} – «тип кризиса».

Универсальными множествами являются: $U_1 = [0; +\infty)$, $U_2 = [0; 100]$; $Y = [0; 6]$, так как значение u_1 неотрицательное, значение u_2 измеряется в процентах; для выходной переменной y принята шкала типа кризиса, включающая значения от нуля до 6.

Лингвистические переменные \tilde{u}_1, \tilde{u}_2 ; \tilde{y} принимают лингвистические значения. Будем использовать обозначения: \tilde{A}_i^j – j -е лингвистическое значение лингвистической переменной \tilde{u}_i ; \tilde{B}^j – j -е лингвистическое значение лингвистической переменной \tilde{y} . Таким образом, лингвистические переменные \tilde{u}_i, \tilde{y} принимают значения из множеств $\tilde{A}_i = \{\tilde{A}_i^j, j = 1, \dots, N_i\}$; $\tilde{B} = \{\tilde{B}^j, j = 1, \dots, M\}$, где N_i, M – число лингвистических значений соответствующих лингвистических переменных. Они называются терм-множествами.

В исследуемой задаче используются следующие лингвистические значения: для переменной \tilde{u}_1 – \tilde{A}_1^1 = “несущественное”; \tilde{A}_1^2 = “приемлемое”; \tilde{A}_1^3 = “недопустимое”; для переменной \tilde{u}_2 – \tilde{A}_2^1 = “малый”; \tilde{A}_2^2 = “ниже среднего”; \tilde{A}_2^3 = “средний”; \tilde{A}_2^4 = “выше среднего”; \tilde{A}_2^5 = “высокий”; для переменной \tilde{y} – \tilde{B}^1 = “локальный”; \tilde{B}^2 = “комплексный”; \tilde{B}^3 = “глобальный”. Таким образом, $N_1 = 3, N_2 = 5; M = 3$.

В дальнейшем вместо лингвистических значений для сокращения записи в некоторых случаях будем использовать просто номера переменных, т.е. в качестве терм-множеств переменных \tilde{u}_1, \tilde{u}_2 будем использовать $T_1 = \{1, 2, 3\}, T_2 = \{1, 2, \dots, 5\}$, а в качестве терм-множества переменной \tilde{y} множество $D = \{1, 2, 3\}$.

Для определения лингвистических переменных используются нечеткие множества и нечеткая логика. С каждым лингвистическим значением \tilde{A}_i^j переменной \tilde{u}_i связывается функция принадлежности $\mu_i^j(u_i) : U_i \rightarrow [0; 1]$. Ее величина описывает определенность, с которой элемент u_i множества U_i может быть ассоциирован с лингвистическим значением \tilde{A}_i^j . В данной работе применяются треугольные функции принадлежности [5, 6] (рис. 5).

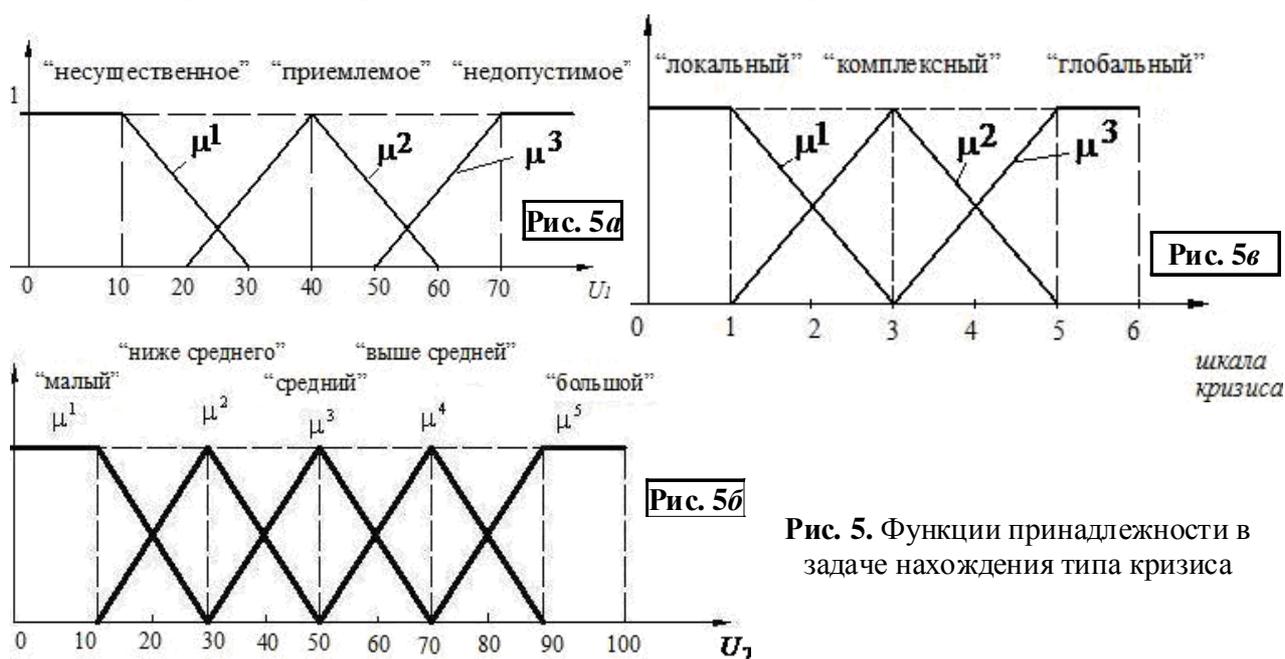


Рис. 5. Функции принадлежности в задаче нахождения типа кризиса

Функционирование нечетких систем опирается на лингвистические правила вывода, отображающие отображение входных переменных в выходные (рис. 4). Лингвистическое правило вывода обычно имеет следующую структуру (если предпосылки, то заключение): если \tilde{u}_1 есть \tilde{A}_1^j , и \tilde{u}_2 есть \tilde{A}_2^k , и ... \tilde{u}_n есть \tilde{A}_n^l , то \tilde{y} есть \tilde{B}^p .

В решаемой задаче правила имеют вид: если \tilde{u}_1 есть \tilde{A}_1^j и \tilde{u}_2 есть \tilde{A}_2^k , то \tilde{y} есть \tilde{B}^p , где $j = 1, \dots, N_1; k = 1, \dots, N_2; p = 1, \dots, M$. Например, если общее число реализовавшихся рисков “несущественное”, процент охваченных кризисом бизнес-единиц “средний”, то тип кризиса y – “комплексный”.

Будем предполагать, что формируется R различных правил, представляемых в форме $(j, k; p)_i, i = 1, \dots, R$, где $j = 1, \dots, N_1; k = 1, \dots, N_2; p = 1, \dots, M$. Если все правила имеют описанную структуру, то их общее число $R = N_1 \cdot N_2$.

Для решаемой задачи удобно формализовать составление лингвистических правил в форме табл. 1.

Таблица 1

Тип кризиса T_k	“Процент охваченных кризисом стратегических бизнес-единиц” = \tilde{u}_2					
		1	2	3	4	5
“Общее число реализовавшихся рисков” = \tilde{u}_1	1	1	1	2	2	2
	2	1	2	2	2	3
	3	2	2	2	3	3

2. Процедура функционирования нечеткой системы

1. *Фаззификация.* На вход системы поступают значения u_1, u_2 . Нечеткая система преобразует эти значения в нечеткие множества. Для этого находят значения всех функций принадлежности по каждой переменной.

2. *Установление соответствий:* определение правил, которые применяются в текущей ситуации, характеризуемой значениями входных переменных. При этом находится степень уверенности в применении правил вывода из имеющегося набора лингвистических правил. С этой целью находится значение функции принадлежности для предпосылки, входящей в правило вывода

$$\mu_{предп(i)} = \mu_i(u_1, u_2, \dots, u_n) = \mu_{A_1^j}(u_1) * \mu_{A_2^k}(u_2) * \dots * \mu_{A_n^l}(u_n).$$

Считается, что i -е правило активно в данный момент, если выполняется условие $\mu_i(u_1, u_2, \dots, u_n) > 0$. Это условие выполняется только в том случае, когда значения функций принадлежности для всех входящих в предпосылку переменных больше нуля. Правила, для которых $\mu_i(u_1, u_2, \dots, u_n) = 0$, называются пассивными. Заметим, что если графики соседних функций принадлежности пересекаются, то в системе с двумя входными переменными максимальное число активных в данный момент правил вывода равно 4.

Для определения степени уверенности в применении i -го правила при решении данной задачи ($n = 2$) будем применять два способа [5].

Первый способ. В качестве операции $*$ применяется операция \min
 $\mu_{предп(i)} = \mu_i(u_1, u_2) = \min\{\mu_{A_1^j}(u_1), \mu_{A_2^k}(u_2)\}.$

Второй способ. В качестве операции $*$ применяется операция “алгебраического произведения”
 $\mu_{предп(i)} = \mu_i(u_1, u_2) = \mu_{A_1^j}(u_1) \cdot \mu_{A_2^k}(u_2).$

3. *Нахождение выводимого нечеткого множества.* Для каждого активного в данный мо-

мент правила найти *выводимое нечеткое множество* с функцией принадлежности $\mu_{B(i)}(y) = \mu_i(u_1, \dots, u_n) * \mu_{B^p}(y) = \mu_{предн(i)} * \mu_{B^p}(y)$, где для получения результата используется i -е правило вида $(j, k; p)_i, i \in \{1, \dots, R\}$.

Будем использовать два способа.

Первый способ. В качестве операции $*$ применить операцию \min $\mu_{B(i)}(y) = \min\{\mu_{предн(i)}, \mu_{B^p}(y)\}$. Тогда процедура сводится к “отсечению” от графика функции принадлежности $\mu_{B^p}(y)$ верхней части выше уровня, определяемого значением $\mu_{предн(i)}$. Заметим, что величина $\mu_{предн(i)}$ характеризует степень уверенности в применении i -го правила. Процедура нахождения выводимого множества вытекает из следующего рассуждения: нельзя быть более уверенным в заключении, чем в предпосылках.

Второй способ. В качестве операции $*$ применить операцию “алгебраического произведения” $\mu_{B(i)}(y) = \mu_{предн(i)} \cdot \mu_{B^p}(y)$.

4. *Агрегирование.* Предлагается применять два способа, отражающие два различных подхода к нахождению результата нечеткого вывода.

Первый способ. Рассматривается *совокупность всех выводимых нечетких множеств* с функциями принадлежности $\mu_{(1)}(y), \mu_{(2)}(y), \dots, \mu_{(R)}(y)$, полученными в результате применения каждого i -го правила. Далее эта совокупность обрабатывается с помощью процедуры дефаззификации.

Второй способ. Получить *обобщенное выводимое нечеткое множество* \tilde{B} с функцией принадлежности $\mu_B(y) = \mu_{(1)}(y) \oplus \mu_{(2)}(y) \oplus \dots \oplus \mu_{(R)}(y)$, где R - общее число лингвистических правил вывода. Это множество отражает результат применения всех правил вывода, имеющих в базе. В качестве операции \oplus будем применять операцию \max .

5. *Дефаззификация.* Целью данной операции является выработка четкого значения выходной переменной y . Методы дефаззификации соответствуют применяемому способу агрегирования.

Если применяется *первый способ агрегирования*, т.е. рассматривается совокупность функций принадлежности $\mu_{(1)}(y), \mu_{(2)}(y), \dots, \mu_{(R)}(y)$, полученных после применения всех правил вывода (на практике чаще не всех, а только активных), то используются:

- *метод центра тяжести*, согласно которому значение выходной переменной находится

$$\text{по формуле } y = \frac{\sum_{i=1}^R b_i \cdot \int_Y \mu_{(i)}(y) dy}{\sum_{i=1}^R \int_Y \mu_{(i)}(y) dy}, \text{ где } b_i - \text{“центр площади” функции принадлежности лингвистического значения } \tilde{B}^p, \text{ соответствующего выводимому множеству } i\text{-го правила вывода } (j, k; p)_i, i \in \{1, \dots, R\}, \int_Y \mu_{(i)}(y) dy - \text{площадь, ограниченная графиком функции принадлежности } \mu_{(i)}(y) \text{ и осью абсцисс;}$$

• *метод усредненного центра*, согласно которому значение выходной переменной находится по формуле $y = \frac{\sum_{i=1}^R b_i \cdot \sup_Y \mu_{(i)}(y)}{\sum_{i=1}^R \sup_Y \mu_{(i)}(y)}$, где $\sup_Y \mu_{(i)}(y)$ - наибольшее значение функции принадлежности.

дится по формуле $y = \frac{\sum_{i=1}^R b_i \cdot \sup_Y \mu_{(i)}(y)}{\sum_{i=1}^R \sup_Y \mu_{(i)}(y)}$, где $\sup_Y \mu_{(i)}(y)$ - наибольшее значение функции принадлежности.

надлежности.

Если применяется *второй способ агрегирования*, т.е. обработке подлежит функция принадлежности $\mu_B(y) = \mu_{(1)}(y) \oplus \mu_{(2)}(y) \oplus \dots \oplus \mu_{(R)}(y)$, соответствующая обобщенному выводимому нечеткому множеству, то используются:

- *метод максимума функции принадлежности*, соответствующей обобщенному выводимому множеству: $y \in \arg \sup_Y \mu_B(y)$, т.е. в качестве выхода y выбирается значение, при кото-

ром функция принадлежности достигает своей наибольшей величины. Если имеется несколько таких значений y , то требуется добавить дополнительную процедуру выбора (например, среди найденных значений определяется наименьшее по модулю и т.п.);

- *метод среднего максимума*, где в качестве y выбирается среднее значение всех элементов, на которых достигается максимальное значение функции принадлежности, соответствующей обобщенному выводимому множеству

- *метод центра площади*, где в качестве y выбирается
$$y = \frac{\int_Y y \cdot \mu_B(y) dy}{\int_Y \mu_B(y) dy}.$$

Заключение

Рассмотрена задача нахождения типа кризиса в функциональной системе управления рисками и кризисами на предприятии АПК. Для ее решения предложен алгоритм формирования нечеткой системы, служащей для нахождения типа кризиса по соответствующей шкале на основании лингвистических значений процента охваченных кризисом стратегических бизнес-единиц (или подразделений предприятия) и общего числа реализовавшихся рисков.

ЛИТЕРАТУРА

1. Бадалова А.Г. Управление рисками производственных систем: теория, методология, механизмы реализации. – М.: ИЦ МГТУ “Станкин”, “Янус-К”, 2006.
2. Бадалова А.Г. Система управления рисками: методология, организационно-информационное обеспечение, эффективность внедрения. – М.: ИЦ МГТУ “Станкин”, “Янус-К”, 2007.
3. Бадалова А.Г., Пантелеев П.А. Функциональная схема системы стратегического управления риском предприятия авиационно-промышленного комплекса на основе стоимостного подхода // Электронный журнал “Труды МАИ”. - 2011. - № 42.
4. Бадалова А.Г., Пантелеев П.А. Формирование системы нечеткого вывода для нахождения уровня риска на предприятии авиационно-промышленного комплекса // Электронный журнал “Труды МАИ”. - 2011. - № 42.
5. Passino K.M., Yurkovich S. Fuzzy control, Addison Wesley, 1998.
6. Заде Л. Понятие лингвистической переменной и его применение к принятию приближенных решений.- М.: Мир, 1976.
7. Кошкин В.И. и др. Антикризисное управление: 17-модульная программа для менеджеров «Управление развитием организации». Модуль 11. - М.: ИНФРА-М, 1999.
8. Бланк И.А. Управление финансовой безопасностью предприятия / И.А. Бланк. - изд. 2-е стереотипное. - К: Эльга, 2009. - (Серия «Библиотека финансового менеджера»).

APPLICATION OF FUZZY INFERENCE FOR DETERMINING THE TYPE OF CRISIS IN ADAPTIVE SYSTEM STRATEGIC RISK MANAGEMENT

Badalova A.G., Panteleev P.A.

The problem of determining the type of crisis in the Avia-industrial complex (AIC), which appears in the absence or poor functioning of the risk management system, resulting, in turn, is the escalating risks in the crises of the enterprise. To

solve this problem the authors proposed algorithm for forming fuzzy system, which is used in order to find the type of crisis on the grade on the basis of linguistic values of the percentage affected by the crisis of strategic business units and the total number of realized risks.

Key words: risk, crisis, the type of crisis, fuzzy sets, fuzzy system, fuzzy inference, linguistic variables, fuzzification, defuzzification.

Сведения об авторах

Бадалова Анна Георгиевна, окончила МАИ (1981), доктор экономических наук, профессор МАИ (государственного технического университета), автор более 50 научных работ, область научных интересов – управление рисками производственных систем, управление развитием.

Пантелеев Петр Андреевич, 1986 г.р., аспирант МАИ, автор 4 научных работ, область научных интересов – математические и инструментальные методы в экономике, конструирование авиационных систем.

УДК 621.396.6

СИСТЕМА УПРАВЛЕНИЯ РАЦИОНАЛЬНЫМ РАСПРЕДЕЛЕНИЕМ РЕСУРСОВ ОРГАНИЗАЦИОННО-ТЕХНИЧЕСКОЙ СИСТЕМЫ

Ю.Т. ЗЫРЯНОВ, О.А. КОНОВАЛОВ, К.А. МАЛЫКОВ

В статье рассмотрена задача рационального распределения ограниченных ресурсов по зависимым операциям. Предложена система управления рациональным распределением ресурсов, которая позволяет осуществлять управление и контроль выполнения проекта в организационно-технической системе.

Ключевые слова: организационно-техническая система, система управления, ограничения ресурсов.

Современный уровень развития телекоммуникационной и электронно-вычислительной техники позволяет автоматизировать широкий класс функций управления распределением ресурсов. Для этого необходима разработка алгоритмов решения комплекса взаимосвязанных задач, составляющих систему управления (СУ), и определение соответствующих ограничений. При использовании моделей и методов сетевого планирования и управления (СПУ) возникает задача рационального распределения ресурсов по множеству зависимых операций в организационно-технической системе (ОТС) [1, 2].

Известные методы недостаточно эффективны для зависимых операций. Оптимизация сетевых моделей без календарной увязки сроков отрицательно сказывается на достоверности и ограниченности результатов. Использование эвристических приемов на стадии планирования превращает метод СПУ в метод планирования. После получения первоначального плана распределения ресурсов система становится неуправляемой методами СПУ, так как ее параметры теряют свой смысл. Становится важным решение задачи составления календарного плана с учетом рационального распределения ресурсов и оборудования в динамически меняющейся обстановке и возможность его контроля и регулирования методами СПУ.

В статье используется терминология теории графов и управления проектами. Проектом будем называть некоторый процесс изменений, то есть не рутинный, не повторяющийся процесс, требующий специальных методов проектного управления [2].

Необходимо выполнить проект Y по подготовке к функционированию некоторой сложной системы, состоящий из множества зависимых и независимых операций, технологическая зависимость между которыми задается в виде сети. Для этого предлагается рассматривать детерминированную модель сетевого графика, для которой необходимо задать первоначальный план распределения ограниченных ресурсов. Каждая операция выполняется с переменной интенсивностью однородными или разнородными исполнителями, распределение которых может быть неоднократным. Перераспределение ресурсов с одной операции на другую не связано с временными затратами и операции можно прерывать до их окончания. Результатом решения является ресурсный граф переменной структуры.

Рассмотрим проект Y , состоящий из набора зависимых и независимых операций, технологическая зависимость между которыми задается в виде сети. На реализацию проекта выделено ограниченное количество ресурсов S различных типов h , каждый из которых назначается $f_{(ij)}$ на выполнение $l_{(ij)}$ -й операции и может ее выполнить $f_{(ij)}^*$. Условимся, что событие j непосредственно следует за событием i . L – множество операций заранее заданных списком, $(ij) \in L$; E^f – множество целевых событий проекта, для каждого из которых срок выполнения превышает заданный; H и Q – множество операций сетевого графа, с каждой из которых будут сниматься и соответственно добавляться исполнители хотя бы одного h -го типа, $h \in f_{(ij)}^{(0)}$;

$f_{(ij)}^{(0)}$ – множество различных типов, исполнители каждого из которых будут переназначаться в процессе решения задачи с одной операции на другую; $\Delta_{(ij)}^*$ – резерв времени (ij) -й операции при заданном распределении с учетом директивного срока γ -го целевого события, $\gamma \in E^f$; $N_{h(ij)}$ и $N_{h(ij)}^*$ – минимально и максимально возможное число исполнителей h -го типа соответственно, $h \in f_{(ij)}$ и $h \in f_{(ij)}^*$; $C_{(ij)}$ – доля выполненной части $l_{(ij)}$ -й операции; $Z_{(ij)}$ – множество ресурсных условий для $l_{(ij)}$ -й операции; $x_{(ij)}$ – количество исполнителей, приступивших к выполнению $l_{(ij)}$ -й операции; $\lambda_{h(ij)}$ – производительность одного исполнителя h -го типа; K_γ – суммарный коэффициент, зависящий от факторов внешней среды; T_γ^{dup} и T_γ^{kp} – директивный и критический сроки γ -го целевого события соответственно; $t(\alpha_{(ij)}^*)$ и $t(\alpha_{(ij)}^{(0)})$ – продолжительность выполнения максимального пути, проходящего через $l_{(ij)}$ -ю операцию множества $Q^{(f)}$ и H соответственно; $t_{(ij)}$ – планируемая продолжительность выполнения операции при заданном распределении исполнителей; $t_{\gamma(ij)}^{\min}$ и $t_{\gamma(ij)}^{\max}$ – минимальное и максимальное возможное время выполнения $l_{(ij)}$ -й операции соответственно.

Для оптимального распределения разнородных исполнителей по операциям сетевого графика $x_{h(ij)}(Z_{(ij)})$ найдем такое их перераспределение, $x_{h,i_1(ij)}^*(Z_{(ij)})$, $i_1 = 1, 2, \dots, M_{(ij)}^*$, $h \in f_{(ij)}^*$ ($M_{(ij)}^*$ – число частей $l_{(ij)}$ -й операции), которое обеспечит максимум функции (1) в области ограничений (3)-(7), а для выполнения проекта за минимальное время – минимум функции (2) в области ограничений (3)-(7).

$$F_\Sigma^* = - \sum_{\gamma \in E^f} \left(\sum_{(ij) \in T_\gamma^{kp}} \sum_{i_1=1}^{M_{(ij)}^*} \frac{C_{\gamma i_1(ij)}^* (1 + K_{\gamma r})}{\sum_{h \in f_{(ij)}} \lambda_{\gamma h i_1(ij)} \cdot x_{\gamma h i_1(ij)}^*(Z_{(ij)})} - T_\gamma^{dup} \right) = - \sum_{\gamma \in E^f} \left(\sum_{(ij) \in T_\gamma^{kp}} \sum_{i_1=1}^{M_{(ij)}^*} F_\gamma - T_\gamma^{dup} \right); \quad (1)$$

$$F_\Sigma^*(T_\gamma^{dup}) = \sum_{\gamma \in E^f} \left| \sum_{(ij) \in Q^{(f)}} \left(\sum_{(ij) \in t(\alpha_{(ij)}^*)} \sum_{i_1=1}^{M_{(ij)}^*} F_\gamma - T_\mu^{dup} \right) - \sum_{(ij) \in H} \left(T_\gamma^{dup} - \sum_{(ij) \in t(\alpha_{(ij)}^{(0)})} \sum_{i_1=1}^{M_{(ij)}^*} F_\gamma \right) \right|. \quad (2)$$

Технологические, организационные и ресурсные ограничения.

$$t_{\gamma(ij)}^{\min} \leq t_{\gamma(ij)} \leq t_{\gamma(ij)}^{\max}, \gamma \in E^f, (ij) \in L; \quad (3)$$

$$\begin{cases} t_{\gamma(ij)}^{\min} \leq \sum_{i_1=1}^{M_{(ij)}^*} t_{\gamma i_1(ij)}^* \leq t_{\gamma(ij)} + \Delta_{\gamma(ij)}^*, & \gamma \in E^f, (ij) \in H, \\ t_{\gamma(ij)}^{\min} \leq \sum_{i_1=1}^{M_{(ij)}^*} t_{\gamma i_1(ij)}^* \leq t_{\gamma(ij)}, & \gamma \in E^f, (ij) \in (Q/H). \end{cases} \quad (4)$$

$$N_{h(ij)} \leq x_{\gamma h i_1(ij)}^*(Z_{(ij)}) \leq N_{h(ij)}^*, h \in f_{(ij)}^*, (ij) \in L; \quad (5)$$

$$\begin{cases} [x_{\gamma h i_1(ij)}^*(Z_{(ij)})] = x_{\gamma h i_1(ij)}^*(Z_{(ij)}), \\ \gamma \in E^f, h \in f_{(ij)}^*, i_1 = 1, 2, \dots, M_{(ij)}^*, (ij) \in Q; \end{cases} \quad (6)$$

$$\begin{cases} \sum_{(ij) \in H} \sum_{\gamma=1}^{x_{\gamma h(ij)} - N_{\gamma h(ij)}} x_{\gamma h \rho(ij)}^{ch}(Z_{(ij)}) t_{\gamma h \rho(ij)}^{ch} = \sum_{(ij) \in Q} \sum_{\gamma=1}^{N_{\gamma h(ij)} - x_{\gamma h(ij)}} x_{\gamma h \rho(ij)}^{naz}(Z_{(ij)}) t_{\gamma h \rho(ij)}^{naz}; \\ \gamma \in E^f, h \in f_{(ij)}^{(0)}. \end{cases} \quad (7)$$

Здесь $x_{(ij)}^{ch}(Z_{(ij)})$ – число снятых исполнителей за время $t_{(ij)}^{ch}$ с $l_{(ij)}$ -й операции, а $x_{(ij)}^{has}(Z_{(ij)})$ – число назначенных исполнителей за время $t_{(ij)}^{has}$ на выполнение $l_{(ij)}$ -й операции соответственно на ρ -й итерации решения задачи, $(ij) \in L$.

Физический смысл функции (1) состоит в том, что для целевых событий определяется разность между критическим и директивным сроком. Просуммировав такие разности по всем целевым событиям, получим функцию (1). В физическом смысле функция (2) означает абсолютную величину суммарного резерва всех операций сетевой модели. Решение задачи с использованием функции (1) сведено к проблеме невыпуклого целочисленного программирования с целевой функцией, численное значение которой зависит от структуры сетевой модели и распределения исполнителей по операциям проекта. Оптимизация сетевой модели с использованием функции (2) выполняется методом наискорейшего градиентного спуска.

При выполнении проекта Y необходимо получить такой план рационального распределения ресурсов Ω^* (8) при фиксированном количестве специалистов x и их различных типов h , который позволит в области ограничений (3)-(7) рационально распределить ресурсы за минимальное время T [3]. Таким образом, основным критерием эффективности является минимизация временных затрат на выполнение проекта

$$[\Omega^*]: T[\Omega^* |_{x,h}] = \min_{\substack{\Omega \in Y \\ x \in X}} T[\Omega |_{x,h}], \text{ при } T < T^{dup}. \quad (8)$$

Под рациональным планом распределения ресурсов понимается такой план, в котором ресурсы распределены так, что при выполнении каждой операции сетевой модели обеспечивается максимальное приращение целевой функции (1) в области (2)-(7) за счет использования резервов времени некритических операций.

Для решения задачи применяется метод последовательных назначений, с процедурой типа динамического программирования, согласно которой состояние системы изменяется в соответствии с одношаговой функцией переходов. Результатом является ресурсный граф переменной структуры. Алгоритм заключается в выполнении следующих этапов:

1. Значение функции (1) на первом шаге принимается равной нулю.
2. Составляется первоначальный план распределения ресурсов в виде матрицы с целочисленными неотрицательными элементами для обеспечения максимума функции (1) при заданных вариантах последовательностей.
3. Вычисляется возможное приращение функции (1).
4. Определяется номер операции, которая обеспечивает максимум функции (1) и производится назначение на нее соответствующих исполнителей.
5. Формируется новый план распределения ресурсов.
6. Вычисляется максимум функции (1).
7. Полученное значение функции (1) сравнивается с первоначальным (до переназначения), и если условие выполняется, то полученный план принимается за рациональный.

Однако такой алгоритм дает значительную погрешность при переназначении специалистов. В связи с этим авторами предлагается модернизированный метод последовательных назначений, который в общем виде реализуется по следующему алгоритму:

1. Исходным планом распределения является полученный план распределения ресурсов по предыдущему алгоритму.
2. Выбирается вариант последовательности номеров типов ресурсов.
3. Рассматривается ряд вариантов последовательностей номеров типов ресурсов и выбирается из них тот, который обеспечивает максимум функции (1).
4. Выполняется деление операции на η частей (рациональным является разделение операции от 3 до 10 частей) и последовательное назначение ресурсов с целью обеспечения максимума приращения функции (1) на каждой части операции.

5. Выполняется проверка целесообразности перераспределения исполнителей на другие части операций. Если условие не выполняется, то исполнители не перераспределяются, если даже такое распределение обеспечивает максимум функции (1).

6. Производится последовательное назначение каждой единицы ресурсов по выбранному варианту последовательности, обеспечивая максимальное приращение функции (1).

7. Определяется максимальное приращение функции (1) при переназначении исполнителей h -го типа с одной операции на другую, и определяется номер операции, с которой снимаются исполнители, и номер операции, на которую они назначаются.

8. Если приращение функции (1) является неотрицательной величиной, то переназначение исполнителей на операции, лежащие на критическом пути, производить целесообразно.

9. Расчет выполняется последовательно и повторяется до тех пор, пока значение приращения функции (1) не станет отрицательной величиной.

10. Затем каждое значение полученного плана распределения умножается на $(1/\eta)$. Полученный новый план будет являться улучшенным планом рационального распределения.

Алгоритм распределения ограниченных ресурсов организационно-технической системы ОТС по множеству зависимых операций представлен в [4]. В [3] показано решение задачи распределения ресурсов с применением модернизированного метода последовательных назначений, а также приведены этапы решения задачи перераспределения ресурсов. При этом следует заметить, что метод последовательных назначений дает оптимальное решение задачи перераспределения исполнителей по операциям сетевой модели за счет использования полных резервов времени работ для случая, когда каждая (ij) -я операция проекта выполняется только одним h -м типом исполнителей, $h = 1, 2, \dots, S$ [5].

Предложенный алгоритм реализован при разработке СУ проектами, ограниченную по времени организационными и техническими ресурсами (СУРРР) для решения прикладной задачи распределения ресурсов [6]. Параметры СУРРР должны иметь тот же смысл, что и в СУ без учета ресурсов. Создание СУРРР обосновано следующим утверждением.

Утверждение. Для того чтобы параметры СУРРР сохраняли тот же смысл, что и в СУ (без учета ресурсов) необходимо и достаточно между операциями сетевой модели установить, кроме технологических связей, связи по ресурсам. Справедливость утверждения достаточно доказать для главного параметра системы – критического пути при постоянном качестве выполнения операций.

Критическим путем называется путь, ведущий от множества входных операций к множеству выходных операций и состоящий из критических операций [7]. Операция называется критической, если ее резерв равен нулю [2].

Необходимость. Пусть в СУРРР критический путь имеет то же самое толкование, что и в СУ (без учета ресурсов). Доказательство проведем от противного. Предположим, что связи по ресурсам между операциями в СУРРР не установлены. В случае дефицита ресурсов суммарная продолжительность выполнения операций, которые обеспечиваются одними и теми же исполнителями, будет больше критического пути, определенного только на основе технологических связей между операциями сетевой модели. Тогда критический путь не определит срок окончания всего проекта, что противоречит утверждению.

Достаточность. Определив критический путь в СУРРР как самый длинный путь на всем множестве путей, образуемых технологией и распределением исполнителей, обеспечивающих выполнение проекта, очевидно, что параметры СУРРР будут иметь тот же смысл, что и параметры СУ (без учета ресурсов).

Из доказанного утверждения следует следующее основное правило построения сетевой модели в СУРРР: в сетевой модели ориентация дуг графа осуществляется в соответствии с технологией этого процесса, а также распределением ресурсов, обеспечивающих его выполнение. Структура СУРРР представлена на рис. 1.

При их несовпадении разница учитывается при очередном цикле функционирования системы, или путем вывода информации из системы.

Разработанная СУРРР отличается тем, что дополнительно введены на начальный момент времени T_1 устройство 4, ячейки 5, 11, 12, 14, 15, 16, 17, а на конечный момент времени T_2 – блок 24 и ячейки 18, 19, 25, 26 (наименования ячеек и блоков приведены на рис. 1). В основу функционирования СУРРР положена методика освоенного объема [9].

На момент времени T_1 в блоке 10 имеется в наличии выделенное количество ресурсов, значение величины которых из ячейки 3 переводится в ячейку 21 блока 2, где осуществляется их фиксирование. На вход 1 блока 1 может дополнительно поступать информация о количестве специалистов, трудоемкости операций, наличии материальных и технических ресурсов, воздействиях внешней среды и т.д. Объем выполненных операций на момент времени T_2 передается в ячейку 18 блока 2, где осуществляется копирование запомненной величины запланированных операций в ячейку 11 блока 1. Величина возврата ресурсов вводится оператором по входу 2 или 3 блока 2 соответственно, вследствие чего перестраивается сетевая модель и производится пересчет параметров модели. При этом в блоке 2 обратные связи с блоком 1 осуществляются по трем отдельным направлениям. Первое направление относится к организации обратных связей по возврату ресурсов, второе – к выполнению операций, третье – к оптимальному перераспределению ресурсов. За счет организации функциональных связей между ячейками блока 2 и блока 1 реализуется асимметричная связь плоскости времени T_2 с T_1 по возврату ресурсов посредством связи ячейки 6, 7 и 21, 22 блоков 1 и 2 соответственно. Информация о ходе выполнения операций поступает входу 3 на вход ячейки 22 и переносится в ячейку сигнализации 23 и далее на выход из системы 2. При этом асимметрично связывается плоскость времени T_2 с плоскостью T_1 , т.е. организуется связь между ячейками 11, 12 и 18, 19 блоков 1 и 2 соответственно. Из ячейки 22 соответствующая информация поступает в ячейку 7 и в элемент 8. При перераспределении ресурсов по входу 2 блока 2 в ячейку 19 поступает информация о ходе выполнения операций и затем переносится по соответствующей связи в ячейку 12 и в элемент 13 блока 1. Для получения оптимального плана распределения ресурсов в системе предусмотрен блок 24, в котором учитывается доступное количество ресурсов в системе, отслеживается ход выполнения операций и рассчитываются параметры сетевой модели. Перераспределение ресурсов осуществляется с помощью разработанного алгоритма для СУРРР [6].

За счет организации функциональных связей между ячейками блоков 2 и 1 реализуется асимметричная связь плоскости времени T_2 с T_1 по рациональному распределению и перераспределению ресурсов посредством связи ячеек 15, 16, 17 блока 1 с блоком 24 и ячейкой 25 блока 2. С выходов ячейки 25 и устройства 4 (понимается ЭВМ) зафиксированные параметры по выбранному критерию ячейки 3 поступают в ячейку 26, в которой проверяется наличие ресурсов и возможность их взаимозаменяемости. В интервале времени T_1 – T_2 для осуществления контроля и планирования предусмотрены средства сравнения плановых показателей с фактическими посредством ячеек 9 и 14. При их несовпадении возникает сигнал рассогласования и учитывается их разница в ячейке сигнализации 20 при очередном цикле функционирования и сигнал поступает на выход из системы 2. Цель проекта достигается за счет совпадения освоенного и суммарного объема операций по проекту. В [9] подробно рассмотрены показатели освоенного объема и производственные суммарные показатели освоенного объема, используемые для описания проекта, а также приведены частные модели для СУ.

Таким образом, в работе предложено решение задачи распределения ограниченных ресурсов по зависимым операциям в ОТС с применением модернизированного метода последовательных назначений. Предложенная СУРРР позволяет осуществлять управление ресурсами, а также контролировать ход выполнения проекта. Рассмотренный подход распределения ресурсов может использоваться при оперативном развертывании техники, ремонте, техническом обслуживании, составлении расписаний и в других задачах управления ОТС.

ЛИТЕРАТУРА

1. **Зырянов Ю.Т., Малыков К.А.** Управление профилактикой в организационно-технических системах: монография / под общ. ред. Ю.Т. Зырянова. – М.: АСТ-ПРЕСС КНИГА, 2005.
2. **Бурков В.Н., Ланда Б.Д., Ловецкий С.Е.** Сетевые модели и задачи управления. – М.: Сов. радио, 1967.
3. **Зырянов Ю.Т., Коновалов О.А.** Алгоритм распределения ресурсов по множеству зависимых операций // Информационно-измерительные и управляющие системы. - 2009. - № 10.
4. Распределение ограниченных ресурсов организационно-технической системы военного назначения по множеству зависимых операций (программа для ЭВМ) / **Зырянов Ю.Т., Сухоруков С.Е., Коновалов О.А.** // Свид. РФ.-№ 2007612327; заявл. 13.04.07; зарегистрировано в Реестре программ для ЭВМ Роспатента 4.06.07.
5. **Костина Л.П.** Постановка проблемы оптимального распределения ресурсов на стохастических сетях со сложной пространственно-временной структурой // Вестник Санкт-Петербургского университета. - 1992. - Сер. 1. - Вып. 2 (8).
6. Алгоритмическое обеспечение системы управления распределением ресурсов в организационно-технической системе по зависимым операциям с переменной интенсивностью выполнения работ (программа для ЭВМ) / **Коновалов О.А., Зырянов Ю.Т.** // Свид. РФ.-№ 2009611379; заявл. 30.04.09; зарегистрировано в Реестре программ для ЭВМ Роспатента 22.05.09.
7. **Бурков В.Н., Моисеенко Г.Е.** Вопросы анализа и оптимизации комплексов операций при учете перемещений ресурсов // Автоматика и телемеханика. – 1969. – №12.
8. Патент 2375750 Российская Федерация, МКИ ⁷ G 06 Q 10/00. Система рационального распределения ресурсов / **Коновалов О.К., Зырянов Ю.Т., Малыков К.А.**; заявители и патентообладатели Коновалов О.А., Зырянов Ю.Т., Малыков К.А. – №2008111831/09; заявл. 19.12.09; опубли. 10.12.09, Бюл. № 34.
9. **Колосова Е.В., Новиков Д.А., Цветков А.В.** Методика освоения объема в оперативном управлении проектами. - М.: ООО «НИЦ «Апостроф», 2000.

**THE SYSTEM OF MANAGEMENT OF RATIONAL RESOURCES
DISTRIBUTION OF ORGANIZATIONAL-TECHNICAL SYSTEM****Zyryanov Y.T., Konovalov O.A., Malykov K.A.**

The problem of rational distribution of the limited resources on depended operations is considered. The system of management of rational resources distribution (SMRRD) is offered by rational distribution of resources which allows to exercise management and the control of the performance project in organizational and technical system.

Key words: organizational-technical system, control system, restrictions of resources.

Сведения об авторах

Зырянов Юрий Трифионович 1960 г.р., окончил Тамбовское ВВАИУ (1983), доктор технических наук, доцент, профессор кафедры конструирования радиоэлектронных и микропроцессорных систем ФГУ ВПО Тамбовского технического государственного университета, автор свыше 100 научных работ, область научных интересов – управление состоянием организационно-технических систем при ограниченных ресурсах.

Коновалов Олег Анатольевич 1982 г.р., окончил Тамбовское ВВАИУ (2005), кандидат технических наук, преподаватель кафедры организации связи (и технической эксплуатации средств связи) ВАИУ (г. Воронеж), автор 16 научных работ, область научных интересов – распределение ресурсов по зависимым операциям в организационно-технических системах военного назначения.

Малыков Константин Анатольевич 1965 г.р., окончил Тамбовское ВВАИУ (1987), кандидат технических наук, доцент, заместитель начальника кафедры организации связи (и технической эксплуатации средств связи) ВАИУ (г. Воронеж), автор свыше 50 научных работ, область научных интересов – совершенствование программ технической эксплуатации сложных радиотехнических систем.

УДК 656.7.052:004.4

АНАЛИЗ СЕЗОННОГО РАСПИСАНИЯ НА СООТВЕТСТВИЕ НОРМАТИВАМ ПРОПУСКНОЙ СПОСОБНОСТИ

В.В. БЫКОВА, О.Н. МОИСЕЕВ, Л.Е. РУДЕЛЬСОН, М.А. ЧЕРНИКОВА

Оптимизационные модели организации потоков воздушного движения для построения бесконфликтного сезонного расписания недостаточно эффективны вследствие неизбежных флуктуаций переменных в процессе его реализации. Предпочтителен поиск устойчивого локального экстремума функции (максимум удовлетворенных заявок) при ограничениях на показатели безопасности, регулярности и экономичности полетов.

Ключевые слова: сезонное расписание, использование воздушного пространства, имитационная модель.

Введение

В целях обеспечения безопасности, регулярности и экономичности полетов в гражданской авиации (ГА) практикуется составление планов использования воздушного пространства (ИВП). Процесс планирования состоит из ряда последовательных этапов: долгосрочный (составление сезонного расписания), стратегический (предварительный), предтактический (суточный) и тактический (текущий) [1]. Завершающий этап – обеспечение диспетчеров плановой информацией в процессе непосредственного управления воздушным движением (УВД). Общая задача планирования ИВП согласованно решается во взаимодействующих центрах Единой системы организации воздушного движения (ЕС ОрВД) различного уровня.

Цель долгосрочного планирования полетов состоит в удовлетворении сезонных потребностей народного хозяйства и населения в авиаперевозках. Учитываются статистика прежних лет и прогноз экономического развития, а также заявки на выполнение специальных заданий. Осуществляются сбор и обработка информации о предполагаемых объемах перевозок на полгода вперед, составление и координация расписания. Согласуются мероприятия, требующие специальной организации воздушного пространства (ВП). Разрабатывается долгосрочный план ИВП, согласованный по ряду критериев (например, по дням недели, удобству транзитных пассажиров и т.д.). Результат служит основой для проведения в последующем суточного планирования воздушного движения (ВД) и учета интересов ГА при распределении ВП. Применительно к России задача имеет большую размерность параметров [2]:

- более 3000 пунктов обязательных донесений (ПОД);
- более 5000 участков трасс, около 750 воздушных трасс;
- около 40000 возможных маршрутов полета;
- около двух тысяч в сутки заявок на полеты;
- до 800 изменений, вносимых за сутки в существующий план ИВП.

Напомним, что в период плановой экономики, когда в стране существовала единственная авиакомпания, задача составления расписания облегчалась отсутствием конкуренции в сфере воздушных перевозок. Согласованные с Госпланом заявки ведомств на услуги ГА поступали в центральную планово-диспетчерскую службу, специалисты составляли сбалансированное с потребностями населения сезонное расписание полетов, которое наделялось статусом закона и распределялось по авиаотрядам для неукоснительного исполнения.

С переходом к рыночным отношениям в России образовались сотни авиакомпаний (к настоящему времени их количество сократилось до десятков). Несовпадение финансовых интересов между поставщиками услуг воздушного транспорта породило известные процессы распределения сфер их влияния как в пространстве (аэродромы и маршруты), так и во времени (удобные для потребителей услуг времени вылета и посадки).

1. Прогнозирование спроса на рынке авиаперевозок

Авиакомпания, вступающая на рынок, должна реально представлять, на какую долю перевозок на данном рынке она может претендовать в данное время и каковы ее шансы на успех в конкурентной борьбе с другими авиакомпаниями и другими видами транспорта. Доля конкретной авиакомпании на рынке определяется количеством грузовых и пассажирских мест в парке ВС авиакомпании. Прогнозирование объемов перевозок основано на качественных и количественных методах принятия оптимальных решений [3]. Известные математические модели построены на анализе процентного изменения критерия к процентному изменению перевозок (эластичность показателя объема к различным факторам). Например, эластичность по тарифу имеет отрицательный коэффициент эластичности $-(0.5 \div 1.0)$, показывающий, что объем перевозок увеличивается при снижении тарифа. Для валового национального продукта коэффициент ценовой эластичности составляет $(1.5 \div 2.5)$ с большей эластичностью при неразвитом рынке. Логарифмическая форма модели, используемой для анализа перевозок во всем мире [4], следующая: $\ln Y = a + b \ln A + c \ln Z$, где Y – реализованные пассажиро-километры; A – валовой внутренний продукт в реальном выражении; Z – доход от пассажирских перевозок на пассажиро-километр; a, b, c – постоянные коэффициенты, значения которых получают методом статистической оценки.

Перед каждой авиакомпанией встает задача определения своей доли в общегосударственном (и в мировом) объеме перевозок. Необходимо собрать полную информацию о рынке. Рынок состоит из потребителей, имеющих схожие потребности. Однако совокупность покупателей никогда не бывает однородной. Они различаются по своим интересам, по денежной сумме, которую готовы заплатить перевозчику, по требуемому количеству рейсов и по источникам информации. Следовательно, авиакомпании целесообразно разделить рынок на сегменты, а затем выбрать один или несколько из них для освоения.

Единого способа сегментации рынка не существует. Считается полезным использовать разные сегментации в их сочетаниях. Различаются [5]: 1. Географическая сегментация (место жительства, тип населенного пункта, плотность населения, климат). 2. Демографическая сегментация (пол, возраст, количество членов семьи, уровень доходов, род занятий, образование, религиозные убеждения, национальность). 3. Психологическая сегментация (социальный класс, стиль жизни, тип личности). 4. Поведенческая сегментация.

Высокая эффективность сегментации достигается, если сегменты обладают свойствами измеримости (известен способ измерения), доступности (сегмент не интересуется конкурентов), доходности, возможности освоения (сегмент «по плечу» авиакомпании). Спрос конкретен в своих проявлениях, он привязан к месту и времени, услугам, группам населения, имеющимся тарифам, денежным доходам. Изучение спроса связано с определением емкости рынка услуги в целях определения ее объема для той или иной компании. Емкость рынка определяется объемом реализованных перевозок обычно в течение года (в физических единицах или стоимостном выражении), то есть в конечном итоге – объемом потребляемых услуг авиации. Для оценки привлекательности сегмента используются такие факторы, как: размер сегмента; возможности роста сегмента; прибыльность сегмента; конкуренция на данном сегменте, настоящая и потенциальная; возможности самой компании. На практике используется также сегментация по дальности полета, по культурным традициям страны проживания или рождения пассажира, по целям путешествий.

Отличия рынка грузовых перевозок от пассажирского рынка: 1. Груз всегда доставляется в одну сторону, то есть спрос по направлениям несбалансирован. На незагруженных направлениях для привлечения загрузки снижаются тарифы. 2. Груз различен по размерам, весу, плотности, требованиям к условиям хранения и погрузки. 3. На грузовом рынке (при любых расстояниях) высока конкуренция с наземными видами транспорта.

2. Постановка и формализация задачи анализа сезонного расписания

В условиях рыночной конкуренции централизованное планирование полетов в воздушном пространстве России фактически свелось к координации поступающих заявок на ИВП. Заявки на сезонное расписание поступают из аэропортов. Отношения между владельцами аэродромов и владельцами парка ВС определяются платежеспособностью авиакомпаний, предпочтения при предоставлении услуг технического обслуживания и по запрошенному времени выполнения аэродромных операций вылета и посадки диктуются в первую очередь финансовой привлекательностью эксплуатанта. Каждой аэродромной операции выделяется так называемый *slot*, временное окно, т.е. интервал около устанавливаемого владельцем аэропорта момента вылета или посадки (минус 5 плюс 10 минут). Значение слова *slot* (щель) соответствует смыслу термина, однако для большей строгости ему сначала присвоили расшифровку *Selected Local Operational Time* (выделенное для операции местное время), а затем переопределили в *Selected Take-Off Time* (STOT) – выделенное для взлета время.

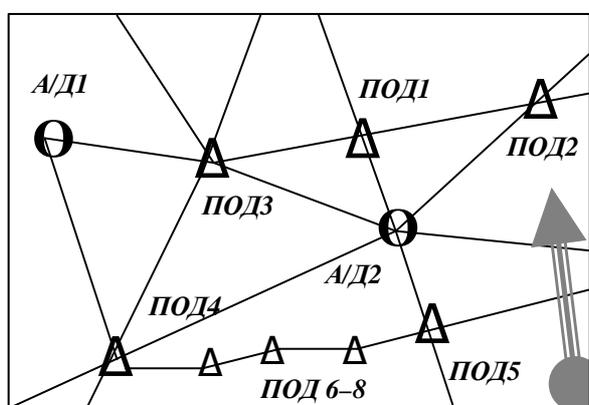
Аэропорт распределяет слоты между авиакомпаниями, исходя из собственных интересов, однако эти интересы нужно согласовать с взаимодействующими органами государственного регулирования ИВП. Если рейс по сезонному расписанию затрагивает ВП только своей аэродромной зоны, то задача координации с другими участниками движения снимается. В противном случае, если рейс затрагивает ВП района УВД, на территории которого расположен аэродром, условия передачи рейса от органов регулирования ВД аэродрома в районный центр (РЦ) должны согласовываться с районом на весь период действия сезонного расписания. Аналогично, план полета рейса, затрагивающего ВП нескольких РЦ, должен согласовываться с зональным центром (ЗЦ), а рейсы, затрагивающие ВП нескольких ЗЦ, должны координироваться с Главным центром (ГЦ) ЕС ОрВД.

Задача каждого аэропорта как поставщика услуг авиакомпаниям состоит в максимальном удовлетворении всех поступивших заявок на ИВП. В случаях столкновения интересов различных потребителей, при запросе одного и того же слота разными компаниями, предпочтение отдается более состоятельному клиенту. Однако даже если удастся безболезненно разрешить конфликты и согласовать слоты на аэродроме, это еще не значит, что составленное расписание окажется сбалансированным вне аэродромной зоны, т.е. в трассовом ВП. Каждый рейс должен без нарушений федеральных авиационных правил вписаться в формируемые потоки ВД на всей территории страны, и вопросы безопасности требуют согласовывать последствия назначения каждого слота на всю глубину полета.

Сезонное расписание содержит повторяющиеся каждую неделю планы рейсов. Наряду с ежедневными регулярными полетами в нем представлены рейсы, действующие в отдельные дни недели. Фактически задача составления оптимального сбалансированного плана ставится заново ежедневно, так как каждый день в течение всего сезона в движении участвуют разные рейсы. Оптимальное по экономическим и по любым другим критериям решение становится уникальным для каждого дня действия расписания. Однако во всем мировом авиационном сообществе действует жесткое ограничение на так называемую регулярность полетов ГА. Назначенное рейсу время вылета должно соблюдаться независимо от дня недели.

Попытки математического обоснования эмпирических приемов и процедур обеспечения равномерной загрузки воздушного пространства предпринимались [6] с помощью методов теории календарного планирования (расписаний), теории очередей (массового обслуживания), линейного и нелинейного программирования (например, потоки в сетях). Результаты, как правило, нивелировались большой размерностью задачи и высокой чувствительностью методов к неизбежным флуктуациям значений параметров. Теория очередей, достаточно популярная в сфере оценки вероятностных характеристик потоков движения, вообще говоря, является инструментом анализа случайного процесса, но не его синтеза и тем более оптимизации. В [7] предложен подход к решению задачи с помощью имитационной модели.

Пусть имеем представленный на рис. 1 связный планарный гиперграф $G(v_i, r_{ij})$, отображающий содержимое Каталога зарегистрированных маршрутов России. Аббревиатуры ПОД (пункт обязательных донесений) и А/Д (аэродром) общеупотребительны. Вершины графа v_i ($i = 1, \dots, I$) сопоставлены I навигационным пунктам, точкам пересечения трасс и аэродромам, а соединяющие вершины v_i и v_j ребра r_{ij} ($i, j = 1, \dots, I; i \neq j$) – участкам трасс, включенных в [8]. Пусть каждое ребро r_{ij} взвешено показателями g_{ij} пропускной способности и c_{ij} стоимости движения на участке r_{ij} . Оба весовых коэффициента переменные, их значения поддерживаются по результатам обновления данных о техническом, метеорологическом и организационном (режимном) состоянии соответствующих отрезков воздушных трасс. Отказы средств наблюдения и связи, ограничения полетов, опасные явления погоды снижают пропускную способность g_{ij} трассового пространства r_{ij} . Встречные ветровые потоки повышают, а попутные – уменьшают расход топлива и, как следствие, коэффициент c_{ij} стоимости полета. Значения весовых показателей, вообще говоря, в каждой точке отрезка трассы являются функциями высоты и расстояния.



1	2	3	4	5	A1	A2	
	c_{12}	c_{13}	∞	∞	∞	c_{1a}	1
		∞	∞	∞	∞	c_{22}	2
Элементы			c_{34}	∞	c_{31}	c_{32}	3
матрицы равны				c_{45}	c_{41}	c_{42}	4
∞ , если вершины не					∞	c_{52}	5
связаны друг с другом						∞	A1
							A2

Рис. 1. Машинное отображение графа $G(v_i, r_{ij})$ маршрутов на матрицу M . Для сокращения размерности в матрицу не включены вершины степени 2 (ПОД 6 – 8). Они отображаются ребром 4 – 5 (r_{45}) с суммарной стоимостью c_{45} и минимальной пропускной способностью g_{ij} среди смежных участков 4-6, 6-7, 7-8, 8-5. В правом нижнем углу стрелкой указано направление ветра. Для маршрута А/Д1-ПОД3-ПОД1-А/Д2 ветер на заключительном участке маршрута встречный, для А/Д1-ПОД4-...-ПОД5-А/Д2 – попутный, т.е. $c_{52} < c_{1a}$

Среди десятков тысяч возможных путей достижения любой вершины графа из любой другой его вершины априорно заданы в качестве ограничений единицы допускаемых [5] маршрутов, ранжированные по предпочтительности при следовании из одного отправного пункта в другой. По протяженности полета приоритетность остается постоянной (с точностью до ввода новых аэролиний), по технической оснащенности трасс изменяется с годами, по воздействию сезонных ветровых потоков отслеживается корректировка расписания, по внезапным отказам средств обеспечения полетов, явлениям погоды, вводам ограничений зависит от интенсивности этих событий. Машинным представлением совокупности данных о среде, в которой моделируется процесс ИВП, становится отображение графа $G(v_i, r_{ij})$ на матрицу M , каждый элемент которой m_{ij} сопоставлен набору обновляемых значений g_{ij} и c_{ij} .

Должен быть разработан метод составления бесконфликтного плана ИВП в процессе его формирования, позволяющий удовлетворить максимально возможное количество заявок с учетом государственных приоритетов при выполнении ограничений на показатели безопасности, экономичности и регулярности полетов, а также на потребляемые компьютерные ресурсы. Метод должен использовать целенаправленную пошаговую процедуру, позволяющую находить рациональный (экономически допустимый) план полетов по сезонному расписанию на любой день его действия.

3. Модель использования воздушного пространства

Модель относится к классу имитационных. Структурно она реализована как композиция трех фильтров (маршрутов, загрузки, конфликтов) и эксперта рекомендаций, формулирующего предложения по оптимизации плана ИВП в процессе его составления. Над структурной составляющей надстроены алгоритмические процедуры проверки логических условий. Входной информацией является поток поступающих в систему заявок на выполнение полетов по расписанию. В качестве результата либо используется диалог с диспетчером организации потоков (рекомендации), либо заявка включается в формируемое расписание. Обработка осуществляется последовательно, по мере поступления заявок аэродромов вылета.

Фильтр маршрутов производит проверку поступающей заявки на соответствие условиям использования воздушного пространства. Его информационной базой являются:

- правила составления заявок на использование воздушного пространства;
- описание структуры воздушного пространства России;
- летно-технические характеристики воздушных судов;
- прогноз метеорологической обстановки на аэродромах.

Алгоритмическая схема фильтрации включает в себя:

- форматно-логический контроль заявки на ИВП;
- анализ возможности сформировать маршрут от точки входа в воздушное пространство России (или взлета на ее территории) до точки выхода за границу (посадки);
- штурманский расчет заявки на ИВП и распределение информации по элементам структуры воздушного пространства (построение модели ИВП);

Нарушение любого из перечисленных условий приводит к формированию и выдаче на отображение диспетчеру диагностического сообщения, облегчающего редактирование заявки. Простейший пример – обнаружение ошибок формата полей плана. Корректно составленные заявки на ИВП передаются следующему фильтру. В информационном поле модели фиксируются отметки о пролете каждого из секторов и пунктов, затрагиваемых маршрутом.

Фильтр загрузки производит проверку поступающей заявки на соответствие пропускной способности элементов ВП: секторов и участков трасс по маршруту. Его основу составляет отображение вершин и ребер графа $G(v_i, r_{ij})$ на информационное поле базы данных. Каждому навигационному пункту, аэродрому и каждому сектору управления воздушным движением на территории России сопоставлены функции распределения полетов через эти элементы ВП, представленные гистограммами почасовой загрузки. Совокупность создаваемых по мере формирования расписания функций распределения загрузки, называемая далее его информационным образом, позволяет в реальном масштабе времени решать задачу определения загрузки элементов системы. Почасовая загрузка каждого затрагиваемого маршрутом сектора и участка трассы отображена высотой столбца гистограммы, соответствующего расчетному времени движения в границах этих элементов пространства. Превышение столбца над допустимым значением (порогом) сигнализирует о необходимости пересмотра расписания. По каждой заявке на ИВП, для которой зафиксировано превышение порога, формируется запрос на перераспределение потоков алгоритмами эксперта рекомендаций.

Фильтр конфликтов действует по трехступенчатой схеме. Сначала анализируется совместное распределение ранее включенных в расписание заявок с очередной обрабатываемой заявкой в затрагиваемых вводимым маршрутом навигационных пунктах и в точках пересечения трасс. Алгоритм последовательно обращается к гистограммам распределения загрузки указанных пунктов, отслеживая соблюдение норм эшелонирования и адресуюсь к соответствующим по времени их пролета столбцам. Две последующие ступени анализируют предпосылки к потенциальным конфликтным ситуациям.

Эксперт рекомендаций анализирует альтернативные варианты по критерию равномерности

загрузки с учетом государственных приоритетов рейсов и предлагает их персоналу для утверждения. Информационную основу алгоритмов составляет упомянутая ранее треугольная матрица без диагонали M графа $G(v_i, r_{ij})$, каждый элемент m_{ij} которой представляется парой значений g_{ij} и c_{ij} . Сопровождение обоих коэффициентов основано на прогнозе метеорологической обстановки на период планирования, действующих в его пределах ограничений полетов и данных о работоспособности технических средств. Для сокращения размерности матрицы M используется подграф $G_c(v_i, r_{ij})$ исходного графа $G(v_i, r_{ij})$, из которого исключены вершины степени два, соответствующие навигационным пунктам, не являющимся узлами пересечения или разветвления воздушных трасс России.

4. Требования к методу компьютерной поддержки процедур составления сезонного расписания

1. Метод должен строиться как унифицированный инструмент формирования плана ИВП на этапах составления сезонного расписания, суточного и текущего планирования на территории страны, зоны, района в целях сохранения преемственности решений о распределении полетов по элементам ВП.

2. Программная реализация метода должна требовать для своей работы не более 10% компьютерных ресурсов производительности и памяти.

3. Должна обеспечиваться работа персонала в реальном масштабе времени (оптимизация времени анализа сезонного расписания на соответствие нормативам пропускной способности – НПС).

4. Информационное обеспечение метода должно формироваться как компьютерная модель полетных данных, способная компактно представлять рельеф загрузки элементов ВП с учетом действующих в системе на период планирования связей и ограничений.

5. Модель полетных данных, создаваемая для сезонного расписания, должна формироваться на каждый сутки периода действия расписания по мере поступления повторяющихся планов и сохраняться в таком виде в памяти для анализа влияния каждой новой вводимой заявки на сбалансированность и непротиворечивость сводного плана ИВП на каждые сутки.

6. На этапе составления сезонного расписания при построении модели ИВП должна учитываться статистика метеорологической обстановки по маршруту каждого рейса по итогам многолетних наблюдений погоды.

7. На предварительном этапе модель ИВП должна преобразовываться с учетом поступления в систему заявок на рейсы вне расписания, а также в соответствии с изменениями планов рейсов по расписанию и с прогнозом состояния системы, в том числе технических средств, режимных ограничений, атмосферных явлений, т.е. адаптироваться к новым условиям. Адаптивная модель должна создаваться до начала полетов и приспосабливаться к ситуации по мере их выполнения.

8. На этапе краткосрочного планирования и УВД программная реализация метода должна включаться автоматически как реакция на поступление в ПО информации об изменении условий выполнения полетов. При этом должны обнаруживаться все рейсы, попавшие (а также планирующие оказаться) в неблагоприятных условиях, для немедленной выработки рекомендаций персоналу по регулированию потока ВС в создавшейся ситуации.

9. Процедура регулирования должна реализовываться последовательными шагами; сначала все планы рейсов, затронутые неблагоприятными условиями, удаляются из ПО, затем один за другим вводятся в систему вновь в порядке их приоритетности, подвергаются новому штурманскому расчету и распределению по элементам ВП.

10. Присоединение планов регулируемых рейсов к действующему текущему сводному плану ИВП должно осуществляться последовательно в режиме диалога с ответственным должностным лицом системы. По каждому изменяемому рейсу на отображение должны выдаваться рекомендации по переводу на обходные маршруты.

11. На этапе предварительного планирования экономичные маршруты полетов должны выбираться из Каталога зарегистрированных маршрутов РФ, если это не противоречит цели полета. На этапе оперативного регулирования, когда стоимости пролета участков трасс резко изменяются, должны включаться процедуры нахождения на существующей сети трасс кратчайшего маршрута между текущим местоположением ВС и аэродромом назначения.

12. Выбор и утверждение одного из предложенных вариантов осуществляет уполномоченное ответственное лицо, которому должна предоставляться возможность не только выбирать, но и диктовать системе альтернативные решения, а также откладывать регулирование рейсов, по которым необходимо согласование с органами УВД по дальнейшему маршруту.

13. Метод должен обнаруживать перегрузку любого элемента ВП, затрагиваемого корректируемым рейсом, если она возникает, и находить альтернативные варианты для выработки рекомендаций диспетчерскому персоналу.

14. В целях обеспечения требований по безопасности должно проверяться соблюдение безопасных интервалов следования и безопасных расстояний между ВС при пролете ПОД, а также в ситуациях обгона и движения на пересекающихся курсах. Должна использоваться процедура обнаружения опасных сближений ВС между собой и с зонами ограничений.

15. Для анализа сезонного расписания на соответствие нормативам пропускной способности прогнозируемое на время выполнения полета значение норматива должно вычисляться программно по действующей методике оценки НПС для каждого сектора, затрагиваемого каждым рейсом.

16. Анализ сезонного расписания на соответствие нормативам пропускной способности должно производиться программно для каждого очередного повторяющегося плана полета при его вводе в систему (т.е. при его присоединении к формируемым рельефам загрузки ВП для каждого дня действия вводимого плана).

Заключение

Переход от оптимизационной к имитационной модели позволяет избежать экспоненциального роста сложности анализа сезонного расписания на соответствие нормативам пропускной способности элементов воздушного пространства. В изложенной постановке ее реализация требует достаточно умеренных для аппаратуры серверного класса компьютерных ресурсов. В оперативной памяти должны храниться описания около трехсот секторов управления, шестнадцать тысяч навигационных пунктов и аэродромов, сотен трасс и сорока тысяч построенных на них зарегистрированных маршрутов. Самый объемный элемент модели – информационный образ использования воздушного пространства – занимает около шестидесяти четырех мегабайтов на каждый день действия анализируемого расписания. Основными препятствиями применению предлагаемой адаптационной технологии планирования и регулирования потоков выступают не столько технические, сколько организационные проблемы.

Рассмотренный подход реализован в программном обеспечении комплекса средств автоматизации для планирования использования воздушного пространства, разрабатываемого в рамках модернизации Единой системы организации воздушного движения в России [9].

ЛИТЕРАТУРА

1. Федеральные авиационные правила. Подготовка и выполнение полетов в гражданской авиации Российской Федерации. Утверждены приказом Минтранса России от 31.07.2009 № 128. - М., 2009.
2. Автоматизированные системы управления воздушным движением: Новые информационные технологии в авиации: учебн. пособие / Р.М. Ахмедов, А.А. Бибутов, А.В. Васильев и др. / под ред. С.Г. Пятко и А.И. Красова. - СПб.: Политехника, 2004.
3. SESAR Consortium. SESAR Definition Phase: the Concept of Operation at a glance. EUROCONTROL. 2007.

4. Егорова Н.Е., Мудунов А.С. Применение моделей и методов прогнозирования спроса на продукцию сферы услуг. - М.: ЦЭМИ РАН, 2000.
5. Бронникова Т.С., Чернявский А.Г. Сегментирование рынка. - Таганрог: ТРТУ, 2008.
6. Урубков А.Р. Курс MBA по оптимизации управленческих решений: Практическое руководство по использованию моделей линейного программирования. - М.: Альпина Бизнес-Букс, 2006.
7. Гальков М.А., Рудельсон Л.Е., Тверитнев М.М. Имитационная модель использования воздушного пространства // Известия Российской академии наук, Теория и системы управления. - 2003. - № 4.
8. Каталог зарегистрированных маршрутов. - М.: ГЦ ППВД, 2003.
9. Автоматизированная система планирования использования воздушного пространства и организации потоков воздушного движения для Главного центра ЕС ОрВД. Технический проект ПАВУ.466453.001ПЗ. - М., 2010.

THE ANALYSIS OF THE SEASONAL SCHEDULE ON CONFORMITY TO THROUGHPUT SPECIFICATIONS

Bykova V.V., Moiseev O.N., Rudelson L.E., Chernikova M.A.

Optimizing models of the organization of air traffic flows for construction of the frictionless seasonal schedule are insufficiently effective owing to inevitable fluctuations of variables in the course of its realization. Search of a steady local optimum of function (a maximum of the satisfied demands) is preferable at restrictions on indicators of safety, a regularity and profitability of flights.

Key words: the seasonal schedule, use of air space, imitation model.

Сведения об авторах

Быкова Вера Викторовна, окончила МГТУ ГА (1989), научный сотрудник ГосНИИ ГА, автор около десяти научных работ, область научных интересов – программное обеспечение автоматизированных систем организации воздушного движения.

Моисеев Олег Николаевич, 1989 г.р., окончил МГТУ ГА (2010), аспирант кафедры ВМКСС МГТУ ГА, область научных интересов – вычислительные системы и методы, распределенные системы, организация вычислений.

Рудельсон Лев Ефимович, 1944 г.р., окончил МЭИ (1968), доктор технических наук, профессор МГТУ ГА, автор более 140 печатных работ, область научных интересов – программное обеспечение автоматизированных систем организации воздушного движения.

Черникова Марина Александровна, окончила МГТУ ГА (2004), аспирантка кафедры ВМКСС МГТУ ГА, автор 20 научных работ, область научных интересов – программное обеспечение планирования полетов воздушных судов.

УДК 519.81

РЕШЕНИЕ ЗАДАЧИ РАССТАНОВКИ ПАРКА ВОЗДУШНЫХ СУДОВ ПО РЕЙСАМ ПОЛЕТНОГО РАСПИСАНИЯ С ПОМОЩЬЮ ЦЕПОЧЕК РЕЙСОВ

Г.М. ФРИДМАН, Н.А. ЗЕНКОВА

Представлена математическая модель, описывающая процесс назначения парка воздушных судов на рейсы полетного расписания авиакомпании, основанная на использовании цепочек рейсов, проведен анализ модели, а также даны результаты расчетов для модельного и полномасштабного набора данных.

Ключевые слова: цепочки рейсов, оптимальная расстановка, целочисленное программирование.

1. Введение

Выделяют следующие этапы планирования пассажирских авиаперевозок [1, 2]:

- планирование сети маршрутов и состава парка воздушных судов (ВС): определение перспективных, стратегически важных рынков; определение необходимой доли присутствия на рынке (по емкостям, рейсам); определение наиболее привлекательного для пассажира времени вылета и прилета;
- расстановка парка ВС по рейсам: оптимальное распределение самолетов различных типов по рейсам в соответствии со спросом и техническими характеристиками топливной эффективности на данном расстоянии;
- составление цепочек рейсов и определение оптимальной последовательности полетов для каждого самолета;
- составление оптимального расписания для экипажей ВС с учетом нормативов по безопасности полетов.

До недавнего времени все четыре задачи сначала решались для стандартного дня недели, а затем расписание изменялось вручную, чтобы точнее смоделировать специфику отдельных дней недели. Особенно больших изменений требовало составление расписания на выходные. Сегодня наиболее развитые из предлагаемых на рынке решений позволяют рассматривать задачу составления расписания не только на целую неделю, с учетом особенностей каждого дня, но и на весь сезон. Естественным с точки зрения математического моделирования было бы проводить все этапы планирования совместно [1, 3], однако это в настоящее время оказывается невозможным вследствие огромных размеров возникающей оптимизационной задачи. Поэтому обычно используется итерационная процедура, часто, к сожалению, требующая возвращения на предыдущий этап из-за отсутствия допустимых (даже не оптимальных) решений на следующем.

2. Задача расстановки парка воздушных судов на базе «цепочек» рейсов

В статье рассмотрена одна из ключевых задач планирования, а именно задача оптимальной расстановки парка ВС по рейсам полетного расписания с помощью цепочек рейсов; традиционно эта задача носит название Fleet Assignment Model (FAM). Под термином «рейс» ниже понимается беспосадочный перелет из аэропорта А в аэропорт Б с заранее присвоенным уникальным номером и заданным временем вылета и прилета.

«Цепочка рейсов», на базе которой в статье решается задача FAM – это последовательность стыкующихся между собой рейсов. Рейсы в цепочке стыкуются с учетом выполнения всех норм технического обслуживания и обратного времени и выполняются на одном самолете. Естественным критерием стыковки рейсов в цепочке служит совпадение аэропорта прилета предыдущего рейса с аэропортом вылета следующего.

Целью решения задачи FAM является максимизация суммарной прибыли от выполнения всех запланированных рейсов при использовании ограниченных ресурсов, доступных авиакомпании (собственного парка самолетов). Входными данными служат расписание полетов, запланированное к выполнению, прогноз общего спроса на каждый из рейсов, оборотные времена (минимальное время, необходимое данному типу ВС в данном аэропорту для проведения всех технологических процедур, совершаемых между посадкой и последующим за ней вылетом), а также доходы и расходы на выполнение каждого рейса различными типами ВС. Предполагается, что расписание составлено на некий период времени (например, одни сутки) и циклически повторяется.

В практической деятельности авиакомпаний наиболее распространен эвристический метод составления цепочек, основанный на использовании ближайшей доступной стыковки для получения начального решения и дальнейшем итерационном его улучшении. В статье использован строгий математический алгоритм построения цепочек с помощью алгоритма поиска в глубину в ориентированном графе (орграфе) [4]. Пусть мы начинаем поиск с некоторой фиксированной вершины v_0 , затем выбираем произвольную вершину u , смежную с v_0 , и повторяем наш процесс от u . В общем случае предположим, что мы находимся в вершине v . Если существует новая (еще непросмотренная) вершина u , $u - v$, мы рассматриваем эту вершину (она перестает быть новой), и начиная с нее продолжаем поиск. Если же не существует ни одной новой вершины, смежной с v , то мы говорим, что вершина v использована, возвращаемся в вершину, из которой мы попали в v и продолжаем процесс (поиск закончен, если $v = v_0$). Другими словами, поиск в глубину из вершины v основывается на поиске в глубину из всех новых вершин, смежных с v .

Итак, необходимо по заданному однодневному расписанию рейсов для каждого из типов самолетов составить все возможные однодневные цепочки рейсов. Однодневной будем считать цепочку «первый и последний рейс», которой начинается промежуток с 00:00 до 23:59. На подготовительном этапе полетное расписание клонируется на два дня. Далее, полученное двухдневное расписание представляется в виде орграфа, где вершины соответствуют рейсам. Между двумя вершинами есть ребро, если один самолет может последовательно выполнить два рейса (аэропорт прибытия одного рейса совпадает с аэропортом отправления второго и т.д.). Используя заранее оговоренные критерии стыковки рейсов, составляется матрица смежности этого графа, и определяются все пути от всех вершин, соответствующих рейсам первого дня, до какой-либо вершины, соответствующей рейсу из второго дня. Отбрасывая вершины, соответствующие рейсам второго дня (т.е. последнюю вершину каждой цепочки), в результате находим все однодневные цепочки рейсов. Итак, получено множество всех возможных цепочек $s \in S$, состоящих из стыкующихся между собой рейсов $l \in L$ заданного полетного расписания L .

В то время как в классической формулировке задача FAM [5] решается «на уровне рейсов», а затем из найденной оптимальной расстановки воздушного флота по рейсам определяется последовательность выполнения рейсов для каждого ВС данного типа (т.е. составляются цепочки), в альтернативном подходе изначально ставится вопрос о назначении того или иного типа ВС сразу на всю цепочку. Ниже приведена общая математическая формулировка соответствующей задачи.

Множества:

A : множество аэропортов с индексом $a \in A$.

L : множество рейсов с индексом $i \in L$.

F : множество типов судов с индексом $f \in F$.

T : множество всех событий вылета и прилета с индексом $t \in T$.

S : множество всех цепочек рейсов с индексом $s \in S$.

$IS(f, a)$: множество цепочек рейсов, оканчивающихся в аэропорту $a \in A$, на которые назначен самолет типа $f \in F$.

$OS(f, a)$: множество цепочек рейсов, начинающихся в аэропорту $a \in A$, на которые назначен самолет типа $f \in F$.

Неизвестные:

$$x_{f,s} = \begin{cases} 1, & \text{если самолет типа } f \in F \text{ назначен на цепочку } s \in S \\ 0, & \text{если самолет типа } f \in F \text{ не назначен на цепочку } s \in S \end{cases}$$

Каждой цепочке $s \in S$, таким образом, соответствует столько неизвестных $x_{f,s}$, сколько типов самолета $f \in F$ может быть назначено для выполнения всех рейсов, входящих в данную цепочку. Проверка возможности назначения самолета типа f на цепочку s должна быть проведена непосредственно после генерирования множества всех цепочек S . Это позволит сократить общее количество переменных в оптимизационной задаче в случае, если на какие-либо цепочки невозможно назначить ВС определенного типа/типов.

Параметры:

N_f : имеющееся количество ВС типа $f \in F$.

Cap_f : вместимость (емкость) ВС типа $f \in F$.

$C_{f,s}$: операционные расходы на присвоение судна типа $f \in F$ цепочке $s \in S$.

$R_{f,s}$: доход, получаемый от назначения судна типа $f \in F$ на цепочку $s \in S$.

Функция цели:

$$\sum_{s \in S} \sum_{f \in F} (R_{f,s} - C_{f,s}) \cdot x_{f,s} \rightarrow \max. \quad (1)$$

Функция цели – прибыль авиакомпании от назначения самолетов определенного типа на определенную последовательность рейсов (цепочку), т.е. доход от проданных билетов минус затраты на перевозку и обслуживание самолетов. Ожидаемые доходы и расходы для цепочки рассчитываются как сумма соответствующих величин для всех входящих в нее рейсов. Доход отдельного рейса $i \in L$ определяется, исходя из средней стоимости билета на него, вместимости ВС и прогноза спроса по формуле

$$R_{f,i} = \min(Cap_f, D_i) \times p_i, \quad (2)$$

где D_i – прогноз общего количества пассажиров, желающих полететь данным рейсом (прогноз спроса); p_i – средняя цена билета на данный рейс. Цена билета умножается на минимум между вместимостью самолета и прогнозом спроса, что, естественно, определяет ожидаемое количество проданных на рейс билетов. Величина расхода для рейса $i \in L$ включает общие операционные затраты, связанные с назначением на рейс $i \in L$ судна $f \in F$. Они состоят из расходов на топливо, затрат на экипаж, обслуживание самолета и пассажиров, расходные материалы, амортизацию и т.п.

Ограничения:

Условие покрытия – на одну цепочку может быть назначено ВС не более, чем одного типа

$$\sum_{f \in F} x_{f,s} \leq 1, \quad \forall s \in S. \quad (3)$$

Условие разбиения – в оптимальной совокупности цепочек каждый рейс, может присутствовать в одной и только одной цепочке

$$\sum_{f \in F} \sum_{s \in S} \sigma_{i,s} x_{f,s} = 1 \quad \forall i \in L \quad (4)$$

где

$$\sigma_{i,s} = \begin{cases} 1, & \text{если цепочка } s \text{ включает рейс } i \in L \\ 0, & \text{если цепочка } s \text{ не включает рейс } i \in L \end{cases}$$

Условия мощности парка – количество цепочек, на которые назначен самолет данного типа, не должно превосходить количества имеющихся самолетов данного типа

$$\sum_{s \in S} x_{f,s} \leq N_f \quad \forall f \in F. \quad (5)$$

Условие баланса – для любого типа самолета $f \in F$ и любого аэропорта $a \in A$ количество цепочек, начинающихся в данном аэропорту, должно быть равно количеству цепочек, заканчивающихся в нем, при условии, что на эти цепочки назначен самолет данного типа

$$\sum_{s \in IS(f,a)} x_{f,s} = \sum_{s \in OS(f,a)} x_{f,s} \quad \forall f \in F; \forall a \in A. \quad (6)$$

Исходные данные:

- полетное расписание (A, L, T) ;
- имеющийся в наличии авиационный парк (F, N_f, Cap_f) ;
- удельные затраты $C_{f,i}$ на полеты по рейсу $i \in L$ самолета типа $f \in F$;
- доход $R_{f,i}$, полученный от назначения самолета $f \in F$ на рейс $i \in L$.

Задача (1) – (6) является задачей целочисленного программирования и может быть решена, например, методом ветвей и границ [6]. Результатом решения задачи является оптимальный, с точки зрения общей прибыли, набор цепочек рейсов, на каждую из которых назначен самолет определенного типа.

3. Сравнительный анализ предложенной формулировки задачи и возможности ее расширения

Цепочки рейсов довольно успешно используются в задачах планирования расписаний, поскольку предоставляют возможности по их интеграции, прежде всего, по совместному решению задач расстановки типов ВС по рейсам и составлению бортовых полетных планов и связей для экипажей [7]. Применение цепочек также оказалось эффективным при построении расписаний, устойчивых к воздействию случайных факторов, таких как погодные условия, поломка самолета или болезнь экипажа, забастовки персонала и т. д. [8].

В математических формулировках задач, рассмотренных в работах, связанных с использованием цепочек рейсов, в том числе и в постановке (1) – (6) данной статьи, каждая цепочка порождает столько бинарных неизвестных, сколько типов ВС может быть на нее назначено. При этом, однако, задача (1) – (6) выгодно отличается от уже существующих тем, что не содержит никаких других неизвестных, которые отвечают за стоянку ВС в аэропортах полетного расписания и т.п. [5, 7, 8].

В предлагаемой постановке условие покрытия (3) предполагает, что на цепочку назначается не более одного типа ВС, что в сочетании с условием (4) обеспечивает назначение в точности одного самолета на каждый рейс полетного расписания. Условия (5) и (6) гарантируют, что в каждый момент времени используется не более, чем имеющееся количество N_f ВС каждого типа $f \in F$. Кроме того, условие (6) обеспечивает сбалансированность всего расписания. Отсутствие дополнительных неизвестных делает задачу (1) – (6) более «экономной» с вычислительной точки зрения, что чрезвычайно важно при проведении полномасштабных расчетов. Это, как уже было упомянуто, связано с тем, что в задаче, решаемой на основе цепочек, существенно увеличивается, по сравнению с «рейсовой» задачей расстановки типов ВС [5], количество неизвестных, поскольку число всех возможных цепочек много больше числа исходных рейсов расписания. Например, в проведенном в данной статье расчете полномасштабной задачи расстановки типов ВС с однодневным полетным расписанием, состоящим из 84 рейсов, было построено более 30000 цепочек, что, в свою очередь, привело к задаче целочисленного программирования, содержащей около 230000 неизвестных (число различных типов ВС, которые использовала авиакомпания, было равно 7).

Математическая формулировка задачи FАM на основе цепочек (1) – (6) предоставляет дополнительные широкие возможности по установлению ограничений при составлении цепочек, позволяющие решать значительное количество технических и экономических задач. Примером может служить *условие цикличности*, т.е. условие о том, что цепочка должна начинаться и заканчиваться в одном и том же пункте; это может быть базовый аэропорт или аэропорт с более выгодными коммерческими условиями длительного технического обслуживания. Широко распространен этот подход в случае, если необходимо более гибко подходить к определению стыковочного времени, или реагировать на возможности аэропортов по приему определенных типов ВС.

Кроме указанных условий, в математическую постановку задачи (1) – (6) можно добавлять различные дополнительные ограничения. Например, для авиакомпании принципиально важно, какое количество часов тот или иной тип ВС проводит в воздухе. Этот параметр, так называемый «налет», фигурирует в лизинговых контрактах. За счет управления этим параметром для разных типов ВС при формировании условий оптимизационной задачи появляется возможность более приоритетного использования одного из типов ВС.

Большой интерес представляет собой возможность включения требований по выполнению необходимых форм технического обслуживания в задачу расстановки типов ВС. Учет подобных требований имеет смысл проводить еще при генерировании множества допустимых цепочек рейсов, т.к. это значительно снижает их общее количество [7].

Авиакомпании сами выбирают наиболее удобный для них способ решения задачи расстановки типов ВС, в зависимости от собственных возможностей, приоритетов и технических нужд, при этом в основном упор делается на известные алгоритмы FIFO/LIFO. В настоящее время ни одна российская авиакомпания не имеет в своем распоряжении работающей автоматизированной системы для поиска *оптимальной* расстановки типов ВС, поэтому результаты работы в данном направлении будут приносить ощутимое увеличение как дохода авиакомпании, так и эффективности принимаемых управленческих решений.

4. Модельный пример

В качестве расчетного примера было использовано полетное расписание условной авиакомпании, осуществляющей рейсы между четырьмя аэропортами: CDG (Charles De Gaulle, Париж, Франция), LED (Пулково, С.Петербург, Россия), OVB (Толмачево, Новосибирск, Россия), SVO (Шереметьево, Москва, Россия). Всего в расписании содержится 22 рейса; парк ВС состоит из A320 (Airbus A320-210) – 4 борта; B735 (Boeing 737-500) – 2 борта; B772 (Boeing 777-200) – 3 борта. Все исходные данные для рейсов полетного расписания приведены в табл. 1. Указан идентификационный номер каждого рейса, аэропорты его вылета и прилета, времена вылета и прилета в формате GMT+0 (т.е. время «по Гринвичу»), а также прогноз общего пассажирского спроса на рейс.

С точки зрения покупателя спрос на рейс есть величина совершенно бессмысленная, однако для проведения расчетов прогноз спроса на рейс определялся как сумма прогнозов спроса на все маршруты, в которые данный рейс входит. Под маршрутом здесь понимается выбранная пассажиром комбинация одного либо нескольких рейсов для совершения перелета из выбранного им начального аэропорта в конечный. Зная прогнозы спроса на все указанные маршруты, несложно вычислить прогноз спроса и на данный рейс. Значения прогноза спроса на маршруты, а, значит, и на рейсы часто оказываются не целыми, а вещественными числами, так как они основаны на статистической обработке исторических данных для общего числа проданных билетов.

Табл. 2 содержит информацию по вместимости типов ВС и удельным затратам по перевозке одного пассажира на 1 км пути самолетом данного типа. Обратные времена каждого типа ВС в каждом аэропорту полетного расписания собраны в табл. 3.

Таблица 1

Рейсы полетного расписания и прогноз спроса

Номер рейса	Аэропорт вылета	Аэропорт прибытия	Время вылета	Время прибытия	Прогноз спроса
1	LED	CDG	08:00	11:00	144.653
2	LED	CDG	16:00	19:00	132.975
3	LED	OVB	13:00	17:00	48.9623
4	LED	OVB	20:00	00:00	48.9623
5	LED	SVO	07:00	08:00	288.094
6	LED	SVO	15:00	16:00	369.597
11	CDG	LED	08:00	11:00	174.
12	CDG	LED	15:00	18:00	165.
13	CDG	SVO	07:00	11:00	135.
14	CDG	SVO	13:00	17:00	133.5
15	CDG	OVB	12:00	18:00	22.5
21	OVB	LED	02:00	06:00	78.4751
22	OVB	LED	10:00	14:00	89.0911
23	OVB	SVO	01:00	05:00	89.2936
24	OVB	SVO	11:00	15:00	53.6776
25	OVB	CDG	04:00	10:00	19.4626
31	SVO	LED	06:00	07:00	281.678
32	SVO	LED	13:00	14:00	245.
33	SVO	OVB	12:00	16:00	100.094
34	SVO	OVB	19:00	23:00	109.481
35	SVO	CDG	08:00	12:00	130.616
36	SVO	CDG	07:00	21:00	202.294

Таблица 2

Исходные данные по типам ВС

Тип ВС	Кол-во	Удельные затраты	Сар
A320	4	0.08	164
B735	2	0.0775	138
B772	3	0.077	305

Таблица 3

Оборотное время (мин.)

Аэропорт	A320	B735	B772
CDG	35	40	60
LED	40	45	55
OVB	45	50	70
SVO	40	45	60

Для выполнения расчетов по описанным выше математическим моделям был создан программный комплекс со встроенным оптимизатором. В результате для «рейсовой» модели задачи FAM найдено оптимальное решение (табл. 4). Общее количество переменных в задаче целочисленного программирования составило 124, а количество ограничений оказалось равно 115. Время счета программы составило 3.7 с на PC с процессором Intel®Core™2 Duo CPU T7500@2.20GHz и RAM 2 Gb.

Таблица 4

Оптимальное решение «рейсовой» задачи FAM для модельного примера

Тип ВС	Номера рейсов
A320	4, 12, 15, 24, 25, 35
B735	2, 13, 22, 33
B772	1, 3, 5, 6, 11, 14, 21, 23, 31, 32, 34, 36

Для каждого подмножества рейсов, на которые в полученном оптимальном решении назначен один и тот же тип ВС, построены цепочки стыкующихся между собой рейсов. Количество цепочек не должно превышать числа имеющихся ВС данного типа и каждый рейс подмножества должен содержаться в одной, и только в одной цепочке. В результате получено более 20 различных наборов цепочек, причем в одном случае оказалось возможным использовать только три из четырех имеющихся ВС типа A320, что, конечно же, является дополнительным преимуществом такого набора (табл. 5).

Таблица 5

Набор цепочек, полученный из решения «рейсовой» задачи FAM

Тип ВС	Кол-во ВС	Цепочки рейсов
A320	4	24, 25 – 15, 35 – 12 – 4
B735	2	13 – 33, 22 – 2
B772	3	11 – 3, 21 – 5 – 32 – 6 – 36, 23 – 31 – 1 – 14 – 34

Для альтернативного способа решения задачи расстановки парка ВС, основанного на предварительном построении цепочек рейсов, также было получено оптимальное решение в созданном программном комплексе. Из 22 рейсов полетного расписания было образовано 137 цепочек, что соответствует 411 неизвестным в задаче целочисленного программирования. Матрица ограничений включает 174 строки. Время расчета на том же компьютере, что был использован и в предыдущем случае, оказалось равным 0.73 с., а сам результат, т.е. оптимальный набор цепочек, совпал с указанным в табл. 5 как по расстановке типов ВС, так и по значению функции цели.

Если в качестве дополнительного условия при формировании цепочек потребовать их цикличности (замкнутости), т.е. совпадения начального и конечного аэропортов, то их количество окажется равным всего 39. Для 117 неизвестных в оптимизационной задаче определены 64 ограничения, и время счета составило менее 0.5 с. Полученный оптимальный набор цепочек отличается от предыдущего (табл. 6), а функция цели (ожидаемая прибыль, связанная с расстановкой типов ВС) уменьшилась примерно на 6%. Последнее вызвано, очевидно, значительным сужением области допустимых решений.

Таблица 6

Набор замкнутых оптимальных цепочек

Тип ВС	Кол-во ВС	Цепочки рейсов
A320	4	25 – 15, 24 – 34, 22 – 4, 1 – 12
B735	2	35 – 14, 11 – 2
B772	3	13 – 32 – 6 – 36, 21 – 5 – 33, 23 – 31 – 3

Кроме расчетов для модельного примера были проведены и вычисления оптимальной расстановки типов ВС по рейсам однодневного полетного расписания для одной из крупнейших российских авиакомпаний. В рассмотренном расписании содержалось 84 рейса, авиакомпания использовала 7 различных типов ВС. Расчет основан на реальных экономических данных и учитывает следующие операционные ограничения:

- по имеющемуся количеству ВС каждого типа;
- по емкости и дальности полета типов ВС;
- по возможности взлета/посадки типов ВС в аэропортах полетного расписания;
- по оборотному времени для каждого типа ВС в аэропортах полетного расписания.

Расчетная оптимальная расстановка типов ВС показала увеличение прибыли более чем на 25% в день по сравнению с расстановкой, реализованной в самой авиакомпании, а также привела к сокращению числа используемых в день самолетов и повышению коэффициента утилизации (отношению общего времени, в течение которого ВС данного типа находятся в воздухе, к временному интервалу рассматриваемого расписания).

5. Заключение

В статье дана математическая формулировка задачи оптимальной расстановки парка ВС по рейсам полетного расписания авиакомпании, основанной на использовании цепочек стыкующихся между собой рейсов. Проведен сравнительный анализ задачи по результатам расчетов, для модельного примера и для однодневного полетного расписания одной из крупнейших российских авиакомпаний в созданном авторами программном комплексе. Показано, что применение оптимизационных подходов к задачам составления расписаний в деятельности авиакомпаний дает высокую экономическую отдачу.

ЛИТЕРАТУРА

1. **Виноградов Л.В., Фридман Г.М., Шебалов С.М.** Математическое моделирование в оптимизации авиационных перевозок: перспективы развития и эффект от использования // Научный Вестник МГТУ ГА. - 2008. - № 132.
2. **Barnhart C., Belobaba P., and Odoni A.** Application operation research in the air transport industry. Transportation Science vol. 37, No. 4, 2003.
3. **Егорова А.А., Козлов С.А.** Информационные системы: методы и средства проектирования // Научный Вестник МГТУ ГА. - 2006. - № 105.
4. **Липский В.** Комбинаторика для программистов. - М.: Мир, 1988.
5. **Smith B.C.** Robust Airline Fleet Assignment. Ph.D. thesis. Georgia Institute of Technology, USA, 2004.
6. **Корбут А.А., Финкельштейн Ю.Ю.** Дискретное программирование - М.: Наука, 1969.
7. **Barnhart, C., Boland N.L., Clarke L.W., Johnson E.L., Nemhauser G.L., Sheno R.G.** Flight String Models for Aircraft Fleeting and Routing. Transportation Science, 32, 3, 1998.
8. **Ageeva Y.** Approaches to Incorporating Robustness into Airline Scheduling. Masters Thesis, Operations Research Center, Massachusetts Institute of Technology, 2000.

FLEET ASSIGNMENT PROBLEM SOLUTION BASED ON FLIGHT STRINGS APPROACH

Fridman G.M., Zenkova N.A.

An approach to the fleet assignment problem solution based on flight flows generation is presented. The analysis reveals that such an approach has its own advantages which are discussed in the paper. Numerical results are presented for the model input data as well as for the real-scale data of a major Russian airline.

Key words: flight flows, optimal fleet assignment problem, integer programming.

Сведения об авторах

Фридман Григорий Морицович, 1964 г.р., окончил СПбГМТУ (1986), доктор технических наук, профессор кафедры экономической кибернетики и математических методов в экономике СПбГУЭФ и кафедры прикладной математики и математического моделирования СПбГМТУ, автор более 60 научных работ, область научных интересов – методы оптимизации, оптимизация планирования операций в авиатранспортной индустрии, механика жидкости.

Зенкова Надежда Александровна, окончила СПбГМТУ (2008), аспирантка СПбГМТУ, эксперт по управлению доходностью ЗАО «Сирена-Трэвел», автор 3 научных работ, область научных интересов – математическое программирование, методы оптимального управления, оптимизация планирования операций в авиатранспортной индустрии.

УДК 621.313.684

ВОССТАНОВЛЕНИЕ СТЁРТЫХ ПАКЕТОВ В КОМПЬЮТЕРНЫХ СЕТЯХ

В.И. КУБИЦКИЙ

Статья представлена доктором технических наук, профессором Рудельсоном Л.Е.

Предлагается метод восстановления информационных пакетов при их потере в компьютерных сетях, основанный на использовании кодов Лагранжа. Разработаны алгоритмы вычисления контрольных пакетов, передаваемых по сети в дополнение к информационным пакетам, и алгоритмы восстановления стёртых пакетов. Определено количество операций в конечных полях и объёмы памяти для хранения постоянных величин при реализации этих алгоритмов.

Ключевые слова: компьютерные сети, кодирование, восстановление пакетов, стирающие коды.

Введение

В компьютерных сетях обработки данных для автоматизированных систем управления воздушным движением (АС УВД) существуют помехи, шумы, ведущие к потерям и искажениям данных. Одной из актуальных проблем является проблема потерь и задержек пакетов, что вызвано перегрузками в узлах (коммутаторов, маршрутизаторов и т.п.), коллизиями (наложением пакетов от разных абонентов), «проседанием» пропускной способности каналов передачи данных вследствие слишком большого количества одновременно подключённых пользователей, искажений в пакете, переполнения входных буферов сетевых устройств и др. В системах передачи данных и сетевых приложениях, работающих в режиме реального времени, задержки пакетов равнозначны потерям, поскольку нет возможности приостанавливать процесс обработки, передачи и отображения данных в ожидании пакета или его повторной передачи. Кроме того, запросы на пересылку, особенно в топологии «один-многим», приводят к дополнительным нагрузкам на канал передачи данных, что уменьшает объём трафика в сети. К тому же механизм обратной связи даёт положительный результат, если ошибки или потери пакетов носят случайный характер. При систематических ошибках и потерях пакетов их восстановление в отведённое время с использованием обратной связи весьма затруднительно. Поэтому в современных компьютерных сетях с большими объёмами цифровых данных необходимы методы обмена без использования обратных каналов. Нужны новые методы защиты от ошибок, которые не предусматривают наличие обратного канала.

Рассмотрим некоторые существующие и предложим новые методы восстановления пакетов. Будем использовать следующие термины:

сообщение (message) – блок информации, состоящий из последовательности пакетов;

пакет (packet) – блок информации, содержащий ограничители начала и конца пакета, заголовков, передаваемые данные и код для контроля ошибок.

1. Методы восстановления пакетов в компьютерных сетях

В реальных компьютерных сетях АС УВД часть пакетов может быть не доставлена получателю (потеряна), а часть – доставлена с обнаруженными получателем ошибками. Такие пакеты (потерянные и с ошибками) в теории помехоустойчивого кодирования считаются стёртыми [1], а в качестве базовой математической модели для описания канала с помехами в компьютерных сетях используется модель канала со стираниями пакетов.

Для обнаружения ошибок в пакетах получателем отправитель обычно использует коды. Простейшим примером является контрольная сумма бит пакета или проверка на чётность, что недостаточно надёжно, особенно при появлении пачек ошибок. В большинстве сетевых приложений применяются более сложные, но и более надёжные циклические коды, к числу которых относятся коды CRC (Cyclic Redundancy Code), способные обнаруживать как независимые ошибки, так и серии ошибок в последовательно расположенных битах, и гарантированно обнаруживают серии ошибок длины r (где $r=4, 8, 16, 32$) [1].

Среди протоколов с обнаружением ошибок можно выделить каналные протоколы передачи данных по телефонным каналам, протокол XModem – использует контрольную сумму, протоколы XModem-CRC, XModem-1k, YModem, ZModem – используют циклический код с 16-разрядной проверяющей комбинацией.

Для обнаружения потерянных пакетов могут использоваться протоколы, предусматривающие в структуре пакета его номер. Реализацию порядковой нумерации пакетов данных обеспечивают, например, протоколы RTP, IP, входящие в стек протоколов TCP/IP.

После обнаружения ошибочных и потерянных пакетов включаются механизмы восстановления пакетов с использованием протоколов, основанных на наличии обратного канала от получателя к отправителю. Такие протоколы традиционно используются для повышения достоверности доставки сообщений в каналах со стиранием.

Примером протоколов с использованием обратного канала являются протоколы с гарантированной доставкой данных, такие как TCP (Transmission Control Protocol), FTP (File Transfer Protocol), HTTP (Hypertext Transfer Protocol) и другие. Позитивное подтверждение с повторной передачей, положенное в основу рассмотренных протоколов, является одним из методов, гарантирующих надёжную доставку данных.

Однако в приложениях, работающих в режиме реального времени и подверженных риску потерь пакетов при передаче (например, цифровое телевидение через Интернет, Интернет-телефония, распределенные сетевые игры и т.д.), протоколы с гарантированной доставкой данных не применяются, так как используемая в них повторная пересылка данных приводит к увеличению задержек. А в этих приложениях введены жёсткие ограничения на максимально допустимую задержку данных в цепи «отправитель-получатель». Для таких приложений используются протоколы с негарантированной доставкой пакетов, такие как IP (Internet Protocol), UDP (User Datagram Protocol), RTP (Real-time Transport Protocol), не использующие повторную пересылку данных в случае потерь. Протоколы с негарантированной доставкой пакетов разрабатывались для приложений, в которых приемлемы потери пакетов и ошибки передачи, поэтому они не могут полностью исключить такие события. В то же время для многих приложений недопустимы потери или порча данных, как и повторный запрос. Например, при потере в сетях хранения данных (Storage Area Networks – SAN) большинство ленточных устройств прерывают операцию резервного копирования.

Как отмечено выше, контрольная сумма и коды CRC используются для обнаружения ошибок в пакетах. Для исправления ошибок в пакете используют корректирующий код Хэмминга, при этом обратный канал не используется. Однако применение кодов CRC, Хэмминга и т. п. для кодирования пакетов не решает проблему восстановления всего стёртого пакета без использования обратного канала. Не могут решить проблему другие классические помехоустойчивые коды: свёрточные коды исправляют отдельные битовые ошибки, многие блочные коды способны исправлять пачки ошибок в отдельно взятом пакете [2].

Для восстановления целого пакета в случае его потери созданы новые помехоустойчивые стирающие коды (коды Лаби) [3], которые позволяют закодировать исходное сообщение конечной длины, состоящее из символов (или пакетов) любой одинаковой длины, потенциально неограниченным потоком кодовых символов (пакетов). Это свойство позволяет говорить о стирающих кодах, как о кодах с нефиксированной скоростью. В случае необходимости кодер стирающего кода по запросу всегда может добавить "на лету" (on-the-fly, on-line) небольшое число

кодовых символов. Для классических кодов добавить "на лету" несколько контрольных символов не всегда удаётся, контрольные символы оказываются зависимыми друг от друга. Такая проблема существенна для блочного кода Рида-Соломона. При переходе к большему числу контрольных символов требуется повторное вычисление всех ранее вычисленных контрольных символов. Без учёта этого недостатка код Рида-Соломона можно использовать для доставки сообщений в канале со стираниями, если символы этого кода, состоящие из m бит, рассматривать как пакеты сообщения.

Следует отметить некоторые недостатки, которые препятствуют широкому внедрению кодов Лаби (кодов ЛТ). Для кодов ЛТ, использующих случайную выборку соседей (исходные символы, использованные для генерации некоторого кодового символа), вероятность восстановления пакетов может существенно (и неконтролируемо) варьироваться. Некоторые пакеты, в случае неудачной выборки соседей, могут оказаться невозможными даже при полном отсутствии потерь пакетов.

Ввиду того, что в основе теории кодов ЛТ лежит статистическая задача о мячах, бросаемых случайным образом в корзины, эффективность этих кодов проявляется при достаточно большом числе исходных символов (пакетов). Для кодов ЛТ количество информационных пакетов (K) рекомендуется порядка 10000 [3], при низких значениях K (порядка 100-500 пакетов) не работают статистические свойства, и эффективность кодов ЛТ может существенно снизиться. Снижение эффективности кодов ЛТ может наблюдаться, например, в сетях Ethernet, FDDI, X.25. Это вызвано тем, что для этих сетей максимальный размер поля данных кадра или пакета (MTU - Maximum Transfer Unit), в которые должен инкапсулировать свой пакет протокол IP, меньше количества информационных пакетов (10000), для которого проявляется эффективность кодов ЛТ. Сети Ethernet имеют значение MTU, равное 1500 байт, сети FDDI - 4096 байт, а сети X.25 чаще всего работают с MTU в 128 байт.

Для полного декодирования всех исходных пакетов коды ЛТ предполагают возможность досылки кодовых пакетов до тех пор, пока все исходные пакеты не будут восстановлены. Для потоковых сетей в режиме реального времени (например, для «живого» IPTV) возможность ожидания досылки отсутствует. Каждое сообщение должно быть получено (обработано) в строго определённое время, и ждать, пока придёт дополнительный пакет для восстановления данных сообщения, нельзя.

Необходимость буферизации большого числа исходных пакетов (т. е. эффективность кодов ЛТ начинает проявляться для большого числа исходных пакетов) вызывает длительную (до нескольких десятков секунд [4]) задержку сигнала между приёмником и передатчиком.

Внедрение новых стирающих кодов в компьютерных сетях потребует внесения изменений в принятую технологию обмена информацией.

Отмеченные недостатки рассмотренных кодов приводят к необходимости разработки для вычислительных систем реального времени кодов, лишённых этих недостатков.

2. Требования к кодам для вычислительных систем

Помехоустойчивые коды для АС УВД и других систем, работающих в реальном масштабе времени, должны отвечать следующим требованиям:

1) обладать способностью восстановления данных (сообщения) в случае стирания целых пакетов при передаче и при этом не требовать наличия канала обратной связи между отправителем и получателем, так как запрос и повторная передача пакетов в системах реального времени (особенно в системах дальней космической связи) недопустима;

2) для восстановления исходного сообщения получателем должны быть использованы любые полученные пакеты;

3) иметь минимальную избыточность при высоких корректирующих возможностях;

4) обеспечивать по возможности гибкую схему кодирования (добавлять "на лету" контрольные символы), в зависимости от степени зашумлённости канала передачи;

- 5) контрольные символы (пакеты) должны формироваться независимо друг от друга;
- 6) кодирование/декодирование кода должно осуществляться с использованием минимального объёма вычислительных ресурсов;
- 7) учитывать байтовую структуру данных;
- 8) обладать способностью обнаружения и/или исправления ошибок при передаче, обработке и хранении информации;
- 9) обладать способностью защиты функциональных преобразований при обработке;
- 10) внедрение помехоустойчивого кодирования в сетях должно обеспечивать совместимость с существующими процедурами передачи данных в компьютерных сетях.

Исследования показывают, что наиболее полно удовлетворяют этим требованиям полиномиальные коды Лагранжа, введённые в [5].

3. Коды Лагранжа для восстановления стёртых пакетов

Пусть исходное сообщение, состоящее из k исходных пакетов, передаётся по каналу со стиранием с помощью n кодовых пакетов. Длины всех пакетов одинаковы. При передаче некоторая часть пакетов, предположительно не большая ($n-k$), будет стёрта. Оставшаяся часть, предположительно не меньшая k , будет правильно принята получателем. Под стиранием пакетов будем понимать неполученные адресатом пакеты и пакеты с обнаруженными ошибками. В сетях, где используется сквозная нумерация передаваемых пакетов, существует возможность определить номер потерянного пакета. Таким образом, стиранием является ошибка, местоположение которой в сообщении известно, но неизвестна её величина.

Для обеспечения возможности восстановления стёртых пакетов представляется необходимым кодировать всё сообщение, состоящее из пакетов.

3.1. Общая схема кодирования

Код Лагранжа представляет собой блочный код, в котором символы состоят из m бит. Если эти символы рассматривать как пакеты сообщения, то код может быть использован для доставки сообщений в канале со стираниями [6]. При этом основным свойством кода является следующее: для доставки k информационных пакетов (символов) достаточно принять любые k пакетов (символов) из n , то есть стёртыми могут быть любые из $r=n-k$ пакетов (символов). Для кода Лагранжа не существует проблемы добавления «на лету» небольшого числа контрольных пакетов (символов), так как число контрольных символов $r \leq n-k$ (где k – фиксировано для конкретного сообщения, $N=p^m$ – количество элементов конечного поля) и контрольные символы независимы друг от друга. Поэтому при переходе от $r=n-k$ ($r+k=n \leq N$) контрольных пакетов (символов) к большему числу $r'=n'-k$ ($r'+k=n' \leq N$) все контрольные пакеты (символы) не требуют повторного вычисления. Это свойство кодов Лагранжа вытекает из положений работы [7].

Если каждый из пакетов сообщения не рассматривать как m -битовый символ, то необходимо применять следующую процедуру кодирования сообщения.

1. Формируются информационные пакеты исходного сообщения:

- а) исходное сообщение, подлежащее передаче, разбивается на части (блоки данных – блоки данных информационных пакетов), количество которых равно k ;
- б) каждая часть (блок данных) разбивается на части одинаковой длины m (элементы), кодируется и дополняется контрольным кодом, который будем называть внутренним кодом.

Количество элементов i -го блока данных равно $\lambda_i = \frac{L_i}{m}$, $i = \overline{1, k}$, где L_i – длина i -го блока данных (бит);

- в) каждый закодированный блок данных исходного сообщения дополняется атрибутами для передачи по сети.

В результате получим k информационных пакетов.

Реализация 1-го шага процедуры кодирования исходного сообщения может быть заложена в существующих сетях.

2. В дополнение к k информационным пакетам формируются r контрольных пакетов:

а) по правилам кодирования выбранного кода производится формирование блоков данных контрольных пакетов.

При этом каждый p -й ($p = \overline{1, \lambda}$) элемент блока данных в каждом j -м ($j = \overline{1, r}$) контрольном пакете кодируется кодом, который будем называть внешним кодом. Этот код, каждый символ которого состоит из t двоичных бит, является недвоичным.

Длина каждого p -го кодового слова внешнего кода составляет $n_p \leq 2^m$;

б) блок данных каждого контрольного пакета кодируется и дополняется контрольным кодом (внутренним кодом) и атрибутами для передачи по сети в соответствии с процедурой, которая может быть заложена в существующих сетях и применяемой для информационных пакетов. Таким образом, получаем структуру контрольных пакетов, аналогичную структуре информационных пакетов, что очень важно для адаптации предлагаемого метода к использованию в существующих сетях.

Количество контрольных пакетов (r) определяется максимальным количеством пакетов, которое необходимо будет восстановить после стирания. Общее количество передаваемых пакетов составит $n = k + r$.

Отметим, что формирование контрольных пакетов будет происходить без задержки передачи исходных информационных пакетов.

Таким образом, получена 2-хуровневая каскадная конструкция. 1-й уровень этой конструкции содержит внутренний код, полученный на 1-м шаге процедуры кодирования, 2-й уровень – внешний код, полученный на 2-м шаге. Внутренним кодом (кодом 1-го уровня) может быть используемый в существующих сетях код (контрольная сумма, CRC, Хэмминга и др.). В качестве внешнего кода (кода 2-го уровня) предлагается выбрать код Лагранжа.

3.2. Кодирование сообщения кодом Лагранжа

Формирование блоков данных контрольных пакетов будет производиться поэлементно (по-символьно) по интерполяционной формуле Лагранжа с применением одного из алгоритмов: параллельного, последовательного или параллельно-последовательного [8].

Поскольку последовательный алгоритм требует меньшего количества операций в конечном поле, то кодирование будет выполняться с применением этого алгоритма в соответствии с выражением

$$f^{(p)}(\beta_j) = - \sum_{i=0}^{s+j-1} f_i^{(p)} \prod_{l=j+1}^r \frac{x_i - \beta_l}{\beta_j - \beta_l}, \quad j = \overline{1, r}, \quad p = \overline{1, \lambda}, \quad (1)$$

где $\prod_{l=j+1}^r \frac{x_i - \beta_l}{\beta_j - \beta_l} = -L_{S_{j-1}}^{(i)}(\beta_j) = L_{T_j}^{(j)}(x_i)$, $S_{j-1} = S \cup \{\beta_1, \dots, \beta_{j-1}\}$, $T_j = T \setminus \{\beta_1, \dots, \beta_j\}$,

$$T_r = \emptyset, \quad x_i \in S_{j-1}, \quad \beta_j \in T_j;$$

$S = \{x_0, \dots, x_s\}$ – множество информационных узлов мощности k ;

$T = \{\beta_1, \dots, \beta_r\}$ – множество контрольных узлов мощности r ;

$f^{(p)}(\beta_j) = f_{s+j}^{(p)}$ – вычисленное значение p -го элемента в j -м контрольном пакете;

$f_{s+j-1}^{(p)}$ – вычисленные значения p -го элемента в предыдущих $(j-1)$ -х пакетах;

$f_i^{(p)}$ – значение p -го элемента в i -м информационном пакете ($i = \overline{0, s}$);

λ – количество элементов в блоке данных пакета (длина блока данных пакета).

В [6] доказано, что при любом n ($2e \leq n \leq q$, где n – длина кодового слова, $q = p^m$ – количество элементов конечного поля $\text{GF}(p^m)$) для исправления e стираний кодом Лагранжа достаточное количество контрольных узлов (пакетов), то есть $r=e$.

Алгоритм вычисления всех (r) контрольных символов по каждому p -му элементу ($p = \overline{1, \lambda}$) в соответствии с формулой (1) будет следующим:

1. Для каждого $i = \overline{0, s}$, задавая $l = \overline{r, 2}$, вычисляем величины $a_{il} = x_i - \beta_l$ и определяем суммы: $\sum_{i=0}^s f_i a_{ir}$, $\sum_{i=0}^s (f_i a_{ir}) a_{i,r-1}$, ..., $\sum_{i=0}^s ((\dots (f_i a_{ir}) \dots) a_{i3}) a_{i2}$. Отдельно определяется $\sum_{i=0}^s f_i$.

2. Вычисляем $\gamma_j = -1 / \prod_{l=j+1}^r (\beta_j - \beta_l)$. На первом шаге $j = 1$, в дальнейшем каждый раз при обращении к этому пункту увеличиваем j на единицу до $(r-1)$.

3. Определяем контрольные символы f_{s+j} . При $j = r-1$ переходим к пункту 5.

4. Для каждого $j = \overline{1, r-1}$, задавая значения $l = \overline{r, j+2}$, вычисляем произведения $\alpha_j^{(l)} b_{jl}$, где $\alpha_j^{(l)}$: $\alpha_j^{(r)} = f_{s+j}$, $\alpha_j^{(r-1)} = f_{s+j} b_{jr}$, ..., $\alpha_j^{(j+2)} = (\dots (f_{s+j} b_{jr}) \dots) b_{j,j+3}$; $b_{jl} = \beta_j - \beta_l$. При этом прибавляем величины f_{s+j} к сумме $\sum_{i=0}^s f_i$. После каждого изменения j выполняем пункт 2.

5. Суммируя f_{s+r-1} и $\sum_{i=0}^{s+r-2} f_i$, вычисляем f_{s+r} .

Здесь $b_{12} = \beta_1 - \beta_2$, $b_{13} = \beta_1 - \beta_3$, $b_{23} = \beta_2 - \beta_3$, ..., $b_{1r} = \beta_1 - \beta_r$,
 $b_{2r} = \beta_2 - \beta_r$, ..., $b_{r-1,r} = \beta_{r-1} - \beta_r$, $a_{i2} = x_i - \beta_2$, $a_{i3} = x_i - \beta_3$, ..., $a_{ir} = x_i - \beta_r$,
 $\gamma_1 = -1/(b_{12} b_{13} \dots b_{1r})$, $\gamma_2 = -1/(b_{23} b_{24} \dots b_{2r})$, ..., $\gamma_{r-1} = -1/b_{r-1,r}$.

Для вычисления всех (r) контрольных символов по каждому p -му элементу ($p = \overline{1, \lambda}$) в соответствии с этим алгоритмом потребуется выполнить операции сложения, умножения и инвертирования в конечном поле соответственно в количестве

$$N_{\oplus} = n(2r-1) - r(r+1), \quad N_{\otimes} = (n-1)(r-1), \quad N_{\ominus} = r-1.$$

Расчёт количества операций производился с учётом наличия $(r-1)$ регистров памяти для хранения величин b_{jl} . С учётом наличия памяти для хранения n узлов интерполирования необходимо иметь $(n+r-1)$ регистров.

Для получения всех элементов блоков данных контрольных пакетов необходимо применить приведенный алгоритм кодирования λ раз.

При формировании блоков данных контрольных пакетов исходного сообщения для уменьшения количества операций можно использовать параллельно-последовательный алгоритм кодирования кода Лагранжа [8]. В этом случае узлы интерполирования являются фиксированными и фундаментальные полиномы Лагранжа $L^{(i)}(x)$ можно вычислить заранее и хранить в памяти как постоянные величины. Значения элементов блоков данных контрольных пакетов вычисляются по формулам

$$f^{(p)}(\beta_j) = \sum_{i=0}^s f_i^{(p)} A_{ji}, \quad f^{(p)}(\beta_r) = -[\sum_{i=0}^s f_i^{(p)} + \sum_{j=1}^{r-1} f^{(p)}(\beta_j)], \quad (2)$$

где $A_{ji} = L_s^{(i)}(\beta_j) = \text{const}$, $j = \overline{1, r-1}$.

Для вычисления всех (r) контрольных символов по каждому p -му элементу ($p = \overline{1, \lambda}$) в соответствии с формулой (2) потребуется выполнить операции сложения и умножения в конечном поле соответственно в количестве $N_{\oplus} = r(n-r) - 1$, $N_{\otimes} = (r-1)(n-r)$.

Для хранения коэффициентов $L^{(i)}(x)$ необходимо иметь $(r-1)(n-r)$ регистров.

3.3. Восстановление стёртых пакетов

Определить, какие пакеты в процессе передачи потеряны, можно с использованием механизмов, заложенных в протоколах обмена данными в сетях. Пакеты с ошибками определяются путём декодирования внутреннего кода. Потерянные и ошибочные пакеты отмечаются как стёртые и декодируются внешним кодом.

Из [5] следует, что любые $(n-k)$ символов кода Лагранжа (k – количество информационных символов) могут считаться избыточными (контрольными). Это означает, что в качестве контрольных узлов в случае стирания можно взять стёртые узлы (если их не больше, чем $n-k$). Расширяясь на эти узлы, получим значения символов, которые соответствуют стёртым значениям.

Таким образом, восстановление блоков данных стёртых пакетов можно производить в соответствии с процедурой кодирования кода Лагранжа. Делается это поэлементно для каждого стёртого пакета в соответствии с выражением

$$f^{(p)}(v_\xi) = - \sum_{i=0}^{s'+\xi-1} f_i^{(p)} \prod_{l=\xi+1}^e \frac{x'_i - v_l}{v_\xi - v_l}, \quad \xi = \overline{1, e}, \quad p = \overline{1, \lambda}, \quad (3)$$

где $\prod_{l=\xi+1}^e \frac{x'_i - v_l}{v_\xi - v_l} = -L_{S'_{\xi-1}}^{(i)}(v_\xi) = L_{V_\xi}^{(\xi)}(x'_i)$, $S'_{\xi-1} = S' \cup \{v_1, \dots, v_{\xi-1}\}$, $V_\xi = V \setminus \{v_1, \dots, v_\xi\}$,

$$V_e = \emptyset, \quad x'_i \in S'_{\xi-1}, \quad v_\xi \in V_\xi;$$

$S' = \{x'_0, \dots, x'_{s'}\}$ – множество узлов мощности k' принятых (нестёртых) пакетов;

$V = \{v_1, \dots, v_e\}$ – множество узлов мощности e стёртых пакетов;

$f^{(p)}(v_\xi) = f_{s'+\xi}^{(p)}$ – вычисленное значение p -го элемента в ξ -м восстанавливаемом пакете;

$f_{s'+\xi-1}^{(p)}$ – вычисленные значения p -го элемента в предыдущих $(\xi-1)$ -х восстановленных пакетах;

$f_i^{(p)}$ – значение p -го элемента в i -м нестёртом пакете ($i = \overline{0, s'}$);

λ – количество элементов в блоке данных пакета (длина блока данных пакета).

При необходимости восстановленные блоки данных стёртых пакетов кодируются и дополняются контрольным кодом (внутренним кодом) и атрибутами для передачи по сети в соответствии с процедурой, которая может быть заложена в существующих сетях и применяемой для информационных пакетов.

Заключение

Рассмотрена возможность применения корректирующих кодов для восстановления данных при потерях пакетов в компьютерных сетях. Установлено, что применение для этих целей стирающих кодов, основанных на методах статистического кодирования, ограничено свойствами этих кодов. Разработаны требования к кодам для использования в вычислительных системах. Предложенные для исправления стираний коды Лагранжа, удовлетворяющие этим требованиям, имеют простые алгоритмы декодирования, сводящиеся к вычислению стёртых пакетов (символов) посредством алгоритмов кодирования. Эти коды, являясь блоковыми кодами, позволяют добавлять «на лету» контрольные символы. Тем самым можно получить код с нефиксированной скоростью. Коды могут быть использованы в любом канале со стираниями, независимо от статистики стираний. Алгоритмы кодирования и декодирования принципиально не зависят от размера пакета.

Предложенный метод восстановления стертых пакетов не нарушает принятые процедуры передачи данных в компьютерных сетях, что очень важно для его реализации в существующих системах. Метод позволяет восстанавливать стёртые пакеты без запроса на их повторную передачу. Это разгружает каналы связи от ненужных обменов данными, а также обеспечивает работоспособность сетевых приложений, работающих в режиме реального времени, для которых задержки пакетов равнозначны потерям, поскольку нет возможности приостанавливать процесс обработки, передачи и отображения данных в ожидании пакета или его повторной передачи.

Разработанные алгоритмы дают возможность выполнить программную и аппаратную реализацию восстановления стёртых пакетов.

ЛИТЕРАТУРА

1. Варгазин В.А. Помехоустойчивое кодирование в пакетных сетях // ТелеМультиМедиа. - 2005. - № 3.
2. Питерсон У., Уэлдон Э. Коды, исправляющие ошибки / пер. с англ. / под ред. Р.Л. Добрушина и С.И. Самойленко. - М.: Мир, 1976.
3. Luby M. LT Codes // Proc. of the 43rd Annual IEEE Symposium on Foundations of Computer Science (FOCS). - 2002.
4. Шинкаренко К.В., Кориков А.М. Помехоустойчивое кодирование мультимедиа данных в компьютерных сетях // Известия Томского политехнического университета. - 2008. - Т. 313. - № 5.
5. Амербаев В.М. Теоретические основы машинной арифметики. - Алма-Ата: Наука, 1976.
6. Кубицкий В.И. Процедуры декодирования в каналах с ошибками и стираниями // Проблеми інформатизації та управління: Збірник наукових праць. - К.: НАУ, 2007. - Вип. 3 (21).
7. Кубицкий В. И. Кодирование для неполного кода Лагранжа // Проблеми інформатизації та управління: Збірник наукових праць. - К.: НАУ, 2005. - Вип. 4 (15).
8. Кубицкий В.И. Процедуры кодирования и декодирования для полиномиальных кодов // Эксплуатация программного обеспечения систем реального времени, построенных на базе микро - и мини-ЭВМ: сб. научн. тр. - Киев: КИИГА, 1989.

RENEWAL OF THE LOST PACKETS IN COMPUTER NETWORKS

Kubitskiy V.I.

The method of renewal of informative packets is offered at their loss in computer networks, based on the use of Lagrange codes. The algorithms of calculation of check packets, transferred on a network in addition to the informative packets, and algorithms of renewal of the lost packets, are developed. The amount of operations in the finite fields and memory capacity are defined for storage of constant values during realization of these algorithms.

Key words: computer networks, coding, renewal packets, erasure codes.

Сведения об авторе

Кубицкий Валерий Иванович, 1951 г. р., окончил КИИГА (1973), начальник сектора ФГУП ГосНИИ «Аэронавигация», автор более 40 научных работ, область научных интересов – проектирование, эксплуатация, надёжность и помехоустойчивость АС УВД, компьютерные системы и компоненты.

УДК 002.5:65

ПРИНЦИПЫ ОЦЕНКИ СОСТОЯНИЯ СИСТЕМЫ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ АВИАПРЕДПРИЯТИЯ

А.И. ТЕРЕНТЬЕВ

Статья представлена доктором технических наук, профессором Соломенцевым В.В.

Излагаются некоторые подходы и принципы формирования частных и комплексных (итоговых) оценок состояния системы информационной безопасности авиапредприятия.

Ключевые слова: аудит информационной безопасности авиапредприятия, уровни безопасности.

Совершенный 24 января 2011 года террористический акт в аэропорту Домодедово убедительно свидетельствует об особой актуальности работ по оценке состояния и последующему совершенствованию системы комплексной безопасности объектов критической инфраструктуры Российской Федерации, к которым, несомненно, относятся гражданские и военные авиапредприятия. При осуществлении этой деятельности необходимо исключить возможность получения злоумышленниками результатов такой оценки, а также другой значимой информации о функционировании и фактическом состоянии системы безопасности, управления и жизнеобеспечения соответствующих объектов. Достижение этой цели связано с созданием, контролем текущего состояния и поддержанием на адекватном угрозам уровне системы информационной безопасности авиапредприятия.

1. Условные уровни безопасности авиапредприятия и их объектов

Общеизвестно, что даже самая развитая нормативно-правовая база, адекватная и безупречная политика безопасности определяют только потенциальную способность системы комплексной безопасности защищаемых объектов противостоять принятой модели угроз. На практике реальный уровень безопасности зависит от совокупности различных по своей природе факторов, в том числе от эффективности организационно-управленческой деятельности, квалификации специалистов и морально-психологического состояния коллектива, что в совокупности принято называть "человеческим фактором".

Необходимо различать следующие условные уровни информационной безопасности авиапредприятий и их объектов:

потенциальный, являющийся максимально достижимым уровнем безопасности при полном и наиболее эффективном использовании имеющихся материальных, технических, интеллектуальных и других ресурсов, а также стабильной общественно-политической, экологической, техногенной и т.п. ситуации в регионе. Этот уровень безопасности определяется сторонними (внешними) системными аналитиками на основе формальных данных о защищаемом объекте;

декларируемый (желаемый, самопровозглашенный) уровень безопасности, обусловленный спецификой защищаемого объекта, наличием необходимых материально-технических ресурсов и собственным представлением работающих на нем руководителей и специалистов об эффективности используемых методов и средств обеспечения его безопасности. Этот уровень определяется в процессе самоанализа, который условно можно назвать "аудитом первой стороной";

реальный ("квазиобъективный") уровень безопасности, являющийся действующим в конкретный момент времени уровнем безопасности защищаемого объекта, обусловленный влиянием труднопредсказуемого "человеческого фактора", объективным ослаблением со временем от-

ветственности, бдительности, защитных механизмов и другими специфическими факторами. Этот уровень безопасности определяется приглашенными (внешними) специалистами в процессе "аудита второй (или третьей) стороной".

Интуитивно понятно, что указанные уровни существенно отличаются друг от друга. При этом, реальный уровень безопасности всегда существенно ниже декларируемого и, тем более, потенциального. У авиапредприятий, способных обеспечить примерное равенство (соответствие) указанных уровней, так называемый коэффициент полезного действия использования затраченных на безопасность средств и ресурсов приближается к 100%.

2. Аудит информационной безопасности авиапредприятий и формализация его результатов

На практике важнейшее значение приобретает периодический квалифицированный и целенаправленный анализ состояния (уровня) информационной защищенности критических объектов, адекватности принятой политики безопасности выявленным угрозам и наличия потенциально уязвимых мест в действующей системе (подсистеме) безопасности, проводимый по результатам протоколирования (документирования) значимых и предварительно определенных для этих целей событий и состояний, отмеченных на защищаемом объекте за определенный период времени.

Указанный анализ принято называть аудитом безопасности (состояния защищенности). Известно, что он является эффективным профилактическим и сдерживающим средством против нарушений требований по обеспечению безопасности, поскольку позволяет осуществлять действенный контроль текущего состояния безопасности и эффективности деятельности персонала, а также прогнозировать потенциальные нарушения принятой политики безопасности и разрабатывать адекватные меры противодействия.

К сожалению, в настоящее время оценка состояния системы информационной безопасности защищаемых объектов, как правило, носит описательный характер. Она выражается сравнительными категориями и лексическими конструкциями вида "недостаточно полно", "не в полной мере", "с отдельными отступлениями" и т.п. Такая оценка не всегда является достаточным основанием и весомым аргументом для принятия высшим менеджментом или собственниками авиапредприятий решений, связанных с экономической целесообразностью и оправданностью затрат на обеспечение безопасности соответствующих объектов.

В связи с этим в современных условиях возникла объективная необходимость внедрения унифицированного и системного подхода к формированию количественно-качественной оценки состояния безопасности авиапредприятий и их объектов, основанного на единых руководящих документах в этой области.

Разработка методики такой оценки является в настоящее время актуальной многомерной управленческой задачей. При этом важнейшее значение имеет политика аудита, заключающаяся в совокупности подходов и правил, определяющих типы событий, происходящих в контролируемой системе (на защищаемом объекте), подлежащих регистрации и протоколированию (документированию) для последующего квалифицированного анализа. Не менее значимы условия и методы выявления нарушений установленной политики безопасности, анализа признаков атак на защищаемые активы и скрытых (новых) угроз безопасности, анализа уязвимых мест в действующей системе (подсистеме) защиты. Естественно, протоколы событий, называемые свидетельствами аудита, должны быть представлены в пригодном для понимания и анализа формате. Политика аудита зависит от специфики контролируемых объектов и может несколько различаться для разнотипных авиапредприятий, действующих на территории Российской Федерации.

В целом, методика аудита (анализа состояния) информационной безопасности авиапредприятий должна детально определять взаимоотношения и роли сторон, систему используемых критериев, а также этапы проведения аудита и подходы к формированию частных и комплекс-

ных (итоговых) оценок. При этом количественные и качественные оценки состояния системы безопасности авиапредприятий должны базироваться в общем случае на следующих частных составляющих:

- оценке адекватности политики безопасности принятой модели угроз;
- оценке уровня технической оснащенности;
- оценке эффективности организационно-управленческой деятельности.

Указанные частные оценки могут формироваться для всех предметных областей, связанных с комплексной безопасностью авиапредприятий. Поэтому их условно можно называть отраслевыми.

3. Некоторые подходы и принципы формирования оценок состояния системы информационной безопасности авиапредприятия

Рассмотрим некоторые подходы и принципы формирования частных (отраслевых) и комплексных (итоговых) оценок состояния системы информационной безопасности авиапредприятия.

1. В силу присутствия в развитых человеко-машинных комплексах и системах множества несовместимых с каким-либо числовым эквивалентом категорий и объектов, точная абсолютная оценка их состояния, как правило, невозможна. В связи с этим целесообразно оперировать относительной оценкой, характеризующей соответствие элементов системы защиты каким-либо установленным требованиям, например, адекватность действующей политики безопасности принятой модели угроз. При этом в задачу аналитиков может не входить актуализация самой политики безопасности или изменение принятой модели угроз вследствие совершенствования средств и технологий нападения или объективного ослабления со временем защитных механизмов.

2. Оценка может быть линейной или нелинейной, что определяется типом исследуемой системы. Очевидно, что система безопасности авиапредприятия обладает свойствами нелинейной системы. Вместе с тем, математический аппарат и методы исследования линейных систем более простые, а изменение частных составляющих будет линейно влиять на итоговые оценки, что, несомненно, удобно в процессе их возможной корректировки. Поэтому, несмотря на неизбежные погрешности, целесообразно считать исследуемую систему безопасности авиапредприятий линейной и соответственно аппроксимировать используемый математический аппарат с учетом этого соглашения. При этом, если в будущем (с получением практических результатов) указанный подход себя не оправдает, то в этом случае будет оправдана необходимость использования математического аппарата, учитывающего фактор нелинейности.

3. Для формирования частной оценки необходимо определить исследуемые и анализируемые критерии (показатели), а также определить меру, на основании которой можно построить числовую шкалу оценок (рейтинга) безопасности защищаемых объектов. При этом для снижения субъективизма, вносимого в результате аппроксимации математического аппарата и экспертной деятельности аналитиков, принятая шкала оценок должна быть в достаточной степени дифференцированной.

4. Оценки могут быть количественными или качественными. Как правило, количественные оценки показывают обеспеченность объекта необходимыми ресурсами для защиты, а качественные – эффективность их использования. Соотношение количества и качества в определенной степени есть оценка разумной достаточности.

При этом необходимо понимать, что в случае существенного недостатка необходимых материально-технических ресурсов для построения адекватной угрозам системы защиты, даже при исключительно высоком профессионализме специалистов и отличной организации их работы, обеспечить удовлетворительный уровень защищенности объектов практически невозможно. Соответствующая аналогия уместна и в сравнении "наоборот", когда при наличии самых современных и совершенных материально-технических ресурсов их нерациональное или некомпетентное использование критически снижает уровень безопасности защищаемых объектов.

5. Для обработки выбранных критериев (показателей) и формирования частных (отраслевых) и комплексных оценок необходима соответствующая методика. Желательно, чтобы она была универсальной (принцип "единства подходов") и пригодной для исследования различных предметных областей, что обеспечит ее легкую тиражируемость и переносимость на различные направления деятельности. В общем случае, процесс формирования оценки может базироваться на следующих функциях:

функции формального наличия защитных механизмов, которая реализуется на основе отчетных документов и ответов на тестовые задания;

функции соответствия (адекватности) защитных механизмов принятой модели угроз;

функции правильности и эффективности использования защитных механизмов.

Процесс последующей корректировки оценки должен учитывать функцию изменения (совершенствования или ослабления) показателей системы защиты со временем, а также вектор развития системы в положительную или отрицательную сторону.

При этом "адресность" полученных оценок должна обеспечиваться за счет набора критериев, адекватно отражающих специфику конкретного объекта. Для снижения субъективизма (п.п. 2, 3 данного раздела) методика оценки должна быть, по возможности, более простой для использования и понимания, а количество критериев достаточно большим, поскольку, согласно теории вероятности, при постоянном коэффициенте некомпетентности аудитора (вероятности ошибки эксперта) с увеличением ансамбля событий (оцениваемых критериев и параметров) вероятность ошибки при формировании оценки снижается.

6. Полученные частные (отраслевые) оценки должны быть независимы (автономны) друг от друга, что позволит использовать их для анализа отдельных направлений деятельности. При этом комплексная (итоговая) оценка должна быть интегральной, поскольку анализ отдельных факторов не дает целостного представления об исследуемом объекте. Такой подход подразумевает в частности использование при анализе эффективности системы защиты выводов, оценок и результатов проверок, полученных, в том числе, различными федеральными контролирующими органами.

7. При оценке критических объектов необходимо соблюдать принцип "естественной конфиденциальности детализированных выводов", который реализуется посредством однонаправленных (необратимых) преобразований полученных частных результатов, что исключает возможность выявления злоумышленником по итоговой (интегральной) оценке слабых (уязвимых) мест в системе защиты.

Таким образом, знание критериев и методики проведения аудита, а также итоговой оценки состояния системы комплексной безопасности объекта не должно позволять злоумышленнику выявлять уязвимые места в системе защиты без знания полных данных, необходимых для получения частных оценок.

8. При использовании полученных оценок целесообразно соблюдать принцип "ограниченной гласности", согласно которому обнародованию подлежат не все результаты проведенного аудита, что исключает раскрытие уязвимых мест в системе безопасности авиапредприятия, а также деструктивные обсуждения и домыслы (допущения) по этому поводу.

9. В процессе использования методики аудита безопасности необходимо обеспечить возможность корректировки полученных оценок вследствие устранения отмеченных ранее недостатков или выявления новых обстоятельств, связанных с обеспечением безопасности защищаемых объектов. Соответственно корректировка оценки может производиться как в лучшую, так и худшую сторону, в том числе по результатам контрольных мероприятий (проверок) и дистанционного контроля, основанного, например, на различных отчетных документах и тестовых заданиях.

10. При аудите состояния системы безопасности авиапредприятий необходимо учитывать фактор наличия "остаточных рисков", согласно которому даже самые высокие оценки (рейтинги защищенности) не могут гарантировать абсолютную безопасность соответствующих объектов и исключать возможность преодоления системы защиты в случае превосходства интеллектуально-материальных и временных ресурсов со стороны системы нападения.

ЛИТЕРАТУРА

1. Курило А.П., Терентьев А.И. Некоторые подходы и принципы оценки состояния комплексной безопасности организаций банковской системы Российской Федерации // Информационный бюллетень ГУБ и ЗИ Банка России. - 2004. - № 4 (33).

2. Андрианов В.В., Зефирова С.Л., Голованов В.Б., Голдуев Н.А. Обеспечение информационной безопасности бизнеса. - 2-е изд., перераб. и доп. - М.: ЦИПСИР: Альпина Паблишерз, 2011.

CERTAIN PRINCIPLES OF THE ESTIMATION OF INFORMATION SECURITY SYSTEM OF AIRCRAFTS ENTERPRISE

Terentyev A.I.

Certain approaches and principles of formation of private and complex (total) estimations of a condition of system of information security of aviation enterprise are stated.

Key words: information security of aviation enterprise, safety levels.

Сведения об авторе

Терентьев Андрей Иванович, 1965 г.р., окончил МИРЭА (1993), кандидат технических наук, доцент МГТУ ГА, автор более 30 научных работ, область научных интересов – теория преобразования и передачи информации, помехоустойчивое кодирование, криптография, информационная безопасность.

УДК 517.956.223+517.983

О СИСТЕМЕ НЕЛИНЕЙНЫХ ИНТЕГРАЛЬНЫХ УРАВНЕНИЙ ТИПА ГАММЕРШТЕЙНА

Н.А. ЕРЗАКОВА

Для системы уравнений типа Гаммерштейна доказывается существование решения.

Ключевые слова: оператор Гаммерштейна, компактность по мере, мера некомпактности, уплотняющий оператор, частично аддитивный оператор.

Введение

Нелинейные интегральные операторы

$$Ax(t) = \int_{\Omega} K_0(t, s) f[s, x(s)] ds$$

с ядрами специального вида называют операторами Гаммерштейна [1], а их различные обобщения – операторами типа Гаммерштейна. Соответственно уравнения с такими операторами называют уравнениями типа Гаммерштейна [2].

Уравнения вида

$$u = A(u; \lambda)$$

возникают во многих задачах нелинейной механики: отыскание критических нагрузок и форм потери устойчивости упругих систем, исследование автоколебательных процессов, исследования процесса рождения волн в движущейся жидкости и т.д. В подобных задачах роль параметра λ могут играть нагрузки, частоты автоколебаний, скорости движения жидкости [1]. Кроме того, подобные уравнения имеют приложения в физике, астрономии [3].

1. Постановка и формализация задачи

В настоящей работе исследуется разрешимость системы нелинейных интегральных уравнений типа Гаммерштейна

$$u_i(x) = \lambda \sum_{j=1}^n \int_{\Omega} K_{ij}(x, y) g_j(u_1(y), u_2(y), \dots, u_n(y), y) dy + f_i.$$

Пусть Ω - подмножество конечномерного пространства, причем $\mu(\Omega) < \infty$, μ – непрерывная мера. Пусть θ обозначает нулевую функцию ($\theta \equiv 0$ п.в. на Ω).

Рассмотрим пространства Лебега $L_{p,n} = L_p(\Omega)^n$, $1 \leq p \leq \infty$, состоящие из всевозможных наборов из n функций $\bar{u} = (u_1, u_2, \dots, u_n)$, $u_i \in L_p(\Omega)$ для всех $i = 1, 2, \dots, n$.

Зададим норму в $L_{p,n}$ следующим образом $\|\bar{u}\|_{L_{p,n}} = \max_{1 \leq i \leq n} \|u_i\|_{L_p}$.

Сразу заметим, что хотя $L_{p,n}$ – это бесконечномерное пространство, но в силу того, что n конечно и $\bar{u}(x) = (u_1(x), u_2(x), \dots, u_n(x))$ при каждом $x \in \Omega$ – это элемент n -мерного пространства, а в конечномерном пространстве все нормы эквивалентны, ниже следующие результаты справедливы и для всех других норм в $L_{p,n}$, в частности, для

$$\|\bar{u}\|_{L_{p,n}} = \sqrt{\sum_{i=1}^n \|u_i\|_{L_p}^2}.$$

Определение [4]. Пусть E – банахово пространство, а U – ограниченное подмножество E . Мерой некомпактности Хаусдорфа $\chi_E(U)$ множества U называется инфимум всех $\varepsilon > 0$, при которых U имеет в E конечную ε – сеть.

Для относительно компактного множества U справедливо равенство $\chi_E(U) = 0$.

Кроме того, χ_E обладает рядом свойств [4], среди которых полуоднородность

$$\chi_E(tU) = |t| \chi_E(U) \quad (t - \text{число});$$

полуаддитивность

$$\chi_E(U_1 \cup U_2) = \max \{ \chi_E(U_1), \chi_E(U_2) \};$$

инвариантность относительно сдвигов

$$\chi_E(U + b) = \chi_E(U) \quad (b \in E);$$

алгебраическая полуаддитивность

$$\chi_E(U_1 + U_2) \leq \chi_E(U_1) + \chi_E(U_2).$$

Для произвольного подмножества $D \subset \Omega$ и функции $u \in L_p(\Omega)$ обозначим

$$P_D u(s) = \begin{cases} u(s), & s \in D, \\ 0, & s \notin D, \end{cases}$$

соответственно $P_D \bar{u} = (P_D u_1, P_D u_2, \dots, P_D u_n)$.

По аналогии с пространствами $L_p(\Omega)$ при $1 \leq p < \infty$ [5-9] рассмотрим в $L_{p,n}$ меру неравно-
степенной абсолютной непрерывности норм элементов $v_{L_{p,n}}(\bar{U})$ подмножества \bar{U} из $L_{p,n}$, пола-
гая $v_{L_{p,n}}(\bar{U}) = \overline{\lim}_{\mu(D) \rightarrow 0} \sup_{\bar{u} \in \bar{U}} \|P_D \bar{u}\|_{L_{p,n}}$.

Мера $v_{L_{p,n}}(U)$ обладает всеми вышеперечисленными свойствами $\chi_{L_{p,n}}(U)$ и отличается от меры некомпактности Хаусдорфа только тем, что равенство $v_{L_{p,n}}(U) = 0$ возможно на множествах, не являющимися относительно компактными.

Напомним, что банаховы пространства $L_p(\Omega)$, $1 \leq p \leq \infty$ являются идеальными. Следовательно, неравенство $|u(x)| \leq |v(x)|$ п.в., где $v \in L_p(\Omega)$, а u – измеримая функция влечет $u \in L_p(\Omega)$, причем $\|u\|_{L_p} \leq \|v\|_{L_p}$.

Более того, пространства $L_p(\Omega)$, $1 \leq p < \infty$ являются правильными. Это означает, что каждая функция u из $L_p(\Omega)$ имеет абсолютно непрерывную норму, т.е. $\lim_{\mu(D) \rightarrow 0} \|P_D u\|_{L_p} = 0$.

Поэтому, если

$$\bar{u} = (u_1, u_2, \dots, u_n), \bar{v} = (v_1, v_2, \dots, v_n), v_i \in L_p(\Omega), |u_i| \leq |v_i| \text{ п.в. } (i = 1, 2, \dots, n) \Rightarrow \|\bar{u}\|_{L_{p,n}} \leq \|\bar{v}\|_{L_{p,n}} \quad (1)$$

при $1 \leq p \leq \infty$ и $\lim_{\mu(D) \rightarrow 0} \|P_D \bar{u}\|_{L_{p,n}} = 0$ при $1 \leq p < \infty$.

Так же, как в [7-9], назовем подмножества \bar{U}_1, \bar{U}_2 пространства $L_{p,n}$ сравнимыми $U_1 \leq U_2$, если в $L_{p,n}$ существует такой элемент $\bar{b}(x) = (b_1(x), b_2(x), \dots, b_n(x))$, что для произвольного \bar{u} из \bar{U}_1 найдется элемент \bar{v} из \bar{U}_2 , для которого п.в. выполняется неравенство $|u_i(x)| \leq |b_i(x)| + |v_i(x)|$ ($i = 1, 2, \dots, n$).

Аналогично, два оператора \bar{F}_1, \bar{F}_2 , действующие из пространства $L_{p,n}$ в пространство $L_{q,n}$,

назовем сравнимыми $\overline{F}_1 \leq \overline{F}_2$, если для каждого подмножества \overline{U} из $L_{p,n}$ имеет место $\overline{F}_1(\overline{U}) \leq \overline{F}_2(\overline{U})$.

В силу свойства нормы (1) и определения $v_{L_{p,n}}(\overline{U})$ справедливо

$$\overline{U}_1 \leq \overline{U}_2 \Rightarrow v_{L_{p,n}}(\overline{U}_1) \leq v_{L_{p,n}}(\overline{U}_2). \quad (2)$$

Для произвольной функции из $L_p(\Omega)$ и произвольного числа $T > 0$ введем обозначения

$$A(u, T) = \{s : |u(s)| \leq T\}, \quad D(u, T) = \{s : |u(s)| > T\} = \Omega \setminus A(u, T). \quad (3)$$

Аналогично, $A(\overline{u}, T) = \{s : |u_i(s)| \leq T, 1 \leq i \leq n, \overline{u} = (u_1, u_2, \dots, u_n)\}$, $D(\overline{u}, T) = \Omega \setminus A(\overline{u}, T)$.

Лемма [5, 6]. Пусть U – произвольное ограниченное подмножество из $L_p(\Omega)$. Тогда

$$1) \chi_{L_p}(U) \geq v_{L_p}(U);$$

$$2) v_{L_p}(U) = \limsup_{T \rightarrow \infty} \sup_{u \in U} \|P_{D(u, T)} u\|_{L_p};$$

$$3) \text{ если } U \text{ к тому же компактно по мере, то } \chi_{L_p}(U) = v_{L_p}(U).$$

Доказательство. Для множества U и произвольного числа $\varepsilon > 0$, согласно определению $\chi_{L_p}(U)$, найдется конечная $[\chi_{L_p}(U) + \varepsilon]$ -сеть $C_\varepsilon = \{c_1, c_2, \dots, c_m\}$, такая, что $U \subseteq C_\varepsilon + [\chi_{L_p}(U) + \varepsilon]B$, где B - единичный шар $L_p(\Omega)$. Поэтому по определению $v_{L_p}(U)$ получим $v_{L_p}(U) \leq v_{L_p}(C_\varepsilon + [\chi_{L_p}(U) + \varepsilon]B) \leq \chi_{L_p}(U) + \varepsilon$ в силу обычных свойств нормы, а также свойства абсолютной непрерывности нормы в правильном пространстве $L_p(\Omega)$. Учитывая произвольность выбора $\varepsilon > 0$, получим утверждение 1), т.е. $\chi_{L_p}(U) \geq v_{L_p}(U)$.

Утверждения 2) и 3) докажем одновременно. Имеем

$$U \subseteq \{P_{A(u, T)} u + TP_{D(u, T)} \operatorname{sgn} u : u \in U\} + \{P_{D(u, T)}(u - T \operatorname{sgn} u) : u \in U\}.$$

В силу определения v_{L_p} и абсолютной непрерывности нормы в правильном пространстве $L_p(\Omega)$ для любого $T > 0$ имеем $v_{L_p}\{P_{A(u, T)} u + TP_{D(u, T)} \operatorname{sgn} u : u \in U\} = 0$ и в случае компактности по мере U в силу критерия компактности [1] также

$$\chi_{L_p}\{P_{A(u, T)} u + TP_{D(u, T)} \operatorname{sgn} u : u \in U\} = 0.$$

Из аддитивности и алгебраической полуаддитивности $v_{L_p}(U)$, $\chi_{L_p}(U)$ следует, что

$$v_{L_p}(U) \leq v_{L_p}\{P_{D(u, T)}(u - T \operatorname{sgn} u) : u \in U\} \leq v_{L_p}\{P_{D(u, T)} u : u \in U\};$$

$$\chi_{L_p}(U) \leq \chi_{L_p}\{P_{D(u, T)}(u - T \operatorname{sgn} u) : u \in U\} \leq \sup_{u \in U} \|P_{D(u, T)} u\|_{L_p}$$

для любого $T > 0$, фиксированного на момент рассуждений. Устремляя $T \rightarrow \infty$, получим

$$\chi_{L_p}(U) \leq \limsup_{T \rightarrow \infty} \sup_{u \in U} \|P_{D(u, T)} u\|_{L_p}$$

для компактного по мере U и $v_{L_p}(U) \leq \limsup_{T \rightarrow \infty} \sup_{u \in U} \|P_{D(u, T)} u\|_{L_p}$ для любого U . Для ограниченного множества U в $L_p(\Omega)$ имеет место $\limsup_{T \rightarrow \infty} \mu[D(u, T)] = 0$. Поэтому

$$v_{L_p}(U) \leq \limsup_{T \rightarrow \infty} \sup_{u \in U} \|P_{D(u, T)} u\|_{L_p} \leq \overline{\lim}_{\mu(D) \rightarrow 0} \sup_{u \in U} \|P_D u\|_{L_p} = v_{L_p}(U).$$

Равенство 2) доказано. Из 1) и 2) и $\chi_{L_p}(U) \leq \limsup_{T \rightarrow \infty} \sup_{u \in U} \|P_{D(u, T)} u\|_{L_p}$ получим утверждение 3).

Замечание. Рассмотрим для множества \bar{U} из $L_{p,n} = L_p(\Omega)^n$ соответствующие множества $U_i = \{u_i : \bar{u} = (u_1, u_2, \dots, u_n) \in \bar{U}\}$ для всех $1 \leq i \leq n$ из $L_p(\Omega)$.

Так как $\chi_{L_{p,n}}(\bar{U}) = \max_{1 \leq i \leq n} \chi_{L_p}(U_i)$ и

$$v_{L_{p,n}}(\bar{U}) = \overline{\lim}_{\mu(D) \rightarrow 0} \sup_{\bar{u} \in \bar{U}} \|P_D \bar{u}\|_{L_{p,n}} = \overline{\lim}_{\mu(D) \rightarrow 0} \sup_{\bar{u} \in \bar{U}} \max_{1 \leq i \leq n} \|P_D u_i\|_{L_p} = \max_{1 \leq i \leq n} v_{L_p}(U_i),$$

1) $\chi_{L_{p,n}}(\bar{U}) \geq v_{L_{p,n}}(\bar{U});$

2) $v_{L_{p,n}}(\bar{U}) = \lim_{T \rightarrow \infty} \sup_{\bar{u} \in \bar{U}} \|P_{D(\bar{u}, T)} \bar{u}\|_{L_{p,n}};$

3) если \bar{U} к тому же компактно по мере, то $\chi_{L_{p,n}}(\bar{U}) = v_{L_{p,n}}(\bar{U}).$

Здесь компактность по мере \bar{U} означает, что компактны по мере множества $U_i = \{u_i : \bar{u} = (u_1, u_2, \dots, u_n) \in \bar{U}\}$ для всех $1 \leq i \leq n$. Другими словами, предполагается компактность U_i для всех $1 \leq i \leq n$ в нормированном пространстве S всех измеримых почти всюду конечных функций u с нормой

$$\|u\| = \inf_{s > 0} \{s + \mu\{t : |u(t)| \geq s\}\}.$$

Пусть заданы линейные ограниченные интегральные операторы $K_{ij} : L_p(\Omega) \rightarrow L_q(\Omega)$, $i = 1, 2, \dots, n$, $j = 1, 2, \dots, n$. Пусть оператор $\bar{K} : L_{p,n} \rightarrow L_{q,n}$ действует на $\bar{u} = (u_1, u_2, \dots, u_n)$, $u_i \in L_p(\Omega)$ для всех $i = 1, 2, \dots, n$ следующим образом

$$\bar{K}\bar{u} = \left(\sum_{j=1}^n \int_{\Omega} K_{1j}(x, y) u_j(y) dy, \sum_{j=1}^n \int_{\Omega} K_{2j}(x, y) u_j(y) dy, \dots, \sum_{j=1}^n \int_{\Omega} K_{nj}(x, y) u_j(y) dy \right).$$

Очевидно, $\|\bar{K}\| = \max_{1 \leq i, j \leq n} \|K_{ij}\|.$

Оператор \bar{K} компактен по мере, если компактны по мере все множества $K_{ij}(U_i)$ для всех $1 \leq i \leq n$, $1 \leq j \leq n$, где $U_i = \{u_i : \bar{u} = (u_1, u_2, \dots, u_n) \in \bar{U}\}$, \bar{U} – произвольное ограниченное подмножество $L_{p,n}$.

Пусть заданы $g_i(x, u_1, u_2, \dots, u_n)$ – вещественные функции, порождающие непрерывный оператор Немыцкого $\bar{h} = (h_1, h_2, \dots, h_n)$ из пространства Лебега $L_{q,n}$ в пространство $L_{p,n}$, ($1 \leq p < q < \infty$), $h_i \bar{u} = g_i(x, u_1, u_2, \dots, u_n)$ (нелинейный оператор суперпозиции [2]), причем

$$|g_i(x, u_1, u_2, \dots, u_n)| \leq a_i(x) + b \sum_{k=1}^n |u_k(x)|^{q/p}, \quad a_i \in L_p(\Omega), \quad b > 0.$$

Замечание. Так как из неравенства $|h_i \bar{u}(x)| \leq |v(x)|$ почти всюду следует $\|h_i \bar{u}\|_{L_p} \leq \|v\|_{L_p}$, то при всех $1 \leq i \leq n$ имеем

$$\begin{aligned} \|h_i \bar{u}\|_{L_p} &= \left(\int_{\Omega} |g_i(x, u_1(x), u_2(x), \dots, u_n(x))|^p dx \right)^{1/p} \leq \left\| a_i + b \sum_{k=1}^n |u_k|^{q/p} \right\|_{L_p} \leq \\ &\leq \|a_i\|_{L_p} + b \sum_{k=1}^n \| |u_k|^{q/p} \|_{L_p} \leq \|a_i\|_{L_p} + b \sum_{k=1}^n \|u_k\|_{L_q}^{q/p}. \end{aligned}$$

Определение [4]. Пусть E, E_1 – банаховы пространства. Непрерывный оператор $A: G \rightarrow E_1$, где $G \subset E$, называется χ – уплотняющим, если для любого ограниченного подмножества $U \subset G$, замыкание которого не компактно, выполняется неравенство

$$\chi_{E_1}(AU) < \chi_E(U).$$

Уплотняющие операторы, являющиеся обобщением компактных и сжимающих операторов, во многих случаях сохраняют свойства последних.

Непрерывный оператор A называется (k, χ) – ограниченным, если для любого $U \subset G$ выполняется неравенство $\chi_{E_1}(AU) \leq k\chi_E(U)$.

2. Основные теоремы

Теорема 1. Пусть непрерывный оператор \bar{K} действует из $L_{p,n}$ в $L_{q,n}$ ($1 \leq p < q < \infty$). Пусть, кроме того, \bar{K} компактен по мере. Пусть $\bar{h}: L_{q,n} \rightarrow L_{p,n}$ – оператор суперпозиции. Тогда оператор $\bar{K}\bar{h}$ является (k, χ) – ограниченным на шаре $\bar{B}(\bar{u}_1, r) = \left\{ \bar{u} \in L_{q,n} : \|\bar{u} - \bar{u}_1\|_{L_{q,n}} \leq r \right\}$ (\bar{u}_1 и r произвольные) с константой $k = nb \max_{1 \leq i, j \leq n} \|K_{ij}\| r^{\frac{q-1}{p}}$, где константа $b > 0$ зависит только от \bar{h} .

Доказательство. Пусть \bar{U} – подмножество из $\bar{B}(\bar{u}_1, r) = \left\{ \bar{u} \in L_{q,n} : \|\bar{u} - \bar{u}_1\|_{L_{q,n}} \leq r \right\}$, т.е. $\bar{U} \subseteq \bar{u}_1 + r\bar{B}$, где \bar{B} – единичный шар $L_{q,n}$.

В силу частичной аддитивности оператора \bar{h} и линейности оператора \bar{K} для произвольных $\bar{u} \in \bar{U}$ и $T > 0$ имеем

$$\bar{h}\bar{u} = \bar{h}(P_{D(\bar{u}, T)}\bar{u} + P_{A(\bar{u}, T)}\bar{u}) = \bar{h}(P_{D(\bar{u}, T)}\bar{u}) + \bar{h}(P_{A(\bar{u}, T)}\bar{u}) - \bar{h}(\theta).$$

Пусть

$$P_{A(\bar{u}, T)}\bar{U} = \{P_{A(\bar{u}, T)}\bar{u} : \bar{u} \in \bar{U}\}, \quad P_{D(\bar{u}, T)}\bar{U} = \{P_{D(\bar{u}, T)}\bar{u} : \bar{u} \in \bar{U}\};$$

$$P_{A(\bar{u}, T)}\bar{V} \leq \{\bar{v} = (v_1, v_2, \dots, v_n), v_i = \sum_{k=1}^n |P_{A(\bar{u}, T)}u_k(x)|^{q/p}, 1 \leq i \leq n : \bar{u} \in \bar{U}\};$$

$$P_{D(\bar{u}, T)}\bar{V} \leq \{\bar{v} = (v_1, v_2, \dots, v_n), v_i = \sum_{k=1}^n |P_{D(\bar{u}, T)}u_k(x)|^{q/p}, 1 \leq i \leq n : \bar{u} \in \bar{U}\}.$$

Так как

$$|h_i(P_{A(\bar{u}, T)}\bar{u})| = |g_i(x, P_{A(\bar{u}, T)}u_1, P_{A(\bar{u}, T)}u_2, \dots, P_{A(\bar{u}, T)}u_n)| \leq a_i(x) + b \sum_{k=1}^n |P_{A(\bar{u}, T)}u_k(x)|^{q/p},$$

то $\bar{h}(P_{A(\bar{u}, T)}\bar{U}) \leq \bar{a} + bP_{A(\bar{u}, T)}\bar{V}$, где $\bar{a} = (a_1, a_2, \dots, a_n)$, следовательно, в силу (2)

$$\nu_{L_{p,n}}(\bar{h}(P_{A(\bar{u}, T)}\bar{U})) \leq \nu_{L_{p,n}}(\bar{a} + bP_{A(\bar{u}, T)}\bar{V}) = 0.$$

Поэтому

$$\nu_{L_{p,n}}(\bar{h}(\bar{U})) \leq \nu_{L_{p,n}}(\bar{a} + bP_{D(\bar{u}, T)}\bar{V}) \leq nb\nu_{L_{p,n}}^{q/p}(\bar{U}).$$

Из предположения о компактности по мере оператора \bar{K} следует компактность по мере множества $\bar{K}\bar{h}\bar{U}$, откуда $\chi_{L_{q,n}}(\bar{K}\bar{h}\bar{U}) = \nu_{L_{q,n}}(\bar{K}\bar{h}\bar{U})$.

Поскольку

$$v_{L_{q,n}}(\overline{KhU}) = \overline{\lim}_{\mu(D) \rightarrow 0} \sup_{\bar{u} \in \bar{U}} \|P_D \overline{KhU}\|_{L_{q,n}} = \overline{\lim}_{\mu(D) \rightarrow 0} \sup_{\bar{u} \in \bar{U}} \max_{1 \leq i \leq n} \left\| P_D \left(\sum_{j=1}^n \int_{\Omega} K_{ij}(x, y) h_j(\bar{u}(y)) dy \right) \right\|_{L_q} \leq \|\overline{K}\|_{v_{L_{p,n}}}(\bar{hU}),$$

имеем

$$\chi_{L_{q,n}}(\overline{KhU}) \leq \|\overline{K}\|_{v_{L_{p,n}}}(\bar{hU}) \leq nb \|\overline{K}\|_{v_{L_{p,n}}^{\frac{q}{p}}}(\bar{U}) \leq nb \|\overline{K}\|_{v_{L_{p,n}}^{\frac{q-1}{p}}}(\bar{U}) v_{L_{p,n}}(\bar{U}) \leq nb \max_{1 \leq i, j \leq n} \|K_{ij}\|_{v_{L_{p,n}}^{\frac{q-1}{p}}}(\bar{U}) \chi_{L_{p,n}}(\bar{U}),$$

$$v_{L_{p,n}}^{\frac{q-1}{p}}(\bar{U}) \leq r^{\frac{q-1}{p}}$$

для любого $\bar{U} \subseteq \bar{u}_1 + r\bar{B}$, откуда получаем утверждение теоремы.

Рассмотрим систему нелинейных интегральных уравнений ($i = 1, \dots, n$)

$$u_i(x) = \lambda \sum_{j=1}^n \int_{\Omega} K_{ij}(x, y) g_i(u_1(y), u_2(y), \dots, u_n(y), y) dy + f_i \quad (4)$$

или

$$\bar{u} = \lambda \overline{Kh} \bar{u} + \bar{f}.$$

Теорема 2. Пусть непрерывный оператор \overline{K} действует из $L_{p,n}$ в $L_{q,n}$ ($1 \leq p < q < \infty$). Пусть, кроме того, \overline{K} компактен по мере. Пусть $\bar{h}: L_{q,n} \rightarrow L_{p,n}$ – оператор суперпозиции. Тогда существуют шар $B(\bar{f}, r) = \left\{ \bar{u} \in L_{q,n} : \|\bar{u} - \bar{f}\|_{L_{q,n}} \leq r \right\}$ и число $\lambda_0 > 0$, при которых система уравнений (4) имеет хотя бы одно решение в шаре $B(\bar{f}, r)$ при всех $|\lambda| \leq \lambda_0$.

Доказательство. Так как предположения теоремы 1 выполнены, то оператор \overline{Kh} является (k, χ) –ограниченным, в частности, на шаре

$$B(\bar{f}, r) = \left\{ \bar{u} \in L_{q,n} : \|\bar{u} - \bar{f}\|_{L_{q,n}} \leq r \right\}$$

с константой $k = nb \max_{1 \leq i, j \leq n} \|K_{ij}\|_{v_{L_{p,n}}^{\frac{q-1}{p}}}$, где константа $b > 0$ зависит только от \bar{h} . Поэтому для всех r ,

таких, что $nb \max_{1 \leq i, j \leq n} \|K_{ij}\|_{v_{L_{p,n}}^{\frac{q-1}{p}}} r^{\frac{q-1}{p}} < 1$ или $r < \frac{1}{\left(nb \max_{1 \leq i, j \leq n} \|K_{ij}\|_{v_{L_{p,n}}^{\frac{q-1}{p}}} \right)^{\frac{p}{q-p}}}$, оператор \overline{Kh} – уплотняющий оператор.

В силу ограниченности операторов \overline{K} и \bar{h} найдется такое $R > 0$, что $\overline{Kh} B(\bar{f}, r) \subseteq B(\theta, R)$

для всех $r < \frac{1}{\left(nb \max_{1 \leq i, j \leq n} \|K_{ij}\|_{v_{L_{p,n}}^{\frac{q-1}{p}}} \right)^{\frac{p}{q-p}}}$.

Если $\lambda < \min \left\{ \frac{r}{R}, 1 \right\}$, то получим χ – уплотняющий оператор $\lambda \overline{Kh}$ на $B(\bar{f}, r)$. В силу ин-

вариантности относительно сдвигов меры некомпактности Хаусдорфа оператор $\lambda \overline{Kh} + \bar{f}$ также χ – уплотняющий оператор на $B(\bar{f}, r)$.

Так как

$$\|\lambda \overline{Kh} \bar{u} + \bar{f} - \bar{f}\|_{L_{q,n}} = \lambda \|\overline{Kh} \bar{u}\|_{L_{q,n}} \leq \lambda R \leq R \frac{r}{R} \leq r$$

для всех $\bar{u} \in B(\bar{f}, r)$, то для выбранных $\lambda > 0$ шар $B(\bar{f}, r)$ отображается в себя. Применим к оператору $\lambda \bar{K} \bar{h} + \bar{f}$ обобщение теоремы Шаудера о существовании неподвижной точки на \mathcal{X} – уплотняющие операторы [4], получим существование решения в пространстве $L_{q,n}$ системы уравнений (4).

В частности, если $\bar{K} \bar{h} \theta = \theta$, то существует ненулевое решение.
Теорема доказана.

Заключение

Разрешимость системы нелинейных интегральных уравнений типа Гаммерштейна исследована впервые. В [7-9] изучалась разрешимость одного уравнения типа Гаммерштейна как в пространствах Лебега, так и в более общих пространствах. В настоящей работе используется понятие сравнимых множеств и сравнимых операторов для пространств $L_{p,n}$.

ЛИТЕРАТУРА

1. Красносельский М.А., Забрейко П.П. и др. Интегральные операторы в пространствах суммируемых функций. - М.: Наука, 1966.
2. Вайнберг М.М. Вариационный метод и метод монотонных операторов в теории нелинейных уравнений. - М.: Наука, 1972.
3. Evans W.D., Harris D.J. Sobolev embedding for generalized ridged domains // Proc. London Math. Soc. 1987 V. 54. - № 1.
4. Ахмеров Р.Р. и др. Меры некомпактности и уплотняющие операторы. – Новосибирск: Наука, 1986.
5. Yezakova N.A. On Measures of Non-Compactness in Regular Spaces // Zeitschrift für Analysis und ihre Anwendungen. 1996. V. 15. № 2.
6. Ерзакова Н.А. Компактность по мере и мера некомпактности // Сибирский математический журнал. - 1997. - Т. 38. - № 5.
7. Ерзакова Н.А. О нелинейных операторах // Научный Вестник МГТУ ГА. - 2009. - № 140.
8. Ерзакова Н.А. О разрешимости нелинейного уравнения // Научный Вестник МГТУ ГА. - 2009. - № 145.
9. Ерзакова Н.А. О разрешимости уравнений с частично аддитивными операторами // Функциональный анализ и его приложения 44:3. - 2010.

ON SYSTEM OF NONLINEAR INTEGRAL EQUATIONS OF GAMMERSTEIN TYPE

Erzakova N.A.

Existence of a solution of the system of equations of Gammerstein type is proved.

Key words: Gammerstein operator, compactness in a measure, measure of noncompactness, condensing map, partially additive operator.

Сведения об авторе

Ерзакова Нина Александровна, окончила Новосибирский государственный университет (1976), доктор физико-математических наук, профессор, автор более 60 научных работ, область научных интересов – теория неподвижных точек, меры некомпактности, уплотняющие операторы, интегральные операторы, пространства С.Л. Соболева, краевые задачи для уравнений с частными производными.

УДК 517.988.8

О МИНИМИЗАЦИИ ВЫПУКЛЫХ ФУНКЦИОНАЛОВ

А.А. ФОНАРЁВ

Статья представлена доктором технических наук, профессором Кузнецовым В.Л.

С использованием проекционного итерационного процесса, сочетающего в себе проекционный метод и итерационный процесс, строится последовательность, минимизирующая выпуклый функционал, заданный в вещественном нормированном пространстве.

Ключевые слова: проекционный итерационный процесс, выпуклый функционал, минимизация

Введение

В теории экстремальных задач, в разных областях математики: в выпуклом программировании и классическом вариационном исчислении, в математической физике, теории целых функций, математической статистике и т. д. широко применяются понятия и методы выпуклого анализа. Формирование выпуклого анализа как самостоятельного раздела относится к 50-60 гг. XX века. Многие понятия и концепции выпуклого анализа нашли своё завершение в функциональном анализе.

С работы В.Фенхеля [1] начался этап выпуклого анализа, на котором детально исследовались свойства выпуклых функционалов.

Важным направлением выпуклого анализа является построение последовательностей, минимизирующих выпуклые функционалы. В частности, в выпуклом анализе рассматривается задача о минимизации выпуклого функционала, заданного в нормированном пространстве X и дифференцируемого на плотном в X подпространстве пространства X .

В статье с использованием проекционного итерационного процесса (ПИП), сочетающего в себе проекционный метод и итерационный процесс, строится последовательность, минимизирующая выпуклый функционал

$$f: X \rightarrow R^1, \quad (1)$$

где X – вещественное нормированное пространство с нормой $\|x\|$ для $x \in X$ и R^1 – одномерное евклидово пространство.

При построении последовательности, минимизирующей функционал (1), предполагается, что функционал f ограничен снизу на пространстве X и дифференцируем на линейном многообразии $E \subset X$, плотном в X . При этом в линейном многообразии E используется норма, которая может не совпадать с нормой пространства X . В основном результате работы предполагается, что функционал f непрерывен на пространстве X .

1. Постановка и формализация задачи

В статье используется терминология из [2].

Предположим, что в линейном многообразии $E \subset X$, плотном в нормированном пространстве X , задана норма $\|x\|'$ для $x \in E$. И предположим, что задана такая последовательность линейных многообразий $E_i \subseteq E$ ($i=1,2,\dots$), что $E_i \subseteq E_{i+1}$ для каждого $i \geq 1$ и для любого $z \in E$ существует такая последовательность $z_i \in E_i$ ($i=1,2,\dots$), что $\|z_i - z\|' \rightarrow 0$ при $i \rightarrow \infty$.

Рассматривая линейное многообразие E , как нормированное пространство с нормой $\|\cdot\|'$, будем называть его пространством E с нормой $\|\cdot\|'$.

Пусть E^* – сопряженное с пространством E с нормой $\|\cdot\|'$ пространство с нормой

$$\|y\|^* \equiv \sup_{x \in E, \|x\|' \leq 1} \langle y, x \rangle \quad (y \in E^*).$$

Здесь и далее $\langle y, x \rangle$ – значение линейного ограниченного функционала $y \in E^*$ на элементе $x \in E$.

Предположим, что для функционала (1) выполняются следующие условия:

1) функционал f ограничен снизу на X , т. е. существует

$$d \equiv \inf_{x \in X} f(x) \in R^1;$$

2) функционал f дифференцируем по Гато на пространстве E с нормой $\|\cdot\|'$ и для его градиента

$$\nabla f(x) \in E^* \quad (x \in E)$$

имеем неравенство

$$\langle \nabla f(u) - \nabla f(v), u - v \rangle \leq M \left(\max(\|u\|', \|v\|') \right) (\|u - v\|')^\alpha \quad (u, v \in E)$$

с постоянной $\alpha > 1$ и неубывающей неотрицательной функцией $M(t)$, заданной для $t \geq 0$;

3) $F_i: E_i \rightarrow E^*$ ($i \geq 1$) – такая последовательность операторов, что при всяком $i \geq 1$ для каждого $u \in E_i$ норма сужения функционала $F_i u - \nabla f(u) \in E^*$ на E_{i+1} (т.е. $\sup_{v \in E_{i+1}, \|v\|' \leq 1} \langle F_i u - \nabla f(u), v \rangle$) не

превосходит $L(\|u\|') \delta_i$, где $\{\delta_i\}_{i=1}^\infty$ – последовательность неотрицательных чисел, сходящаяся к нулю в R^1 , а $L(t)$ – неубывающая неотрицательная функция, заданная для $t \geq 0$;

4) $\{r_i\}_{i=1}^\infty$ и $\{\tau_i\}_{i=1}^\infty$ – такие последовательности положительных чисел, что $\tau_i \leq 1$ и $r_i + \tau_i r_{i+1} \leq r_{i+1}$ для $i \geq 1$,

$$r_i \rightarrow \infty, \quad \tau_i \rightarrow 0, \quad \gamma_i \equiv L(r_i) r_{i+1} \delta_i + M(r_{i+1}) r_{i+1}^\alpha \tau_i^{\alpha-1} \rightarrow 0 \quad (i \rightarrow \infty)$$

и ряд $\sum_{i=1}^\infty \tau_i$ расходится.

Замечание 1. Если для градиента функционала f имеем неравенство

$$\|\nabla f(u) - \nabla f(v)\|^* \leq M \left(\max(\|u\|', \|v\|') \right) \|u - v\|' \quad (u, v \in E),$$

где $M(t)$ – неубывающая неотрицательная функция, заданная для $t \geq 0$, то выполняется неравенство из условия 2) с постоянной $\alpha = 2$.

С использованием обобщённой формулы Лагранжа [2] и условия 2) имеем неравенство

$$f(u) - f(v) \geq \langle \nabla f(u), u - v \rangle - M \left(\max(\|u\|', \|v\|') \right) (\|u - v\|')^\alpha \quad (u, v \in E).$$

Из выпуклости функционала f вытекает [2], что градиент функционала f является монотонным оператором на пространстве E с нормой $\|\cdot\|'$, т.е.

$$\langle \nabla f(u) - \nabla f(v), u - v \rangle \geq 0 \quad (u, v \in E),$$

и что выполняется неравенство

$$f(v) - f(u) \geq \langle \nabla f(u), v - u \rangle \quad (v, u \in E).$$

Следовательно, имеем неравенства

$$\langle \nabla f(u), u - v \rangle \geq f(u) - f(v) \geq \langle \nabla f(u), u - v \rangle - M \left(\max(\|u\|', \|v\|') \right) (\|u - v\|')^\alpha \quad (u, v \in E),$$

с использованием которых доказывается непрерывность функционала f на пространстве E с нормой $\|\cdot\|'$.

Из условия 1) вытекает, что существует такое число $d' \geq d$, что

$$d' \equiv \inf_{x \in E} f(x).$$

Пусть q — произвольное число из интервала $(0, 1)$. И пусть

$$D_i = \left\{ x \in E_i : \|x\|' \leq r_i \right\} \quad (i \geq 1).$$

Рассмотрим последовательность $\{u_i\}_{i=1}^\infty$ ПИП

$$u_{i+1} = u_i - t_i v_i \quad (i \geq 1) \tag{2}$$

с произвольным начальным элементом $u_1 \in D_1$, где если

$$b_i \equiv \sup_{u \in D_{i+1}} \langle F_i u_i, u \rangle > \gamma_i, \tag{3}$$

то $v_i \in D_{i+1}$, $\langle F_i u_i, v_i \rangle \geq b_i - q(b_i - \gamma_i)$ и $t_i = \tau_i$, а если $b_i \leq \gamma_i$, то $v_i = 0$ и $t_i = 0$.

Итерационный процесс (2) является ПИП, совмещающим в себе проекционный метод и итерационный процесс, ибо $u_i \in E_i$ для каждого $i \geq 1$.

Отметим, что на каждом шаге ПИП (2) рассматривается вспомогательная экстремальная задача, т. е. задача о нахождении b_i в (3).

При исследовании ПИП (2) покажем, что при наличии непрерывности функционала f на пространстве E последовательность $\{u_i\}_{i=1}^\infty$ ПИП (2) минимизирует функционал f на пространстве X , т.е.

$$f(u_i) \rightarrow d \quad (i \rightarrow \infty).$$

2. Основной результат

Предварительно приведём три леммы.

Лемма 1. Для последовательности $\{u_i\}_{i=1}^\infty$ ПИП (2) имеем:

1) последовательность $\{u_i\}_{i=1}^\infty$ релаксационная, т. е.

$$f(u_i) \geq f(u_{i+1}) \quad (i \geq 1);$$

2) нижний предел последовательности $\{b_i\}_{i=1}^\infty$ (см. b_i в (3)) равен нулю, т. е.

$$\liminf_{i \rightarrow \infty} b_i = 0.$$

Доказательство. Для каждого $i \geq 1$ имеем неравенства

$$f(u_i) - f(u_{i+1}) \geq \tau_i \langle \nabla f(u_i), v_i \rangle - M \left(\max(r_i, r_i + \tau_i r_{i+1}) \right) (\tau_i r_{i+1})^\alpha \geq$$

$$\geq \tau_i \left((b_i - q(b_i - \gamma_i)) - L(r_i)r_{i+1}\delta_i \right) - M(r_{i+1})\tau_i^\alpha r_{i+1}^\alpha.$$

Следовательно, для каждого $i \geq 1$ имеем неравенство

$$f(u_i) - f(u_{i+1}) \geq \theta_i,$$

где $\theta_i = (1-q)(b_i - \gamma_i)\tau_i$ при $b_i > \gamma_i$, $\theta_i = 0$ при $b_i \leq \gamma_i$.

Из неравенств

$$f(u_1) - f(u_{i+1}) \geq \sum_{j=1}^i \theta_j \quad (i \geq 1)$$

следует, что ряд

$$\sum_{i=1}^{\infty} \theta_i \quad (4)$$

сходится. А если предположим, что $\liminf_{i \rightarrow \infty} b_i \geq b > 0$, то получим расходимость ряда (4). Лемма 1 доказана.

Лемма 2. Последовательность $\{u_i\}_{i=1}^{\infty}$ ПИП (2) минимизирует функционал f на пространстве E с нормой $\|\cdot\|$, т. е. $f(u_i) \rightarrow d'$ ($i \rightarrow \infty$).

Доказательство. В силу заключения 2) леммы 1 существует такая подпоследовательность $\{u_{i_j}\}_{j=1}^{\infty}$ последовательности $\{u_i\}_{i=1}^{\infty}$ ПИП (2), что

$$\lim_{j \rightarrow \infty} b_{i_j} \leq 0.$$

Пусть $w_j = u_{i_j}$, $Q_j = D_{i_j}$ ($j \geq 1$). В силу $Q_j \subseteq Q_{j+1}$ для $j \geq 1$ существует такая последовательность $z_j \in Q_j$ ($j \geq 1$), что $f(z_j) \rightarrow d'$ при $j \rightarrow \infty$. Имеем $\limsup_{j \rightarrow \infty} \alpha_j \leq 0$, где $\alpha_j = \langle \nabla f(w_j), w_j - z_j \rangle$ ($j \geq 1$). И т. к. $f(w_j) \leq f(z_j) + \alpha_j$ для всех $j \geq 1$, то $f(w_j) \rightarrow d'$ при $j \rightarrow \infty$, что в силу заключения 1) леммы 1 влечёт $f(u_i) \rightarrow d'$ при $i \rightarrow \infty$. Лемма 2 доказана.

Доказательства лемм 1, 2 аналогичны доказательствам предложений 1, 2 в [3].

Аналоги лемм 1 и 2 имеются в работе [4] без доказательств в качестве леммы 1 и теоремы 1 работы [4].

Лемма 3. Если функционал f непрерывен на пространстве X , то $d' = d$.

Лемма 3 доказывается с использованием доказательства от противного.

Из лемм 2, 3 вытекает следующая теорема.

Теорема 1. Если функционал f непрерывен на пространстве X , то последовательность $\{u_i\}_{i=1}^{\infty}$ ПИП (2) минимизирует функционал f на пространстве X .

Теорема 1 является основным результатом статьи.

Замечание 2. Лемма 2 и теорема 1 остаются справедливыми, если в них заменить непрерывность функционала f на пространстве X на такое условие: из сходимости последовательности $\{x_n\}_{n=1}^{\infty} \subset X$ к $x \in X$ в пространстве X , т. е. из $\|x_n - x\| \rightarrow 0$ ($n \rightarrow \infty$), следует, что $\liminf_{n \rightarrow \infty} f(x_n) \leq f(x)$.

Работа выполнена при поддержке АВЦП «Развитие научного потенциала высшей школы».

Заключение

Автором предложен проекционный итерационный процесс, сочетающий в себе проекционный метод и итерационный процесс, последовательность которого минимизирует выпуклый функционал, заданный в вещественном нормированном пространстве X и дифференцируемый на линейном многообразии $E \subset X$, плотном в пространстве X .

ЛИТЕРАТУРА

1. Fenchel W. On conjugate convex functional // Canad. J. Math. Soc. – 1949. – V. 1. – No. 1.
2. Вайнберг М.М. Вариационный метод и метод монотонных операторов в теории нелинейных уравнений. - М.: Наука, 1972.
3. Фонарёв А.А. Об одном новом методе решения вариационных неравенств // Изв. вузов. Математика. - 1988. - № 11.
4. Фонарёв А.А. Об одном методе решения задачи о минимальных поверхностях // Аналитические и конструктивные методы исследования дифференциальных уравнений: сб. научн. тр. - Иркутск: Иркутский государственный университет, 1993.

ABOUT MINIMIZATION OF CONVEX FUNCTIONALS

Fonarev A.A.

The author builds sequence of a projective iterative process which minimizes the convex functional set in real normalized space. Projective iterative process combines a projective method and iterative process.

Key words: projective iterative process, convex functional, minimization.

Сведения об авторе

Фонарёв Анатолий Афанасьевич, 1942 г.р., окончил МГУ (1972), кандидат физико-математических наук, доцент кафедры высшей математики МФТИ, автор более 110 научных работ, область научных интересов – нелинейные уравнения в нормированных пространствах, приближенные методы нелинейного функционального анализа.

КРАТКИЕ СООБЩЕНИЯ

УДК 537.874

МАТЕМАТИЧЕСКАЯ МОДЕЛЬ ОТРАЖЕНИЯ ЭЛЕКТРОМАГНИТНОГО ПОЛЯ ОТ ГОФРИРОВАННОЙ ПОВЕРХНОСТИ С УВЛАЖНЕННЫМ ВЕРХНИМ СЛОЕМ

С.В. БОЯРКИН

Статья представлена доктором технических наук, профессором Кузнецовым В.Л.

Построена математическая модель взаимодействия электромагнитного излучения с периодически возмущенной, увлажненной поверхностью почвы. В основу построения модели положен метод инвариантного погружения.

Ключевые слова: отражение электромагнитных волн, метод инвариантного погружения, гофрированная поверхность, увлажненный слой почвы.

Введение

Дистанционные методы зондирования земной поверхности дают возможность получить интегральные характеристики исследуемого объекта без непосредственного контакта с ним, используя в качестве посредника электромагнитное поле. Получаемая информация «закодирована» в амплитудно-фазовых соотношениях спектра отраженного поля. Ее «декодирование» составляет суть обратной задачи теории рассеяния для сложных геофизических объектов. Очевидно, что для решения этой фундаментальной задачи необходимо понимание особенностей процессов взаимодействия поля с изучаемым объектом, выявление связей интересующих параметров с характеристиками отраженного поля. Для организации сбора и обработки используются математические модели взаимодействия излучения с геофизическими объектами, позволяющие оперативно создавать базы данных для характеристик отраженных сигналов при варьировании различных параметров исследуемой среды.

Настоящая работа посвящена разработке такой модели для исследования содержания влаги в почве, что является актуальным, например, для задач сельского хозяйства. В ее основу положен метод инвариантного погружения [1, 2], хорошо зарекомендовавший себя при решении краевых задач прикладной электродинамики [2, 3]. Основная идея метода состоит в том, чтобы выделить в задаче некоторый параметр, называемый параметром погружения, и рассмотреть множество решений при возможных значениях этого параметра. Если какие-либо два значения параметра близки, то мы вправе ожидать и близость (малые отличия) соответствующих решений. Это означает, что возможно построение дифференциальных уравнений эволюции решений исследуемой задачи при изменении значения параметра, играющего роль времени. Такие уравнения принято называть уравнениями погружения. Если при каком-либо значении параметра решение известно или может быть легко получено, то, решая уравнение погружения с такими начальными условиями, можно определить решения задачи при всех других значениях параметра, в частности, при значении, соответствующем исходной постановке. В качестве параметра погружения удобно выбрать толщину переходного слоя. Поэтому увеличение параметра погружения эквивалентно наложению дополнительного слоя на усеченную поверхность. При нулевом значении параметра усечение идет по нижней «кромке» переходного слоя, и мы имеем задачу об отражении поля от гладкой поверхности. Когда на усеченную поверхность начинают накладываться дополнительные слои, то мы плавно переходим к решению искомой задачи.

1. Электродинамическая модель периодически неоднородной земной поверхности с увлажненным верхним слоем

Рассмотрим участок земной поверхности, который может быть аппроксимирован гофром. Будем полагать также, что верхний приповерхностный слой земной поверхности имеет большую влажность, чем глубинные слои, что может быть обусловлено, например, прошедшими осадками или недавно сошедшим снежным покровом. Повышенная концентрация влаги в почве увеличивает ее диэлектрическую

проницаемость. Так, например, если "сухая" почва характеризуется диэлектрической проницаемостью ϵ_1 , то увлажненная имеет диэлектрическую проницаемость $\epsilon_{почв}$ [4, 5]

$$\epsilon_{почв} = \epsilon_1 + \epsilon_в \frac{\Delta\rho}{\rho_в} \tag{1}$$

Здесь ϵ_1 - диэлектрическая проницаемость «сухой» почвы; $\epsilon_в$ - диэлектрическая проницаемость воды; $\rho_в$ - плотность воды. Смоделируем неоднородность диэлектрической проницаемости увлажненной почвы следующим образом

$$\epsilon(\vec{r}) = \begin{cases} \epsilon_{вл.почв.} & \text{при } \vec{r} \in G_1 \\ \epsilon_{сух.почв.} & \text{при } \vec{r} \in G_2 \end{cases} \tag{2}$$

Области G_1 и G_2 обозначены на рис. 1. Как отмечалось выше, мы имеем гофрированную поверхность. Для определенности представим ее в виде цилиндрической поверхности, порожденной движением образующей вдоль направляющей, имеющей вид косинусоиды. Изменение влажности смоделируем введением конечного слоя с повышенным содержанием влаги. Учтем, что толщина слоя повышенной влажности меняется в направлении оси X , при переходе от горбов гофра к впадинам. Неоднородность по толщине смоделируем зависимостью $h(x) = A_1 \cos(\kappa x) + \delta - A_2 \cos(\kappa x)$.

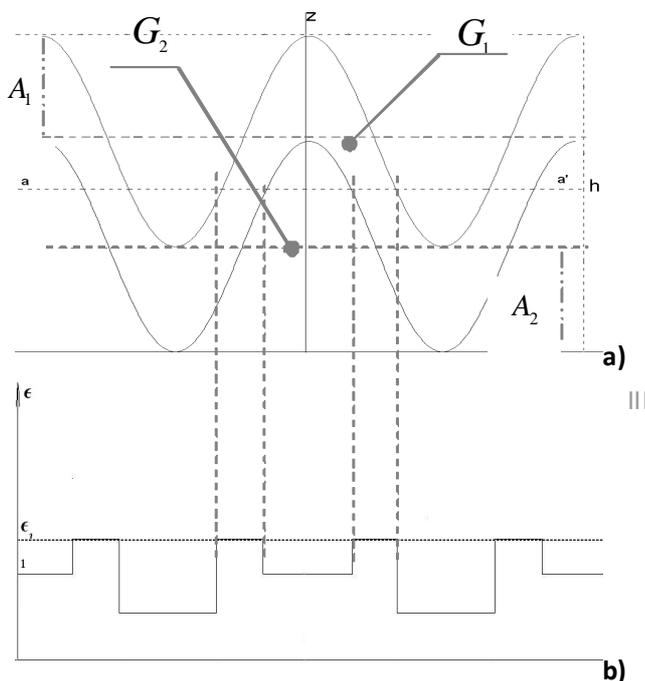


Рис. 1. Схематическая модель промокания почвы. **1а:** профиль переходного слоя на границе раздела двух сред. **1б:** график зависимости диэлектрической проницаемости $\epsilon(\vec{r})$ внутри переходного слоя в горизонтальном сечении aa' .

2. Размерность задачи и представление электромагнитного поля в модели

Свяжем с исследуемой поверхностью правостороннюю декартову систему координат так, что ось Z направлена вверх, а ось Y – параллельно образующей поверхности, ось X вдоль периодичности структуры. Будем полагать, что на переходный слой падает плоская электромагнитная волна (ЭМВ) с волновым вектором $\vec{k} = (q_x, q_y, k_z)$. Известно, что при взаимодействии плоской ЭМВ с периодической структурой в дифрагированном поле возникает пространственный спектр плоских волн с волновыми векторами, отличающимися друг от друга на целое число векторов обратной решетки $\kappa = \frac{2\pi}{\Lambda}$ периодической структуры. Это позволяет представить отраженное поле в виде дискретного разложения по угловому спектру $E = \dots + E_{-1} \cdot e^{i\vec{k}_{-1}\vec{r}} + E_0 \cdot e^{i\vec{k}_0\vec{r}} \dots$ или в виде вектора $\vec{E} = (\dots \vec{E}_{-1} \ E_0 \ E_1 \ \dots)^T$.

В обоих соотношениях \vec{E}_i – i -я амплитуда компоненты углового спектра поля.

Для определенности представим ее в виде цилиндрической поверхности, порожденной движением образующей вдоль направляющей, имеющей вид косинусоиды. Изменение влажности смоделируем введением конечного слоя с повышенным содержанием влаги. Учтем, что толщина слоя повышенной влажности меняется в направлении оси X , при переходе от горбов гофра к впадинам. Неоднородность по толщине смоделируем зависимостью $h(x) = A_1 \cos(\kappa x) + \delta - A_2 \cos(\kappa x)$.

Из рис. 1 видно, что все пространство делится на три области. В верхней среде однородна и характеризуется диэлектрическими проницаемостями $\epsilon \cong 1$, в нижней части она также однородна, но характеризуется диэлектрической проницаемостью $\epsilon_2 = \epsilon_{сух.почв.}$. Для этих частей пространства описание распространения волн не представляет затруднений. Основной интерес представляет центральная область, включающая в себя границу раздела сред. Эту область назовем переходным слоем. Заметим, что среда внутри переходного слоя пространственно неоднородна, и основная проблема здесь заключается в выборе адекватного метода описания процесса рассеяния.

Отметим, что волновые векторы дифрагированного поля не лежат в одной плоскости, т.е. перестает выполняться простой закон Виллброрда Снеллиуса. Это приводит к тому, что при взаимодействии поля с переходным слоем его поляризация существенно искажается, возникают так называемые кросс-поляризационные эффекты.

В нашей модели мы предполагаем, что поле задается в трехмерном пространстве (3D), а сама поверхность имеет 2D структуру, т.е. полная модель взаимодействия поля с периодической поверхностью может классифицироваться как квази-3D.

3. Уравнения погружения для модели переходного слоя

В [1, 3] приведены различные подходы к выводу уравнений погружения для матричных коэффициентов отражения и прозрачности (\hat{R} и \hat{T} соответственно). Нам понадобится лишь первое

$$\frac{d\hat{R}^+(z)}{dz} = \rho^+ + \tau^+ \hat{R}^+ + \hat{R}^+ \tau^+ + \hat{R}^+ \rho^- \hat{R}^+. \quad (1)$$

Здесь ρ^\pm и τ^\pm параметры коэффициентов отражения и прохождения элементарного слоя [7].

Решение дифференциального уравнения (1) должно удовлетворять следующему начальному, по параметру погружения, условию

$$R(0) = \begin{pmatrix} R_{11} & 0 & 0 \\ 0 & \dots & 0 \\ 0 & 0 & R_{nn} \end{pmatrix}, \text{ где } R_{nn} = \frac{\sqrt{k_0^2 - (q + 2\pi n / \Lambda)} - \sqrt{\varepsilon_2 k_0^2 - (q + 2\pi n / \Lambda)}}{\sqrt{k_0^2 - (q + 2\pi n / \Lambda)} + \sqrt{\varepsilon_2 k_0^2 - (q + 2\pi n / \Lambda)}}. \quad (2)$$

Здесь Λ – период регулярной составляющей профиля поверхности; q – проекция волнового вектора на плоскость XOY, а $\varepsilon_2 = \varepsilon_{\text{сух.почв.}}$ – диэлектрическая проницаемость сухой почвы.

Отметим, что общий вид уравнений инвариантен относительно структуры переходного слоя, а все особенности заключены в виде параметров коэффициентов отражения ρ^\pm и прохождения τ^\pm . Нахождение этих коэффициентов и составляет центральную часть нашей работы.

В соответствии с основной идеей метода погружения выделим достаточно тонкий (элементарный)

слой толщины Δz внутри переходного слоя (рис. 2). Сверху и снизу этот элементарный слой отделен от остальной среды бесконечно узкими зазорами, где диэлектрическая проницаемость $\varepsilon = 1$. Будем рассматривать поле в этих зазорах. Поле в верхнем и нижнем зазорах можно представить в виде совокупности плоских волн с положительными и отрицательными проекциями вектора \vec{k} на ось OZ. Первые мы будем обозначать индексом «+», а вторые – индексом «-».

Так как толщина Δz элементарного слоя мала, то рассеяние поля на нем можно рассчитывать в борновском приближении

$$\vec{E}(\vec{r}) = \vec{E}_0(\vec{r}) + \int_{\Delta\Omega} d\vec{r}' \hat{G}(\vec{r}, \vec{r}') [(\varepsilon(\vec{r}') - 1)k^2 \vec{E}_in(\vec{r}')],$$

где \hat{G} – тензорная функция Грина, $\Delta\Omega$ – часть пространства, занимаемая элементарным слоем, а $\vec{E}_0(\vec{r})$ – общее решение однородного уравнения, удовлетворяющее краевым условиям задачи для элементарного слоя. $\vec{E}_0(\vec{r})$ и $\vec{E}(\vec{r})$ – входящие и выходящие из элементарного слоя поля, а $\vec{E}_in(\vec{r})$ – внутреннее поле, определяющиеся по формуле

$$\vec{E}_in(\vec{r}) = \begin{cases} \hat{A} \vec{E}_0(\vec{r}) \text{ } i\partial\partial & \vec{r} \in \Delta\Omega' \\ \vec{E}_0(\vec{r}) \text{ } i\partial\partial & \vec{r} \notin \Delta\Omega' \end{cases}. \quad (3)$$

Отметим, что именно в конкретном виде соотношения (3) и заключено борновское приближение. Можно показать, что

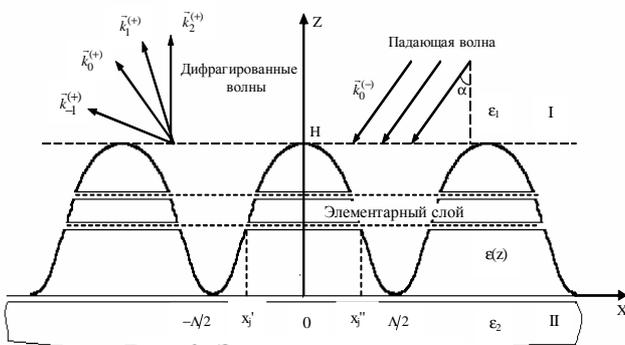


Рис. 2. Элементарный слой

$$E^{\pm}(q, z + \Delta z) = E^{\pm}(q, z)e^{ik_z \Delta z} + \sum_m G^{\pm}(q, \pm 0) \frac{k_0^2}{\pi(n-m)} \left\{ (\varepsilon - 1) \cdot A \cdot (E_n^+(z) + E_m^-(z)) \cdot \sin \left[\frac{2\pi}{\Lambda k} \cdot \arccos \left(\frac{z - \delta}{A_1} \right) \cdot (n - m) \right] + \tilde{\varepsilon} \cdot A \cdot (E_n^+(z) + E_m^-(z)) \cdot \sin \left[\frac{2\pi}{\Lambda k} \cdot \arccos \left(\frac{z}{A_2} \right) \cdot (n - m) \right] \right\} \Delta z \quad (4)$$

Введем связь между E_n^+ и E_m^+ через коэффициент прозрачности $E_n^{\pm} = t_{n,m}^{\pm} E_m^{\pm}$. Тогда из (4)

$$t_{n,m}^{\pm} = \delta_{nm} + iik_z \Delta z + G^{\pm}(q, \pm 0) \frac{k_0^2}{\pi(n-m)} \left\{ (\varepsilon - 1) \cdot A \cdot \sin \left[\frac{2\pi}{\Lambda k} \cdot \arccos \left(\frac{z - \delta}{A_1} \right) \cdot (n - m) \right] + \tilde{\varepsilon} \cdot A \cdot \sin \left[\frac{2\pi}{\Lambda k} \cdot \arccos \left(\frac{z}{A_2} \right) \cdot (n - m) \right] \right\} \Delta z \quad (5)$$

Для коэффициента отражения $r_{n,m}^{\pm}$, определяемого соотношением $E_n^{\pm} = r_{n,m}^{\pm} E_m^{\mp}$, имеем

$$r_{n,m}^{\pm} = G^{\pm}(q, \pm 0) \frac{k_0^2}{\pi(n-m)} \left\{ (\varepsilon - 1) \cdot A \cdot \sin \left[\frac{2\pi}{\Lambda k} \cdot \arccos \left(\frac{z - \delta}{A_1} \right) \cdot (n - m) \right] + \tilde{\varepsilon} \cdot A \cdot \sin \left[\frac{2\pi}{\Lambda k} \cdot \arccos \left(\frac{z}{A_2} \right) \cdot (n - m) \right] \right\} \Delta z \quad (6)$$

Члены, пропорциональные Δz , в (5) и (6), описывают взаимодействие электромагнитного поля с элементарным слоем. Именно они фигурируют в (1) как ρ^{\pm} и τ^{\pm} . Система уравнений (1), (3), (5) и (6) формирует математическую модель взаимодействия ЭМВ с увлажненной неровной поверхностью, которая может быть положена в основу компьютерного эксперимента.

ЛИТЕРАТУРА

1. Barabanenkov Yu.N., Kouznetsov V.L., Barabanenkov M.Yu. Transfer relations for electromagnetic wave scattering from periodic dielectric one-dimension interface: TE polarization. // Progress in Electromagnetic Research: PIRS, VOL. 24, 1999.
2. Кляцкин В.И. Метод погружения в теории распространения радиоволн. - М.: Наука, 1986.
3. Шалашов А.Г., Господчиков Е.Д. Импедансный метод решения задач распространения электромагнитных волн в анизотропных и гиротропных средах // УФН 181 151–172 (2011).
4. Барабаненков Ю. Н., Кузнецов В. Л. Матричное уравнение Риккати для задачи рассеяния векторного поля на двух-масштабной периодической поверхности // Радиотехника и электроника. - 1999. - Т. 44.
5. Mironov V.L., Dobson M.C., Kaupp V.H., Komarov S.A., Kleshchenko V.N. Generalized refractive mixing dielectric model for moist soils // IEEE Trans. Geosci. Remote Sensing, 2004. Vol. 42. No. 4.
6. Беляева Т.А., Бобров А.П., Бобров П.П., Мандрыгина В.Н. О выборе диэлектрической модели при определении диэлектрической проницаемости связанной воды в грунтах // Современные проблемы дистанционного зондирования Земли из космоса. Тезисы докладов Второй открытой Всероссийской конференции. - М.: 16-18 ноября 2004 г.
7. Кузнецов В.Л., Филонов П.В. Уравнения погружения для обобщенной матрицы рассеяния в теории нерегулярных волноводов // Научный Вестник МГТУ ГА. - № 157.
8. Борн М., Вольф Э. Основы оптики. - изд. 2-е / пер. с англ. - М.: Наука, 1973.
9. Басс Ф.Г., Фукс И.М. Рассеяние волн на статистически неровной поверхности. - М.: Наука, 1972.

MATHEMATICAL MODEL OF REFLECTION OF ELECTROMAGNETIC FIELD FROM THE GOFFERED SURFACE WITH MOISTENED THE UPPER LAYER

Boyarkin S.V.

The mathematical model of interaction of electromagnetic radiation with periodically indignant, humidified surface of soil is constructed. The invariant embedding method is put in a basis of construction of model.

Key words: electromagnetic waves reflection, invariant immersing method, goffered surface, earth humidified layer.

Сведения об авторе

Бояркин Сергей Валерьевич, 1987 г.р., окончил МГТУ ГА (2010), аспирант МГТУ ГА, область научных интересов – математическое моделирование в задачах дистанционного зондирования.

УДК 629.7.07

ОБ ОДНОЙ АНАЛИТИЧЕСКОЙ МОДЕЛИ ПРОПУСКНОЙ СПОСОБНОСТИ АЭРОПОРТА

В.Л. КУЗНЕЦОВ, А.А. ЧЕПУРИНА

Рассматривается задача об оценке максимальной пропускной способности аэропорта с одной взлетно-посадочной полосой. Предлагается модель организации воздушного движения в зоне аэропорта, позволяющая формировать аналитические оценки на ограничение пропускной способности, связанные с пересечением траекторий взлетающих и заходящих на посадку воздушных судов. В основе предлагаемой модели лежит аналогия между максимально плотным потоком воздушных судов в зоне аэропорта и ферми-системами.

Ключевые слова: пропускная способность аэропорта, минимумы эшелонирования, квантовые состояния, принцип запрета Паули.

Введение

При растущем объеме авиаперевозок ключевым моментом становится проблема увеличения пропускной способности аэропорта. При этом следует отметить, что пока не существует единой утвержденной методики расчета пропускной способности аэропорта, а есть лишь некоторые рекомендации. Ясно, что и методики, и рекомендации требуют для своего обоснования разработки соответствующих адекватных моделей функционирования аэропорта с включением в рассмотрение и области воздушного пространства аэродромной зоны. Существующие имитационные модели, стремясь максимально точно скопировать систему управления воздушным движением (УВД) в зоне аэропорта, позволяют выработать лишь определенные, возможно, очень важные рекомендации по организации УВД, но не дают аналитических оценок для предельно допустимого числа операций взлетов и посадок в единицу времени [1]. Попытки ассоциировать пропускную способность аэропорта с пропускной способностью взлетно-посадочной полосы (ВПП) также встречают серьезные возражения, связанные с отсутствием учета возможных ограничений, возникающих в воздушном пространстве аэродромной зоны. Эти ограничения связаны с наличием пересечений траекторий взлетающих и заходящих на посадку воздушных судов (ВС) и существующими минимумами эшелонирования. Другими словами, по свидетельству авиадиспетчеров, кроме ВПП, возникает еще одно «узкое место», лимитирующее величину пропускной способности – это область повышенной концентрации ВС в воздушной зоне аэропорта. При абстрагировании от этих проблем задача о пропускной способности аэропорта действительно сводится к вычислению времени занятости ВПП для одной операции типа взлет/посадка.

Целью настоящей статьи является разработка аналитического подхода к задаче расчета потенциальной пропускной способности аэропорта, понимаемой как максимально допустимое среднее число операций взлета и посадки в единицу времени при обеспечении заданного уровня безопасности полетов. Будем далее полагать, что требования по безопасности выполнены, если пространственные интервалы между ВС соответствуют ограничениям на минимумы эшелонирования.

1. Вербальное описание модели и квантовые аналогии

Рассмотрим аэропорт с одной ВПП, работающий по пропускной способности на пределе своих возможностей. Введем понятие «квантового состояния», определяемого как область воздушного пространства, внутри которого может находиться только одно ВС. Появление в этом состоянии второго ВС запрещено, т.к. неизбежно приводит к нарушению минимумов эшелонирования [2]. Отметим сразу, что квантовое состояние может быть как заполненным, так и свободным, т.е. в нем может находиться одно, а может и не находиться ни одного ВС. Проводя аналогию с квантовой теорией можно сказать, что ВС подобны фермионам, для которых справедлив принцип запрета Паули. В нашем случае каждое квантовое состояние представляет собой прямоугольный параллелепипед, центр которого совпадает с авиалинией – траекторией планового движения ВС, взлетевшего или направляющегося на посадку. Очевидно, что эти квантовые состояния не могут накладываться друг на друга, с одной стороны, и, с другой стороны, в предельном случае реализации максимальной пропускной способности аэропорта должны плотно примыкать друг к другу в направлении авиалиний. Таким образом картину воздушного пространства в окрестности аэродрома можно представить в виде цепочек из движущихся друг за другом вдоль авиалиний параллелепипедов.

Авиалинии ВС, заходящих на посадку, сливаются, переходя в конце концов в одну посадочную прямую - глассаду. В соответствии с этим сливаются и соответствующие квантовые состояния. Обратный процесс – рождение квантовых состояний, имеет место для взлетающих ВС. При слиянии двух квантовых состояний образуется новое, в котором может находиться только одно ВС. Следовательно, только одно из двух сливающихся квантовых состояний может быть заполнено, второе должно быть обязательно пустым. При рождении новых квантовых состояний, сопровождающем расщепление авиалиний, выполняется закон сохранения для числа ВС, поэтому при этом появляются новые свободные состояния. Из сказанного следует, что лишь немногие из параллелепипедов, описанных в картине воздушного пространства в окрестности аэродрома, заполнены ВС, большая их доля оказывается свободными. Понятно, что потенциальная пропускная способность аэропорта достигается в случае максимального допустимого заполнения квантовых состояний.

Авиалинии, по которым ВС подходят к аэропорту и, взлетая, уходят на нужных направлениях на соответствующие воздушные коридоры, имеют сложную топологию. Упростить задачу можно, введя путевые координаты, аналогично тому, как это представлено на рис. 1.

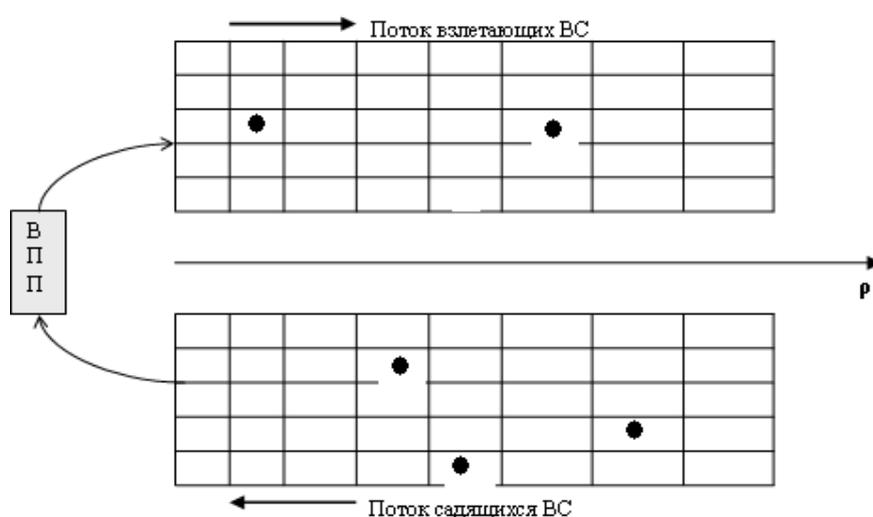


Рис. 1. Распределение квантованных объемов (допустимых квантовых состояний) вдоль путевых координат взлетающих и садящихся ВС: ρ - расстояние от ВПП, измеренное вдоль траекторий ВС. Кружками помечены заполненные в некоторый момент времени t_j квантовые состояния, т.е. состояния, в которых находятся ВС. Другие состояния свободны

Изображенные друг над другом прямоугольники в путевых координатах удалены на одинаковое расстояние от взлетно-посадочной полосы (ВПП) и будут сливаться по мере приближения к глассаде. Это значит, что в столбце состояний может быть заполнено только одно, точнее, не более одного ВС может находиться на данном «удалении» от ВПП. Аналогичное утверждение справедливо и для столбца состояний взлетевших ВС. Каждый такой столбец удобно описывать вектором $\vec{\psi}$, компоненты которого в основном нули, но может появиться и одна единица, т.е. либо $\vec{\psi} = (0,0,0,\dots,0)^T$, либо это вектор вида $\vec{\psi} = (0,1,0,\dots,0)^T$ с одной единицей, стоящей на k -м месте. Необходимость существования нулевых векторов связано уже с тем, что во время взлета ВС ВПП занята, и соответствующее квантовое состояние заходящих на посадку ВС должно быть свободно.

Следующим удобным приемом, упрощающим описание потока ВС, является переход в дискретное время. Это означает, что вместо движения квантовых состояний вдоль авиалиний удобно рассмотреть дискретные переходы ВС от одного столбца к следующему через интервал времени τ , определяемый как

$$\tau = \max \left\{ \tau_{взл}, \tau_{посад}, \left\{ \frac{L_i}{V_i} \right\} \right\}, \tag{1}$$

Здесь $\tau_{взл}, \tau_{посад}$ - времена занятости ВПП при взлете и посадке, L_i - значения минимумов продольного эшелонирования вдоль авиалинии, а V_i - плановая скорость ВС на соответствующем участке траектории.

При переходе от одного дискретного момента времени - t_j к следующему - t_{j+1} вектор $\vec{\psi}$ в сопровождающей системе отсчета может измениться вследствие перехода ВС за это время с одной авиалинии на другую, например, следующим образом: $\vec{\psi}(t_j) = (0, 1, 0, \dots, 0)^T \rightarrow \vec{\psi}(t_{j+1}) = (0, 0, 1, \dots, 0)^T$. Следует отметить, что в сопровождающей «системе отсчета», движущейся вместе с выделенным ВС, $|\vec{\psi}(t_k)| = const$ и не может меняться со временем: в каком-либо состоянии в следующий момент времени наше ВС обязательно будет находиться, а другие ВС в столбце появиться не могут.

Для описания сказанного удобно ввести два поля $\Psi^+(k, t_j)$ и $\Psi^-(m, t_l)$ для заходящих на посадку и взлетающих ВС соответственно. Сечение поля $\Psi^+(k, t_j)$ при $t = t_j = const$ дает описание заполнения квантовых состояний (распределение в воздушном пространстве) ВС, заходящих на посадку в момент времени t_j . Это схематически изображено кружками на рис. 1. Сечение при $k = const$ дает значение вектора $\vec{\psi}(t_j)$ - заполнения квантовых состояний на k - «расстоянии» от ВПП в любой момент времени t_j .

Краевые условия для полей $\Psi^+(k, t_j)$ и $\Psi^-(m, t_l)$ определяются потоками прибывающих - $\Psi^+(K, t_j)$ и взлетающих - $\Psi^-(0, t_j)$ ВС. Моделирование этих величин мы обсудим в этой работе позднее, ибо они, фактически, и определяют решение задачи.

2. Моделирование механизма возникновения конфликтных ситуаций в воздушном пространстве аэропорта

Обратимся к описанию проблем, возникающих из-за пересечения авиалиний взлетающих и заходящих на посадку ВС. Предложенный в работе формализм квантовых состояний и принципа запрета Паули приводит к дополнительным ограничениям на возможности заполнения состояний. Схематически эти ограничения можно понять, анализируя ситуацию, изображенную на рис.2.

Учитывая множественность точек пересечения авиалиний, рис. 1 и рис. 2 можно объединить в одну схему, изображенную на рис. 3. Пунктирными линиями соединены квантовые состояния, соответствующие взлетающим и заходящим на посадку ВС. Наклон этих линий отображает тот факт, что точки пересечения авиалиний могут находиться на разных путевых расстояниях от ВПП, т.е. расстояние, пройденное взлетающим ВС, не совпадает с расстоянием, которое должно пройти до ВПП ВС, заходящее на посадку. Математически эту связь можно описать с помощью некоторого матричного оператора $\hat{R}(k, m)$, элементы которого равны 1, если $m - e$ квантовое состояние для взлетающего потока совпадает с $k - m$ состоянием потока ВС, заходящего на посадку, и нулю в противном случае. Явный вид оператора определяется, естественно, топологией авиалиний в зоне аэропорта.

Дальнейший алгоритм расчета роли конфликтных ситуаций сводится к следующему. Препятствия для потока заходящих на посадку ВС, порожденные взлетающими ВС, можно описать некоторым зависящим от времени полем в пространстве путевых координат

$$\Phi(k, t_j) = \hat{R}(k, m) \Psi^-(m, t_j). \quad (2)$$

Образно назовем это поле минным. Рассмотрим какую-либо из реализаций входного потока $\Psi^+(K, t_i) = \vec{\psi}(t_i) = (0, 1, 0, \dots, 0)^T$ и обозначим множество допустимых авиалиний, приводящих это ВС к ВПП как $L = \{l_p\}$, $p = \overline{1, P}$. Если в этом множестве существует хотя бы одна траектория, минующая $\Phi(k, t_j)$, то будем говорить, что оптимальное управление для рассматриваемого ВС существует. В противном случае такая реализация невозможна.

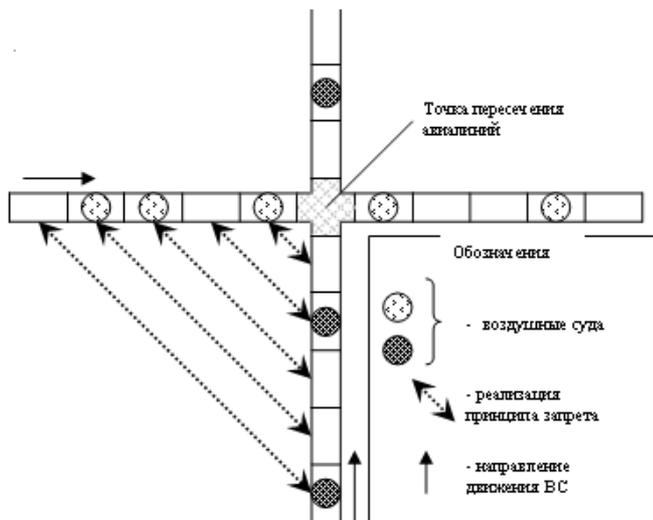


Рис. 2. Иллюстрация ограничений принципа Паули - запрета на заполнение квантовых состояний. В точке пересечения авиалиний только одно из совмещающихся состояний может быть занято

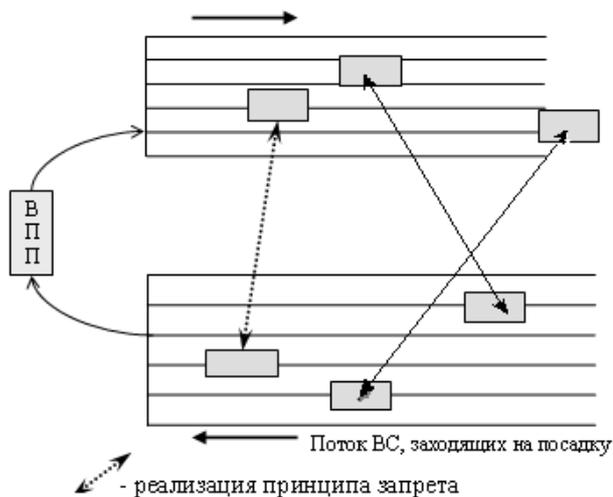


Рис. 3. Схема для расчета ограничений воздушной зоны аэропорта на его пропускную способность

До сих пор наши построения касались различных реализаций (траекторий), вообще говоря, случайного процесса. Поскольку в зону аэропорта ВС поступают в случайные моменты времени, то мы можем говорить лишь о вероятностях различных реализаций случайного вектора $\Psi^+(K, t_i)$. Другими словами, если число входных авиалиний - N , то заданным можно считать вектор $\vec{P} = (p_1, p_2, \dots, p_N)$. Это означает, в частности, что описанный выше пример реализуется с вероятностью p_2 . Конфигурация минного поля, как следует из (2), зависит от потока взлетающих ВС с учетом направления их рейсов - $\Psi^-(0, t_p)$. Аналогично входному потоку, для $\Psi^-(0, t_p)$ будем задавать вектор $\vec{P}' = (p'_1, p'_2, \dots, p'_N)$, описывающий распределение потока взлетающих ВС по соответствующим авиалиниям. Это позволяет рассчитать вероятности реализаций различных конфигураций минного поля $Q = (q_1, q_2, \dots, q_Z)$. Если X - полное число пересечений авиалиний, то число различных конфигураций минного поля (число различных заполнений взлетающими ВС квантовых состояний, предназначенных для ВС, заходящих на посадку) будет равно $Z = 2^X$. Если, например, для ВС, вошедшего в зону аэропорта по 2-й авиалинии, пятая и седьмая конфигурации минного поля окажутся непреодолимыми, т.е. для них не найдется оптимального управления, то ВС, прибывшее с вероятностью p_2 , совершит посадку в аэропорту с вероятностью $p_2 \cdot (1 - q_5 - q_7)$. В общем случае можно записать следующую формулу для вероятности посадки ВС в момент времени t_{j+k}

$$P(t_{j+k}) = \sum_{n=1}^N p_n \cdot \left[1 - \sum_{z \in Z_n} q_z \right]. \tag{3}$$

Здесь Z_n - множество реализаций минных полей, не проходящих для ВС, прибывших в зону аэропорта по n -й авиалинии.

Нетрудно видеть, что при отсутствии пересечений авиалиний вероятность реализации посадки равна $P(t_{j+k}) = \sum_{n=1}^N p_n = 0.5$. Нормировка на 0.5, а не 1 связана с предположением о стационарности работы аэропорта, предполагающей, что среднее число взлетающих ВС совпадает с числом совершающих посадку. Последнее означает, что квантовое состояние, связанное с ВПП, занято садящимися ВС в среднем с вероятностью 0.5.

Из изложенного понятно, что пропускная способность аэропорта C , определяемая как максимальное число операций взлета/посадки за единицу времени T , может быть найдена по формуле

$$C = 2 \left[\frac{T}{\tau} \right] \cdot \left(\sum_{n=1}^N p_n \cdot \left[1 - \sum_{z \in Z_n} q_z \right] \right). \quad (4)$$

Заключение

В работе развивается подход к задаче об оценке максимальной пропускной способности аэропорта, базирующийся на аналогии между максимально плотным потоком ВС в воздушной зоне аэропорта и ферми-системами. Минимумы эшелонирования определяют объемы, интерпретируемые как квантовые состояния, в которых, подобно Ферми-системам, не может находиться более одного объекта.

Предлагаемый подход направлен на вычисление максимально достижимой пропускной способности аэропорта, но не указывает путей достижения этого максимума.

ЛИТЕРАТУРА

1. Ускоренное моделирование процессов организации потоков и управления воздушным движением: современные и будущие возможности / под ред. Е.А. Федосова. - М.: ГосНИИАС, 2008.
2. ICAO, Manual on Airspace Planning Methodology for the Determination of Separating Minima (Doc 9689) First Edition, 1988.

ABOUT ONE ANALYTICAL MODEL OF THE AIRPORT THROUGHPUT

Kuznetsov V.L., Chepurina A.A.

The problem of estimation of the maximum airport throughput with one runway is considered. The model of air traffic organization in the airport zone is offered. It allowing to form analytical estimations for the throughput restriction, connecting a trajectories crossing effects for flying up and sitting down aircrafts. Some analogy between a much as possible dense stream of aircrafts in a zone of the airport and fermi-systems lies at the heart of offered model.

Key words: airport handing capacity, separation minima, quatum state, Pauli exclusion principle.

Сведения об авторах

Кузнецов Валерий Леонидович, 1949 г.р., окончил МГУ им. Ломоносова (1972), доктор технических наук, профессор, заведующий кафедрой прикладной математики МГТУ ГА, автор более 100 научных работ, область научных интересов – методы математического моделирования в задачах распространения излучения в пространственно неоднородных, случайных и периодических средах, безопасность полетов.

Чепурина Анна Александровна, студентка кафедры прикладной математики МГТУ ГА, область научных интересов – моделирование задач УВД, безопасность полетов.

УДК 364.35

ДВУХПЕРИОДНАЯ МОДЕЛЬ ПЕРЕКРЫВАЮЩИХСЯ ПОКОЛЕНИЙ С ЭКЗОГЕННЫМ ПРЕДЛОЖЕНИЕМ ТРУДА: ЭФФЕКТЫ ЗАМЕЩЕНИЯ И ДОХОДА В УСЛОВИЯХ СОЛИДАРНОЙ ПЕНСИОННОЙ СИСТЕМЫ

Ю.Ф. КАСИМОВ, А.С. ЛУЗГИНА

Настоящая статья посвящена изучению эффектов замещения и дохода в модели перекрывающихся поколений Даймонда-Самуэльсона с экзогенным предложением труда. Основное внимание уделено модели с солидарной пенсионной системой; модели с фондируемой пенсионной системой или же не предполагающие пенсионной системы вообще здесь фактически сводятся к ее частному случаю. Полученные результаты дают более четкое представление о поведении индивидов, их реакции на изменение рыночной конъюнктуры.

Ключевые слова: макроэкономические модели, солидарные пенсионные системы, эффект замещения, чувствительность модели.

Введение

Модель перекрывающихся поколений, впервые предложенная Самуэльсоном (1958) и Даймондом (1965), представляет собой динамическую экономическую модель. Модель предполагает, что экономическая ситуация определяется общим поведением индивидов, компаний и правительства.

Главными контролируруемыми переменными являются объем сбережений и предложение труда. Официальная налоговая ставка регулируется политикой правительства. Параметрами модели являются: предельная доля сбережений в доходе, норма временного предпочтения, уровень неприятия риска, норма амортизации капитала, темп прироста населения и темп роста технологического прогресса, темп роста отчислений в пенсионный фонд. Главным параметром состояния является удельный объем основного капитала.

В данной статье мы ограничимся рассмотрением простейшей модели, согласно которой жизненный цикл индивидов состоит из двух жизненных периодов: до и после выхода на пенсию. В этом случае в каждый момент времени имеются два поколения: молодое и пожилое. Предполагается, что индивиды в каждом поколении стремятся максимизировать свою выгоду (функцию полезности), в то время как фирмы стремятся максимизировать прибыль. Также вводится предположение, что индивиды в пределах одного поколения идентичны. Предложение труда является экзогенным параметром, что означает, что все молодые индивиды обеспечены полной занятостью, т.е. затрачивают одну единицу труда. Будем считать, что эластичность функции полезности произвольна, одинакова для обоих жизненных периодов индивида и не меняется с течением времени.

Далее приведен анализ эффектов замещения и дохода, которые возникают при изменении рыночной конъюнктуры (конкретнее, ставки доходности на инвестированный капитал) и обусловлены стремлением индивидов распределить ресурсы во времени наиболее выгодным образом.

1. Эффекты замещения и дохода в условиях отсутствия пенсионной системы

Обратимся сначала к простейшему случаю, когда пенсионная система отсутствует. Этот вопрос был подробно рассмотрен Дэвидом Блэйком в книге "Pension Economics" (2006). Поскольку, как известно, фондируемая пенсионная система (при некоторых предположениях) нейтральна для экономики, то приведенные ниже результаты распространяются также и на этот случай.

Итак, отсутствие пенсионной системы предполагает, что потребление пожилых индивидов финансируется исключительно за счет сбережений, сделанных в первом периоде жизни, и начисленных на них процентов.

Таким образом, индивиды, рожденные в период t , стремятся максимизировать свою прямую функцию полезности

$$\Lambda_t^y(C_t^y, C_{t+1}^o) = U(C_t^y) + \frac{1}{1+p} U(C_{t+1}^o)$$

с учетом бюджетного ограничения

$$W_t = C_t^y + \frac{1}{1+r_{t+1}} C_{t+1}^o.$$

Здесь C_t^y, C_{t+1}^o - уровни потребления в молодом и пожилом возрасте соответственно; p - субъективная ставка дисконтирования; W_t - уровень заработной платы; r_{t+1} - ставка инвестирования; $U(C_t)$ есть некоторая функция полезности. Ограничимся рассмотрением изо-эластичной функции полезности, т.е. для заданного коэффициента эластичности предельной полезности σ^{-1} функцию полезности будем определять по следующей формуле

$$U(C_t) = \begin{cases} \frac{C_t^{1-\frac{1}{\sigma}}}{1-\frac{1}{\sigma}}, & \sigma > 0, \sigma \neq 1 \\ \ln C_t, & \sigma = 1 \end{cases}.$$

Соответственно как решение задачи на условный экстремум находятся оптимальные уровни потребления, которые, очевидно, будут пропорциональны ставке заработной платы с соответствующими коэффициентами, определяющимися параметрами модели

$$(C_t^y)^* = \Theta^y(p, r_{t+1}, \sigma) W_t,$$

$$(C_{t+1}^o)^* = \Theta^o(p, r_{t+1}, \sigma) W_t$$

где

$$\Theta^o(p, r_{t+1}, \sigma) = \left(\frac{1}{1+r_{t+1}} + \left(\frac{1+p}{1+r_{t+1}} \right)^\sigma \right)^{-1},$$

$$\Theta^y(p, r_{t+1}, \sigma) = \left(1 - \frac{\Theta^o(p, r_{t+1}, \sigma)}{1+r_{t+1}} \right)$$

Отсюда можно понять, что с ростом ставки доходности будет увеличиваться потребление во втором жизненном периоде индивида. Однако относительно потребления в первом периоде окончательный вывод можно сделать лишь для конкретно заданных параметров модели: с ростом ставки доходности оно может увеличиваться, уменьшаться или же оставаться неизменным.

Чтобы понять причину такой чувствительности к параметрам модели, рассмотрим функцию сбережения $S_t = S(W_t, r_{t+1})$, для которой выполнено

$$S_t = W_t - C_t^y; \quad C_{t+1}^o = S_t(1+r_{t+1}).$$

Это означает, что, во-первых, сбережения формируются из заработной платы молодых индивидов за вычетом текущего потребления и, во-вторых, потребление пожилых индивидов финансируется за счет сбережений и начисленных на них процентов.

Можно показать, что

$$\frac{\partial}{\partial r_{t+1}} S(W_t, r_{t+1}) - 0 \begin{matrix} > \\ < \end{matrix} \Leftrightarrow \sigma - 1.$$

Таким образом, в зависимости от степени расположенности индивида, распределить потребление равномерно по времени. Имеются 3 случая:

1) $\sigma > 1$.

В этом случае с ростом ставки доходности индивиды увеличат объем сбережений за счет потребления в молодом возрасте, что, в свою очередь, приведет к еще большему увеличению потребления во втором жизненном периоде. Следовательно, мы можем наблюдать эффект замещения: потребление в первом периоде замещается более "дешевым" потреблением в пожилом возрасте;

2) $\sigma < 1$.

В этом случае с ростом ставки доходности индивиды, напротив, уменьшат объем сбережений, израсходовав больше средств на потребление в первом периоде. Здесь мы можем наблюдать эффект дохода: поскольку рост ставки доходности приводит к увеличению суммы начисленных процентов, то для

обеспечения фиксированного уровня потребления в пожилом возрасте будет достаточно меньшего объема инвестированного капитала;

3) $\sigma = 1$.

Случай логарифмической функции полезности. Эффекты сбережения и дохода полностью компенсируют друг друга.

2. Эффекты замещения и дохода в условиях солидарной пенсионной системы

Перейдем теперь к рассмотрению более сложного случая. Как известно, солидарная система предполагает выплату пенсионных компенсаций пожилым индивидам из средств, полученных от молодых индивидов в качестве отчислений в пенсионный фонд; т.е. если L_t - число индивидов, рожденных в период t , P_t и T_t - размер пенсионной компенсации и пенсионных отчислений за соответствующий период, то

$$L_{t-1}P_t = L_tT_t.$$

Пусть g - темп роста отчислений в пенсионный фонд, n - темп прироста населения. Тогда единое бюджетное ограничение для индивида принимает вид

$$\hat{W}_t \stackrel{def}{=} W_t - \frac{(1+r_{t+1}) - (1+g)(1+n)}{1+r_{t+1}} T_t = C_t^y + \frac{1}{1+r_{t+1}} C_{t+1}^o.$$

Аналогичным образом находятся оптимальные уровни потребления

$$(C_t^y)^* = \Theta^y(p, r_{t+1}, \sigma) \hat{W}_t; \quad (C_{t+1}^o)^* = \Theta^0(p, r_{t+1}, \sigma) \hat{W}_t.$$

Очевидно, в условиях данной модели сбережения будут формироваться из располагаемого дохода (т.к. из зарплаты удерживаются пенсионные отчисления) за вычетом текущего потребления. Введем функцию сбережения $S_t = S(W_t, r_{t+1}, T_t)$, такую что

$$\begin{aligned} S_t &= (W_t - T_t) - C_t^y = \\ &= \left\{ 1 - \Theta^y(p, r_{t+1}, \sigma) \right\} W_t - \left\{ 1 - \Theta^y(p, r_{t+1}, \sigma) \frac{(1+r_{t+1}) - (1+g)(1+n)}{1+r_{t+1}} \right\} T_t \end{aligned}$$

и исследуем ее частную производную по второму аргументу.

Можно показать, что

$$\begin{aligned} \frac{\partial}{\partial r_{t+1}} S(W_t, r_{t+1}, T_t) &= \frac{\Theta^y(p, r_{t+1}, \sigma)}{(1+r_{t+1})^2} \{ P_{t+1} - C_{t+1}^o (1-\sigma) \} = \\ &= \frac{\Theta^y(p, r_{t+1}, \sigma)}{(1+r_{t+1})^2} \{ \sigma \cdot C_{t+1}^o - S_t (1+r_{t+1}) \} \end{aligned}$$

Получим

$$\frac{\partial}{\partial r_{t+1}} S(W_t, r_{t+1}, T_t) \geq 0 \Leftrightarrow \frac{P_{t+1}}{C_{t+1}^o} \geq 1 - \sigma,$$

или, что то же самое,

$$\frac{\partial}{\partial r_{t+1}} S(W_t, r_{t+1}, T_t) \geq 0 \Leftrightarrow \frac{S_t (1+r_{t+1})}{C_{t+1}^o} \leq \sigma.$$

Таким образом, параметр σ в данном случае уже не является единственным фактором, определяющим реакцию индивидов на изменение ставки. Имеется положительная зависимость между объемом сбережений индивидов и ставкой доходности в том и только том случае, если после выхода на пенсию доля пенсионной компенсации в их потреблении не менее величины $1 - \sigma$, или, что то же самое, когда доля отложенных на старость средств с учетом процентов в потреблении пожилого индивида не превышает величину σ . Очевидно, что $\forall \sigma \geq 1$ это условие будет выполнено.

Учитывая, что σ характеризует степень расположенности индивидов распределить потребление равномерно по времени, это означает, что индивиды, склонные к дисбалансу, однозначно увеличат объем сбережений при росте ставки доходности, т.к. это сулит им большую прибыль в будущем и наоборот.

С другой стороны, индивиды, склонные распределить потребление равномерно по времени $\sigma \in (0,1)$, принимая решение о сокращении объема сбережений при уменьшении ставки доходности, будут руководствоваться тем, достаточна ли сумма пенсионной компенсации; если нет, то сбережения являются «жизненной необходимостью», и сокращение их объема невозможно, несмотря на падение доходности. Напротив, индивиды вынуждены будут увеличить объем сбережений, чтобы достичь оптимального уровня потребления в пожилом возрасте, несмотря на падение начисляемых на сбережения процентов. Т.е. индивиды могут позволить себе «играть» на ставке доходности лишь в определенной мере, которая зависит от их степени неприятия риска.

Если же оптимальный уровень потребления пожилого индивида практически полностью финансируется за счет выплачиваемой государством пенсии (и необходимость делать сбережения мала), то уменьшение ставки доходности подтолкнет всех индивидов, вне зависимости от их уровня неприятия риска, уменьшить объем сбережений. Коррекция на ставку сбережений, однако, в этом случае будет мала в силу непрерывности функции сбережения по второму аргументу.

3. Заключение

При помощи двухпериодной модели перекрывающихся поколений Даймонда-Самуэльсона с экзогенным предложением труда исследованы эффекты замещения и дохода, которые возникают при изменении рыночной конъюнктуры. Выведены условия их доминирования.

Показано, что помимо коэффициента неприятия риска в модели с солидарной пенсионной системой определяющую роль играет доля пенсионной компенсации в потреблении пожилого индивида.

ЛИТЕРАТУРА

1. **David Blake**, Pension Economics, John Wiley & Sons, 2006.
2. **David de la Croix, Philippe Michel**, A Theory of Economic Growth Dynamics and Policy in Overlapping Generations, Cambridge University Press, 2002.
3. **Birgit Mattil**, The pension systems: Sustainability and Distributional Effects in Germany and the United Kingdom, A Physica Verlag Heidelberg product, 2006.

THE TWO-PERIOD DIAMOND-SAMUELSON OLG MODEL: INCOME AND SUBSTITUTION EFFECTS IN A PAYG SYSTEM

Kasimov Y.F., Luzgina A.S.

The article is devoted to the analyses of income and substitution effects in the two-period Diamond-Samuelson OLG model with exogenous labor supply. We paid particular attention to the PAYG system; the conclusions derived for the fully-funded scheme or a model without state pension scheme can be regarded as its special case. The obtained results give a better indication of the agent's behavior, their response to the turn of market.

Key words: macroeconomics models, solidary pension system, substitution effect, model's sensitivity.

Сведения об авторах

Касимов Юрий Федорович, 1946 г.р., окончил МГУ (1970), заведующий лабораторией МГТУ ГА, автор более 30 научных работ, область научных интересов - теоретическая и прикладная информатика, актуарная и финансовая математика.

Лузгина Анна Сергеевна, студентка РУДН, область научных интересов – математические модели пенсионных систем.

УДК 629.7.07

МОДИФИЦИРОВАННЫЙ МЕТОД ДИСКРЕТНЫХ ВИХРЕЙ ДЛЯ РЕШЕНИЯ СИНГУЛЯРНЫХ ИНТЕГРАЛЬНЫХ УРАВНЕНИЙ НА ОТРЕЗКЕ

А.В. ДВОРАК, Е.М. ИВЕНИНА, С.В. ФИЛИМОНОВ

Статья представлена доктором технических наук, профессором Кузнецовым В.Л.

Рассмотрена модификация метода дискретных вихрей для решения сингулярных интегральных уравнений на отрезке. Приведен пример решения модельной задачи. Сравниваются результаты расчета при различных вариантах неравномерного разбиения отрезка интегрирования, а также расчета классическим методом при аналогичных разбиениях отрезка.

Ключевые слова: модифицированный метод дискретных вихрей, решение сингулярных интегральных уравнений, неравномерное разбиение отрезка интегрирования.

Введение

Использование потенциала двойного слоя при решении задач бесциркуляционного обтекания в двумерном случае приводит к интегральному уравнению вида

$$\int_L \left(\frac{K(\bar{r}, \bar{r}_0)}{|\bar{r} - \bar{r}_0|^2} g(\bar{r}) \right) dl = f(\bar{r}_0), \quad (1)$$

где L – линия профиля, $K(\bar{r}, \bar{r}_0) = \bar{n}(\bar{r}_0) \cdot \bar{n}(\bar{r}) - 2 \frac{(\bar{n}(\bar{r}_0) \cdot (\bar{r} - \bar{r}_0))(\bar{n}(\bar{r}) \cdot (\bar{r} - \bar{r}_0))}{|\bar{r} - \bar{r}_0|^2}$ – гладкое (или кусочно-гладкое) ядро; $\bar{n}(\bar{r}_0)$, $\bar{n}(\bar{r})$ – орты нормали к L в точках \bar{r}_0 , \bar{r} соответственно; $g(\bar{r})$ – искомая плотность потенциала двойного слоя; $f(\bar{r}_0) = -2\pi \bar{u}_0(\bar{r}_0) \cdot \bar{n}(\bar{r}_0)$ – правая часть, определяемая нормальной компонентой скорости невозмущенного потока \bar{u}_0 в точках контура L .

В точках \bar{r}_0 , $\bar{r} \in L$ подынтегральная функция имеет сильную (неинтегрируемую) особенность, и интеграл необходимо понимать в смысле конечного значения по Адамару [3]

$$\int_L \frac{f(\bar{r}, \bar{r}_0)}{|\bar{r} - \bar{r}_0|^2} dl = \lim_{\varepsilon \rightarrow 0} \left(\int_{L \setminus R_\varepsilon} \frac{f(\bar{r}, \bar{r}_0)}{|\bar{r} - \bar{r}_0|^2} dl - \frac{2}{\varepsilon} f(\bar{r}_0, \bar{r}_0) \right), \quad R_\varepsilon = \{\bar{r} \mid |\bar{r}_0 - \bar{r}| < \varepsilon\}.$$

Если L – отрезок прямой ($L = [a, b] \in Ox$), имеем $K(\bar{r}, \bar{r}_0) = 1$ и уравнение (1) принимает вид

$$\int_a^b \frac{g(x)}{|x - x_0|^2} dx = f(x_0). \quad (2)$$

1. Аппроксимация интеграла

Рассмотрению метода дискретных вихрей как метода численного решения сингулярных интегральных уравнений (в том числе и уравнений с сильной особенностью) посвящены многие работы [1-4]. В частности, И.К. Лифановым [4] доказана сходимость квадратурных формул типа прямоугольников к интегралу (2), когда $x_0 \in [a + \delta, b - \delta]$, δ – фиксировано. При этом принципиальное значение имеет характер разбиения отрезка $[a, b]$ на частичные отрезки ($a = x_1 < x_2 < \dots < x_n = b$) и выбор точек коллокации x_{0i} внутри частичных отрезков $[x_i, x_{i+1}]$: разбиение должно быть равномерным, а контрольные точки x_{0i} должны являться серединами частичных отрезков.

Приведем квадратурные формулы, которые позволяют отказаться от столь жестких ограничений.

Зафиксируем произвольно малый шаг $h > 0$ и для $n \geq (b - a) / h$ разобьем отрезок $[a, b]$ точками x_i

($i = \overline{1, n+1}$) так, чтобы $a = x_1 < x_2 < \dots < x_{n+1} = b$, $x_{i+1} - x_i \leq h$, $i = \overline{1, n}$. Внутри каждого частичного отрезка $[x_i, x_{i+1}]$ зафиксируем также произвольно точку x_{0i} . Единственное ограничение на разбиение в том, чтобы при устремлении $h \rightarrow 0$ ($n \rightarrow \infty$) существовала константа $c_1 \in (0, 0.5]$, не зависящая от h , такая, что $\min_{i,j} |x_i - x_{0j}| \geq c_1 h$. Пусть на отрезке $[a, b]$ функция $g(x)$ допускает представление

$$g(x) = g(x_0) + g'(x_0)(x - x_0) + g_2(x, x_0)(x - x_0)^2, \\ |g_2(x, x_0)| \leq c_2, \quad \forall x, x_0 \in [a, b],$$

тогда можно показать, что для произвольной точки $x \in \{x_{0i}\}$ справедливо неравенство

$$\left| \int_a^b \frac{g(x)}{|x - x_0|^2} dx - \sum_{i=1}^n A_i g(x_{0i}) \right| \leq \frac{c_2}{c_1} (c_1 + 6 + 2c_1 |\ln c_1|) h,$$

где коэффициенты квадратурной суммы

$$S(x_0) = \sum_{i=1}^n A_i g(x_{0i}) \quad (3)$$

определяются равенствами

$$a_i = \frac{x_{i+1} - x_i}{(x_{i+1} - x_0)(x_i - x_0)}, \quad b_i = a_i (x_0 - x_{0i}) + \ln \left| \frac{x_{i+1} - x_0}{x_i - x_0} \right|, \\ i = \overline{1, n}; \\ A_i = a_i + \frac{b_{i-1}}{x_{0i} - x_{0i-2}} - \frac{b_{i+1}}{x_{0i+2} - x_{0i}}, \quad i = \overline{2, n-1}; \\ A_1 = a_1 - \frac{b_2}{x_{03} - x_{01}}; \quad A_n = a_n + \frac{b_{n-1}}{x_{0n} - x_{0n-2}}.$$

Вначале рассмотрим аппроксимацию интеграла (2) квадратурной суммой

$$S_1(x_0) = \sum_{i=1}^n (a_i g(x_{0i}) + b_i g'(x_{0i})), \quad (5)$$

где a_i, b_i определяются первыми двумя равенствами (4) и имеют смысл

$$a_i = \int_{x_i}^{x_{i+1}} \frac{dx}{(x - x_0)^2}; \quad b_i = \int_{x_i}^{x_{i+1}} \frac{x - x_{0i}}{(x - x_0)^2} dx.$$

При $x_0 = x_{0i}$ интегралы понимаются в смысле конечного значения по Адамару и главного значения по Коши соответственно. Производные $g'(x_{0i})$ аппроксимируем конечно-разностным аналогом

$$g'(x_{0i}) \approx g'_{0i} = \frac{g(x_{0i+1}) - g(x_{0i-1})}{x_{0i+1} - x_{0i-1}}; \quad i = \overline{2, n-1}. \quad (6)$$

Отметим, что, подставляя правые части (6) в (5) и перегруппировывая слагаемые, приходим к квадратурной сумме (3) с коэффициентами (4). Можно показать, что квадратура (3) обеспечивает равномерную сходимость к значению интеграла (2) на всем интервале (a, b) .

2. Численное решение уравнения на отрезке

Решение уравнения (2) с сильной особенностью в классе функций, обращающихся в нуль на концах отрезка $[a, b]$, $\int_a^b \frac{g(x)}{|x - x_0|^2} dx = f(x_0)$ проводится редукцией к системе линейных алгебраических уравнений следующим образом.

Отрезок $[a, b]$ произвольным образом с учетом указанных ограничений разбивается на n частичных отрезков точками x_i , $a = x_1 < x_2 < \dots < x_{n+1} = b$, $i = \overline{1, n+1}$. Внутри каждого из отрезков достаточно произвольно выбираем точки коллокации $x_{0i} \in (x_i, x_{i+1})$, $i = \overline{1, n}$. Затем, обозначая $g_i = g(x_{0i})$ – искомые неизвестные значения функции, интегральное уравнение (2) заменяем системой линейных алгебраических уравнений $\sum_{i=1}^n A_{ij} g_i = f(x_{0j})$, $j = \overline{1, n}$. Здесь A_{ij} вычисляются в соответствии с (4)

$$A_{ij} = a_{ij} + \frac{b_{i-1j}}{x_{0i} - x_{0i-2}} - \frac{b_{i+1j}}{x_{0i+2} - x_{0i}}, \quad i = \overline{2, n-1}, \quad j = \overline{1, n};$$

$$A_{1j} = a_{j1} - \frac{b_{j2}}{x_{03} - x_{01}}; \quad A_{nj} = a_{jn} + \frac{b_{jn-1}}{x_{0n} - x_{0n-2}}.$$

$$a_{ji} = \frac{x_{i+1} - x_i}{(x_{i+1} - x_{0j})(x_i - x_{0j})}, \quad b_{ji} = a_{ij}(x_{0j} - x_{0i}) + \ln \left| \frac{x_{i+1} - x_{0j}}{x_i - x_{0j}} \right|,$$

$$i = \overline{1, n}, \quad j = \overline{1, n};$$

Матрица $W = \{a_{ij}\}$ является обычной матрицей коэффициентов для метода дискретных вихрей с использованием вихревых пар (МДВ). Отметим, что использование данных квадратурных формул приводит к отличию в матрице модифицированного метода (ММДВ) от матрицы обычного метода дискретных вихрей с использованием вихревых пар лишь введением в нее членов b_{ij} , что позволяет учесть линейные члены разложения $g(x)$ и производить требуемую коррекцию в случае неравномерного распределения вихрей и контрольных точек. Последнее обстоятельство мало важно при решении двумерных задач обтекания, но становится принципиальным в пространственном случае, при сравнительно сложной формуле поверхности, когда равномерное разбиение просто не реализуемо. Одним из немногих случаев, когда применение данного метода оправдано в двумерных задачах, является исследование шарнирных характеристик закрылков (предкрылков), когда использование равномерного разбиения крайне затруднено из-за чрезвычайно мелкого шага, требуемого на закрылке.

3. Решение модельной задачи

Рассмотрим пример применения алгоритма к решению уравнения (2) на отрезке $[0, 1]$

$$\int_0^1 \frac{g(x)}{|x - x_0|^2} dx = f(x_0), \tag{7}$$

где $f(x_0) = -1$. Точное решение уравнения в классе функций, обращающихся в 0 на концах отрезка $[0, 1]$ $g(x) = \frac{1}{\pi} \sqrt{x - x^2}$. На рис. 1 приведен результат решения этой задачи для уравнения (7) в случае равномерного разбиения отрезка $[0, 1]$, когда точки коллокации выбираются случайным образом из δ -окрестностей середин частичных отрезков. Рассматриваются четыре случая δ -окрестностей величин 0.1, 0.2, 0.5 и 0.9 от длины каждого частичного отрезка модифицированного метода (ММДВ) и для сравнения приведен результат расчета с помощью метода дискретных вихрей (МДВ) для δ -окрестности величины 0.9 от длины частичного отрезка. На рис. 1 представлена зависимость среднеквадратичной ошибки $\sigma = \sqrt{\frac{1}{n} \sum_{i=1}^n (g_{\tau}(x) - g_p(x))^2}$ от $\frac{1}{n}$, где $g_{\tau}(x)$ – точное решение, n – число частичных отрезков, $g_p(x)$ – расчетное решение уравнения (7). Как и ожидалось, наилучшая сходимость наблюдается в случаях ММДВ для 0.1 и 0.2, а наихудшую сходимость проявляет МДВ для 0.9.

На рис. 2 приведены результаты расчета среднеквадратичной ошибки для случая неравномерного разбиения отрезка $[0,1]$ с точками коллокации в серединах частичных отрезков. Ожидаемая наилучшая сходимось проявилась в случаях ММДВ для 0.1 и 0.2, а наихудшая в случае МДВ для 0.9.

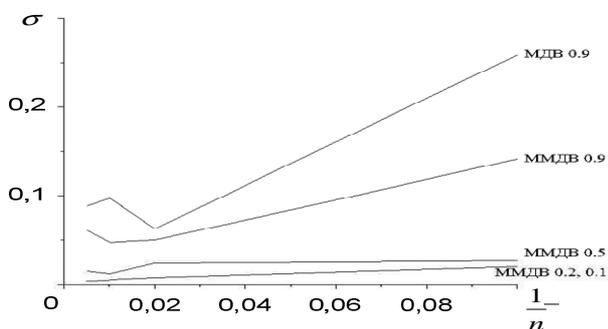


Рис. 1

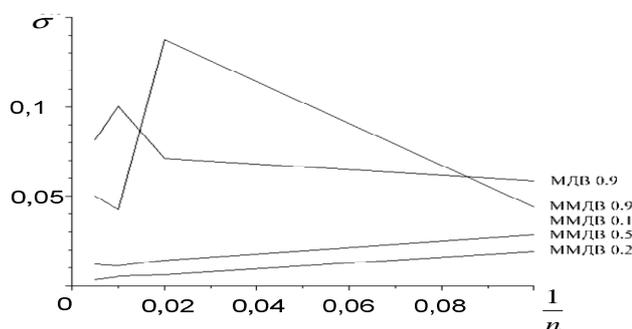


Рис. 2

Заключение

При решении задач проектирования летательных аппаратов часто требуется применение численных методов на неравномерных сетках, так как поверхностные особенности конструкции самолета имеют нерегулярную структуру. В этих случаях для метода дискретных вихрей вопрос о выборе контрольных точек решается исходя из опыта расчетчика, поэтому нахождение квадратур, устойчивых к неравномерности сетки и неопределенности положения контрольных точек, является актуальным. На примере решения модельной задачи было продемонстрировано преимущество модифицированного метода дискретных вихрей на неравномерной сетке при случайном выборе контрольных точек.

ЛИТЕРАТУРА

1. Белоцерковский С.М., Скрипач Б.К. Аэродинамические производные летательного аппарата и крыла при дозвуковых скоростях. - М.: Наука, 1975.
2. Белоцерковский С.М., Ништ М.И. Отрывное и безотрывное обтекание тонких крыльев идеальной жидкостью. - М.: Наука, 1978.
3. Белоцерковский С.М., Лифанов И.К. Численные методы в сингулярных интегральных уравнениях и их применение в аэродинамике, теории упругости, электродинамике. - М.: Наука, 1985.
4. Лифанов И.К. Метод сингулярных интегральных уравнений и численный эксперимент. - М.: Янус, 1995.

MODIFIED DISCRETE VORTICES METHOD FOR SINGULAR INTEGRAL EQUATION ON AN INTERVAL

Dvorak A.V., Ivenina E.M., Filimonov C.V.

Modification of discrete vortices method for solving singular integral equations on an interval is described in the article. The calculation results for different variants of the irregular partition of the integration interval are compared.

Key words: modified discrete vortices method, the solution of singular integral equations, the irregular partition of the interval of integration.

Сведения об авторах

Дворак Александр Владимирович, 1956 г.р., окончил Киевский университет (1973), кандидат физико-математических наук, доцент кафедры ВВА им. Н.Е. Жуковского и Ю.А. Гагарина, автор более 90 научных работ, область научных интересов – математическое моделирование, численные методы и методы математической физики.

Ивенина Елена Михайловна, окончила МГУ (1986), старший преподаватель кафедры прикладной математики МГТУ ГА, автор 8 научных работ, область научных интересов – численные методы, методы математического моделирования и комплексы программ.

Филимонов Сергей Владимирович, 1979 г.р., окончил ВВИА им. проф. Н.Е. Жуковского (2001), кандидат технических наук, доцент кафедры ВУНЦ ВВС «ВВА им. Н.Е. Жуковского и Ю.А. Гагарина», автор более 20 научных работ, область научных интересов – аэроупругость и численные методы в аэродинамике.

УДК 519.683.8

ИСПОЛЬЗОВАНИЕ СПЕЦИАЛИЗИРОВАННОГО КОМПЛЕКСА SCADЕ ДЛЯ РАЗРАБОТКИ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ БОРТОВОЙ ФУНКЦИИ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ

М.В. БАХИРКИН, В.Н. КАНАДИН, Г.Ю. ХРИСТОФОРОВ

Статья представлена доктором технических наук, профессором Рудельсоном Л.Е.

Представлено описание ключевых элементов использования специализированного программного комплекса SCADЕ при разработке бортового программного обеспечения.

Ключевые слова: специализированный программный комплекс, бортовое программное обеспечение.

Введение

Модернизация Единой системы организации воздушного движения (ОрВД) в России рассматривается как развитие сегмента общемировой структуры, базирующееся на согласованных подходах к целям и критериям развития, отраженным в руководящих документах гражданской авиации (ГА). Для проведения исследований, анализа требований и функционального тестирования элементов алгоритмического и программного обеспечения (ПО) перспективных функций обслуживания воздушных судов (ВС) создана модель динамической воздушной обстановки (МДВО).

В простейшем случае математического моделирования динамическая модель движения всех объектов может быть выполнена в рамках единого программного модуля, однако на практике возникает потребность в распределении процесса динамического моделирования.

Во-первых, такая ситуация возникает, когда целью этапа математического моделирования является подготовка к полунатурному моделированию функционирования системы управления ВС. Для организации исследований такого стенда полунатурного моделирования необходимо включение в его состав упрощенной (по сравнению с детализацией полунатурной модели) математической модели внешней среды, окружающей «свой самолет». Переход на такую конфигурацию моделирующего стенда технически упрощается, если уже на стадии математического моделирования будет реализована его будущая структура (отдельные модели ВС и внешней среды, параллельно взаимодействующие по согласованному протоколу).

Во-вторых, математическая или полунатурная модель ВС могут функционировать в различных организациях, в составе различных моделирующих комплексов. Для возможности сравнения получаемых результатов необходимо, чтобы условия моделирования были одинаковыми. Это обеспечивается, если для моделирования используется единая модель внешней обстановки, территориально расположенная в любой из взаимодействующих организаций.

В-третьих, распределение необходимо при разработке бортового программного обеспечения (ПО) с использованием специализированных комплексов, которые предназначены для удовлетворения повышенных требований по безопасности (автопилоты, системы управления полетом и двигателем, тормозные системы и т.д.).

При создании ПО бортовых систем и оборудования, которые выполняют предписанные им функции с уровнем доверия к безопасности, удовлетворяющих требованиям к летной годности, необходимо следовать Квалификационным требованиям, часть 178В (КТ178В) [1].

Квалификационные требования часть 178В[1] являются требованиями к ПО бортовой аппаратуры и систем при сертификации авиационной техники. Основой настоящих Квалификационных требований КТ-178В [1] является документ RTCA/DO-178В [2]. (КТ-178В [1] развивает требования и подходы, приведенные в ранее выпущенном документе КТ-178А). При создании ПО для наземного оборудования возникают схожие проблемы, для решения которых на основе RTCA/DO-178В [2] создан норматив RTCA/DO-278В [1], который ориентирован не на бортовые приложения и системы. Проводимые в настоящее время исследования связаны с разработкой, тестированием и анализом функций и отдельных процедур поддержки принятия решений. Функция поддержки принятия решений предназначена для

предупреждения экипажа ВС о возможности опасных сближений с другими ВС в воздухе и выработки рекомендации к изменению параметров полета ВС для устранения этих конфликтов с учетом действующих правил полетов в используемом воздушном пространстве.

1. Комплекс средств разработки программного обеспечения SCADE

Для создания бортового ПО использовался комплекс средств разработки программного обеспечения (SCADE SUITE TM), который квалифицирован по DO-178B [2] уровня А.

Верификатор дизайна SCADE Suite Design Verifier предоставляет проектировщику возможность всестороннего исследования при поиске «краевых» ошибок, практически не детектируемых традиционными методами тестирования, и исправления их на ранней стадии проекта. Тестовое покрытие моделей SCADE (Model Test Coverage, MTC) оценивает полноту тестовой процедуры, разработанной на основе требований, и помогает достичь 100%-го покрытия по MC/DC (Modified Condition/Decision Coverage). MTC является средством верификации, квалифицированным по DO-178B [2].

Компонент верификации компилятора SCADE (Compiler Verification Kit, CVK) позволяет провести верификацию применяемого компилятора языка С на корректность компиляции кода, сгенерированного в SCADE KCG (Qualified Code Generator). Компонент верификации подтверждает корректность работы сгенерированного кода по утвержденной методике, описанной в руководстве CAST 12. Программный комплекс SCADE представляет собой среду, предназначенную для разработки встроенного ПО с повышенными требованиями по безопасности. Комплекс SCADE:

- представляет собой основанную на моделях среду для разработки встроенного ПО;
- обеспечивает графический ввод и моделирование спецификаций графов потоков данных и конечных автоматов с повышенной надежностью;
- предоставляет уникальную возможность формальной верификации атрибутов безопасности на уровне спецификаций;
- автоматически генерирует портируемый встроенный код промышленного качества;
- устраняет необходимость в тестировании кода нижнего уровня;
- автоматически генерирует проектную документацию;
- дает возможность проводить распределенную работу.

2. Особенности SCADE при разработке бортового программного обеспечения

- Массивы должны быть статическими, недопустимы массивы переменной длины.
- Работу с массивами обеспечивают девять операторов цикла, которые производят действия с элементами массивов каждый по своему алгоритму.
- Задать количество итераций можно только константой.
- Обработка массивов с размерностью, не равной количеству итераций, требует преобразование скаляра в массив (рис. 1).

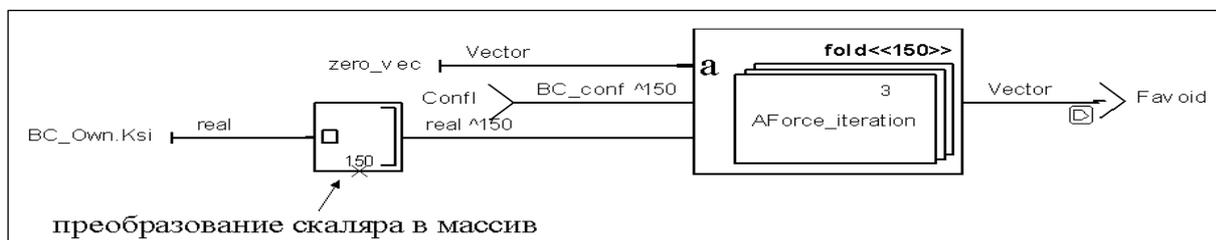


Рис. 1. Лишнее действие – преобразование скаляра в массив

- В зависимости от выбора ОПЕРАТОРА, работающего с элементами массива, выходами могут быть как скаляр, так и массив той же размерности, что и количество итераций.
- При отладке проекта компилятор помогает выявлять недостатки алгоритма, но не всегда явно указывает ветвь проекта, где обнаружена ошибка, что затрудняет поиск ошибки.
- Всем внутренним переменным и выходным параметрам должно быть установлено значение по умолчанию, в противном случае компилятор выдаст ошибку.

- Для простых функций, которые в коде занимают пару строк, в SCADE нужно использовать несколько операторов, на что затрачивается больше времени (рис. 2).

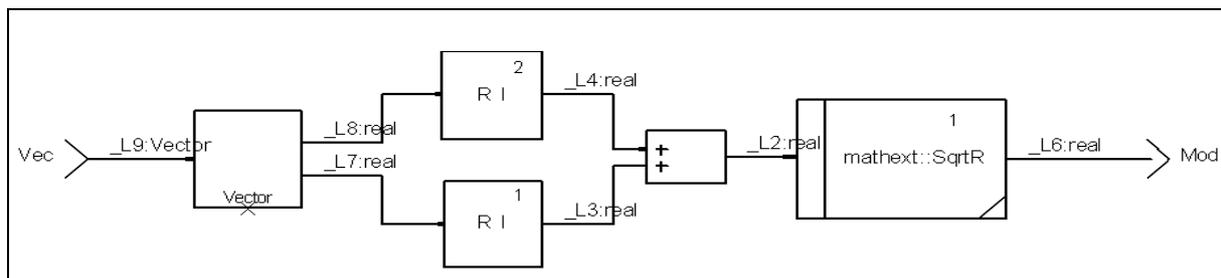


Рис. 2. Вычисление модуля вектора в SCADE

- Для написания отчета в SCADE есть специальная функция генерации отчетов. В этом отчете выводятся блок-схемы алгоритмов, дерево вызовов, описание типов, констант и т.д.
- Использование инструмента Textual expression требует особого внимания при работе с отрицательными числами.
- Не допускаются математические операции между величинами различных типов.
- Возникает трудность, когда необходимо в цикле работать с двумя и более элементами массива (рис. 3). Необходимо дополнительное преобразование массива в двумерный массив, чтобы производить действия с любыми элементами массива (например, соседними).

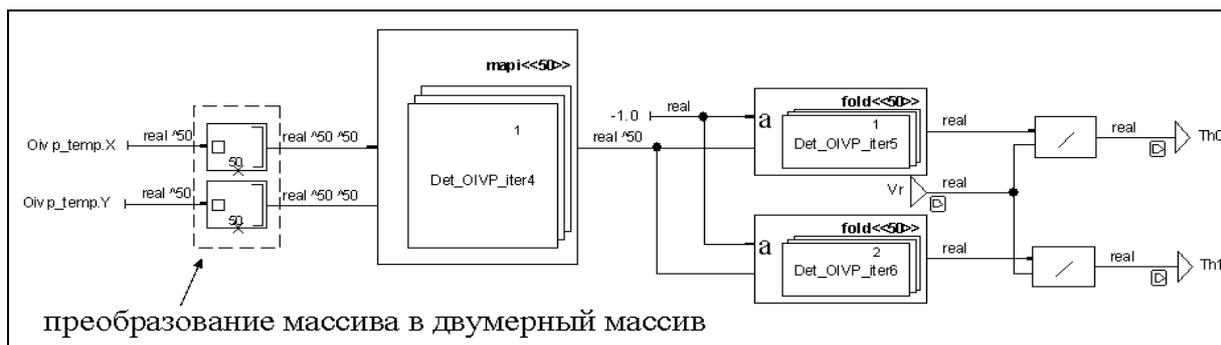


Рис. 3. Дополнительное преобразование массива в двумерный массив

- При генерации кода предусмотрены три уровня оптимизации.
 1. При минимальном уровне оптимизации при генерации кода используются только внутренние переменные SCADE, каждое действие выполняется в отдельной строке, объем кода возрастает.
 2. При максимальном уровне оптимизации генератор кода использует как переменные, объявленные разработчиком, так и свои собственные переменные.

Приведем обнаруженные нами достоинства и недостатки при создании бортовой функции поддержки принятия решений с использованием комплекта средств разработки ПО (SCADE SUITE TM).

Недостатки

- Для анализа и поиска ошибок в автоматически генерируемом коде необходимо привлечение хорошо подготовленного специалиста.
- По сравнению с “классическим” программированием объем автоматически сгенерированного кода существенно больше, чем аналогичный код ручного исполнения.
- Для каждого оператора создается свой заголовочный файл и свой исполняемый файл, что также не всегда удобно для использования.
- Практически полное отсутствие обучающей литературы на русском языке, а в пособиях на английском только поверхностное описание, нет конкретики.
- Генератор кода автоматически создает собственные входные структуры и использует именно их, а не созданные разработчиком (SCADE определяет структуры через структуры разработчика).
- SCADE использует свои локальные переменные, вместо созданных разработчиком при графической реализации алгоритма.

Преимущества

- Графическое представление модулей наглядно, по существу, представляется блок-схемой.
- Основным преимуществом является квалифицированный, стандартизированный генератор кода, что существенно упрощает сертификацию ПО, т.е. достаточно доказать соответствие графических модулей, созданных в SCADE, аналогичным схемам алгоритмов.
- Четкая структурированность языка, не допускается никаких математических операций между величинами разных типов, каждый выход должен быть инициализирован и определен.

Заключение

Неоспоримым преимуществом SCADE, несмотря на перечисленные недостатки, является портируемый и безопасный код промышленного качества, который уже может быть квалифицирован по стандарту DO-178B [2]. В результате работ при помощи SCADE создан прототип функции поддержки принятия решений, работоспособность которого была проверена в МДВО.

ЛИТЕРАТУРА

1. КТ-178В. Требования к программному обеспечению бортовой аппаратуры и систем при сертификации авиационной техники.
2. DO-178B / ED12B, Software Considerations in Airborne Systems and Equipment Certification // RTCA / EUROCAE (December 1992).
3. DO-278 / ED-109 Guidelines for Communication, Navigation, Surveillance, and Air Traffic Management (CNS/ATM) Systems Software Integrity Assurance // RTCA / EUROCAE 2002.
4. SCADE Suite™ Documentation, Product Manuals.

USE OF SPECIAL SOWTWARE PACKAGE SCADE FOR SOWTWARE DEVELOPMENT OF ON-BOARD SUPPORT DECISION-MAKING FUNCTION

Bakhirkin M.V., Kanadin V.N., Hristoforov G.U.

In this article authors share own experience in using of special software package SCADE for development of on-board support decision-making function and try to find pluses and minuses this approach.

Key words: specialized program complex, the onboard software.

Сведения об авторах

Бахиркин Михаил Васильевич, 1983 г.р., окончил МАИ (2006), инженер 2-й категории ФГУП "ГосНИИАС", автор 4 научных работ, область научных интересов – математические методы моделирования, информационные технологии, вычислительные системы и сети, управление проектами.

Канадин Владимир Николаевич, 1971 г.р., окончил МИФИ (1994), начальник сектора ФГУП "ГосНИИАС", автор более 12 научных работ, область научных интересов – математическое моделирование, системы автоматизированного проектирования, системы управления.

Христофоров Григорий Юрьевич, 1981 г.р., окончил МАИ (2004), инженер 1-й категории ФГУП "ГосНИИАС", автор 5 научных работ, область научных интересов – математические методы моделирования, распределенные системы, организация вычислений, человеко-машинный интерфейс.

УДК 629.7.05.07:681.5

ВОПРОСЫ ПОСТРОЕНИЯ КОМПЛЕКСНОЙ СИСТЕМЫ ОБРАБОТКИ АЭРОНАВИГАЦИОННЫХ ДАННЫХ

В.В. СОЛОМЕНЦЕВ, Н.В. РОМАНОВ

Рассматриваются вопросы создания распределенной системы сбора, хранения, распространения аэронавигационных данных на базе глобальных сетевых технологий.

Ключевые слова: распределенная система, сбор, хранения, распространение данных, аэронавигационные данные, сетевые технологии.

Введение

В Аэронавигационной систем России создается комплексная система сбора, хранения и распространения аэронавигационных данных. Она создается как комплексное решение, обеспечивающее переход от службы аэронавигационной информации к управлению аэронавигационной информацией.

Учитывая комплексный характер системы, ее сильное влияние на все авиационное сообщество, выбор варианта построения является ответственным шагом. В статье рассмотрены различные варианты построения, каждый из которых обладает определенными достоинствами и недостатками.

1. Технология подготовки аэронавигационных данных

Общий порядок подготовки аэронавигационной информации определен Федеральными авиационными правилами [1] и другими нормативными документами. Органы аэронавигационной информации, созданные в рамках службы аэронавигационной информации (САИ), выполняют определенные функции и несут ответственность за организацию предоставления аэронавигационной информации пользователям воздушного пространства Российской Федерации (рис. 1).

Обязанности по подготовке элементов объединенного пакета аэронавигационной информации распределены между структурными подразделениями Росавиации, подведомственными ей организациями ФГУП «ЦАИ», ФГУП «Госкорпорация по ОрВД» и территориальными органами Росавиации. Информация, с которой работает Центр аэронавигационной информации, может носить постоянный или временный характер. Источниками информации об изменениях аэронавигационной обстановки в районе аэродрома или на аэродроме являются авиапредприятия/аэропорты со своими службами. В части, касающейся изменений в структуре воздушного пространства или на воздушных трассах, источником информации являются филиалы ФГУП «Госкорпорация по ОрВД» и ее службы.

Информация, подготовленная в виде проектов поправок к документам АНИ или проектов НОТАМ постоянного характера по изменениям на аэродроме или в районе аэродрома, согласованная в установленном порядке, представляется на регистрацию руководителям территориальных органов Росавиации, а затем – во ФГУП «ЦАИ» для издания и рассылки.

Проекты поправок в АИП России (документы АНИ) или проекты НОТАМ постоянного характера об изменениях в структуре воздушного пространства, подготовленные филиалами ФГУП «Госкорпорация по ОрВД», направляются на регистрацию в территориальные органы Росавиации.

Зарегистрированные проекты поправок или НОТАМ по изменениям в структуре воздушного пространства из территориальных управлений возвращаются в филиалы ФГУП «Госкорпорация по ОрВД» и направляются на согласование во ФГУП «Госкорпорация по ОрВД», а после согласования направляются в Росавиацию на утверждение. Утвержденные поправки к АИП или НОТАМ направляются во ФГУП «ЦАИ» для издания и рассылки.

Проекты НОТАМ, которые требуется издать незамедлительно, направляются во ФГУП «ЦАИ» за подписью лиц, имеющих право подписи проектов НОТАМ.

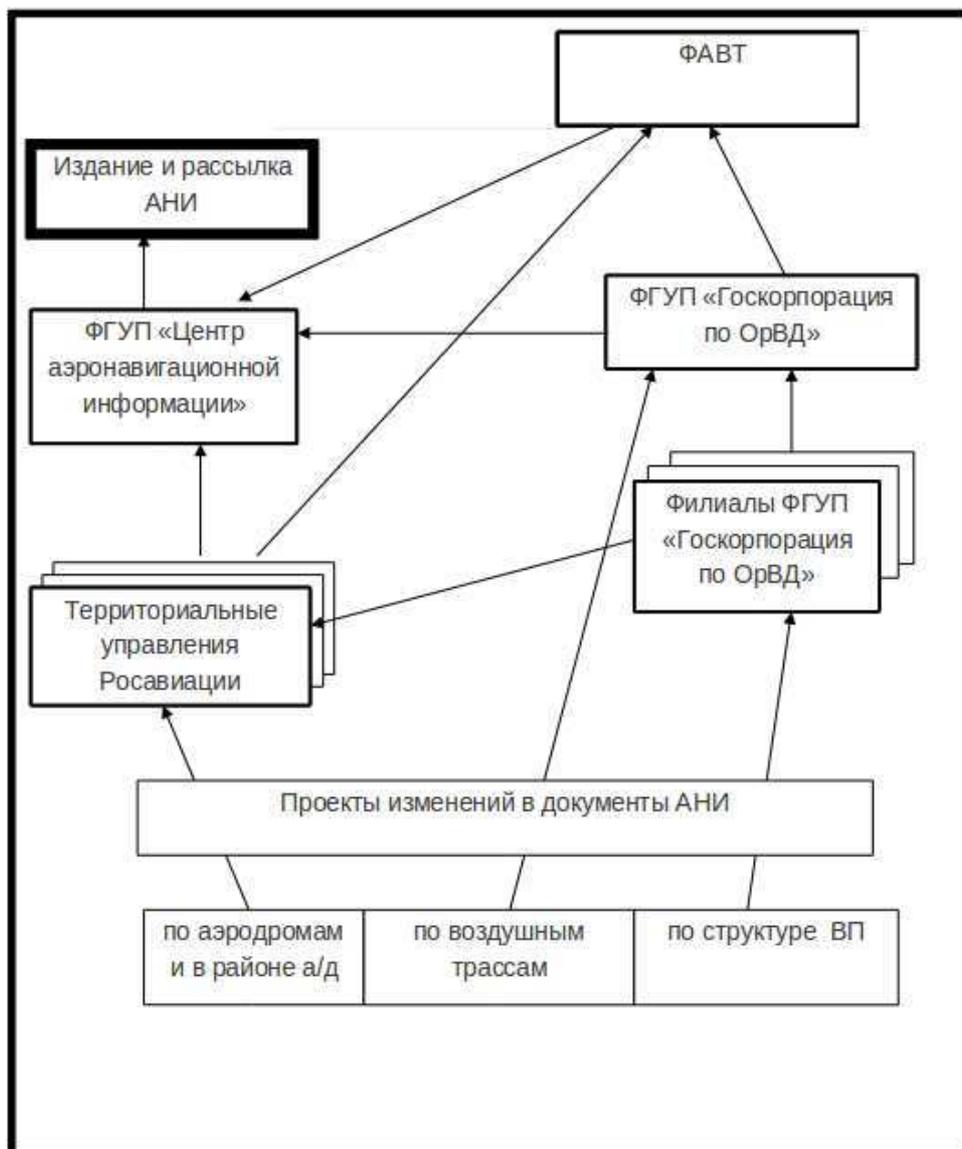


Рис. 1. Структура службы аэронавигационной информации

2. Постановка задачи

Комплексная системы сбора, хранения, распространения аэронавигационной информации должна обеспечить весь технологический цикл подготовки данных, а также распространение аэронавигационных данных потребителям.

С целью унификации систем аэронавигационной информации с учетом накопленного опыта и перспектив внедрения концепции коллективного принятия решений (CDM), международным сообществом разработана аэронавигационная информационная модель обмена АИХМ (Aeronautical Information Exchange Model). Все новые разработки в области создания систем управления аэронавигационной информацией базируются на ее основе. Отличительной особенностью АИХМ является использование временной модели хранения данных, позволяющей хранить информацию о свойствах объекта с учетом всех производимых изменений и периода их действия. Применение временной модели дает возможность организовать автоматическое обновление (синхронизацию) данных между различными хранилищами на основе информации не только о дате создания соответствующих физических записей, но в первую очередь на основе изменений, внесенных собственно в описание объекта аэронавигационной информации.

Вместе с тем, АИХМ оставляет полную свободу разработчикам систем в части архитектурных принципов организации систем управления АНИ на национальном уровне. Единственным условием является возможность обмена данными между системами в формате АИХМ.

Рассмотрим возможные варианты организации системы для Российской Федерации.

3. Варианты организации системы управления аэронавигационной информацией для Российской Федерации

Как было отмечено выше, пользователи системы могут быть условно разделены на четыре основные группы: поставщики АНИ – оригинаторы, органы контроля (регистрации, утверждения информации), Центр аэронавигационной информации, организующий хранение и предоставление АНИ, и наконец, собственно пользователи аэронавигационной информации и данных.

Система условно разделена на две части: подсистема подготовки данных и подсистема предоставления данных.

Первая подсистема может быть организована по централизованному и децентрализованному принципу.

При децентрализованной организации каждый оригинатор и органы контроля (регистрации, утверждения информации) оснащены собственными хранилищами локальных данных, которые по мере обновления информации оригинаторами синхронизируются (обновляют) данные в центральном хранилище ЦАИ.

При централизованной организации соответствующие локальные хранилища организуются виртуально в ЦАИ и по мере изменения оригинаторами и одобрения контролирующими органами обновляют информацию в центральном хранилище. Доступ к виртуальным хранилищам может быть организован с использованием web-технологий. Механизм обновления данных в центральном хранилище идентичен для обоих вариантов и соответствует модели АИХМ.

Оценивая достоинства и недостатки вариантов построения подсистемы подготовки данных, можно отметить следующее.

Децентрализованный вариант требует значительных затрат на оснащение клиентских подсистем хранилищами данных и в целом имитирует существующую бумажную технологию документооборота в службе АНИ. Вместе с тем, использование механизмов синхронизации данных позволяет обмениваться только "разницей" предшествующих и обновленных данных, что существенно снижает нагрузку на сетевую инфраструктуру. Кроме того, органы, осуществляющие подготовку изменений данных (аэропорты, центры ОВД) как правило, являются и потребителями данных. В этом случае локальное хранилище может использоваться как для хранения официальных аэронавигационных данных, необходимых для производственной деятельности, так и для подготовки изменений АНИ.

Варианты построения подсистемы предоставления данных могут отличаться по признаку централизованное хранилище/распределенное хранилище.

В первом случае все аэронавигационные данные должны храниться в центральном хранилище ЦАИ и предоставляться по подписке потребителям на основе механизма синхронизации по определенным профилям пользователей. Такой вариант хранения и управления данными в одном месте также упрощает решение вопросов безопасности и целостности данных.

Вместе с тем, по мере роста объемов аэронавигационных данных, включения в состав АНИ всех дополнительных видов вспомогательной информации, такой механизм становится все менее эффективным и представляется целесообразным разделение функций хранения и функций управления АНИ. Функции хранения данных могут быть предоставлены оригинаторам, а функции управления предоставлением данных могут быть централизованными. Таким образом, центральный сервер управления предоставлением данных будет содержать только тот объем информации, который необходим для поиска места хранения необходимых данных и определения политики доступа к данным и безопасности. Сами же данные могут храниться в локальных хранилищах оригинаторов (или групп оригинаторов) и пересылаться пользователю на основе механизмов синхронизации минуя какие-либо централизованные узлы сети. В качестве основных достоинств такого подхода стоит отметить снижение нагрузки на узкие места сетей передачи данных, а также возможность гибкой модификации и расширения состава и видов аэронавигационных данных. В то же время использование принципа децентрализованного хранения потребует тщательно проработки механизмов обеспечения безопасности данных.

Выводы

Возможны различные подходы к построению системы управления аэронавигационной информацией. Наиболее перспективным представляется подход, основанный на разделении функций управления данными и хранения данных с применением механизма распределенного хранения информации.

Оценка вариантов построения системы управления аэронавигационной информацией является актуальной задачей и должна учитывать необходимость своевременной доставки потребителям корректной и целостной информации.

ЛИТЕРАТУРА

1. ФАП "Аэронавигация и аэронавигационное обеспечение полетов": приказ федеральной службы воздушного транспорта России от 30 декабря 1999 г. - № 151.

PROBLEMS OF CONSTRUCTION INTEGRATED SYSTEM OF PROCESSING AERONAUTICAL DATA

Solomentsev V.V., Romanov N.V.

Reviewed problems of creation distributed system of collection, storing, diffusion air-navigation data based on in global networks technologies.

Key words: distributed system, collect, store, data diffusion, net technology, queuing network.

Сведения об авторах

Соломенцев Виктор Владимирович, 1957 г.р., окончил МИЭМ (1979), доктор технических наук, профессор, заместитель генерального директора ОАО "НТЦ Промтехаэро", заведующий кафедрой МГТУ ГА, автор более 150 научных работ, область научных интересов - системы связи, навигации, наблюдения, моделирование.

Романов Никита Вячеславович, 1985 г.р., окончил МГТУ ГА (2007), аспирант, старший преподаватель МГТУ ГА, автор 7 работ, область научных интересов - вычислительные системы, аэронавигация.

ББК 05
Н 34
Св. план 2011 г.

Научный Вестник МГТУ ГА
№ 169

ISBN 978-5-86311-787-4

Свидетельство о регистрации в Министерстве Российской Федерации по делам печати, телерадиовещания и средств массовых коммуникаций № 77-17906 от 31 марта 2004 г.

Редакторы Т.М. Приорова, И.В. Вилкова
Компьютерная верстка Т.Н. Котиковой

Подписано в печать 22.06.11 г.

Печать офсетная
15,75 усл. печ. л.

Формат 60x90/8
Заказ № 1243/

14,46 уч.-изд. л.
Тираж 100 экз.

Московский государственный технический университет ГА
125993 Москва, Кронштадтский бульвар, д. 20
Редакционно-издательский отдел
125493 Москва, ул. Пулковская, д. 6а

Подписной индекс в каталоге Роспечати 84254
© Московский государственный
технический университет ГА, 2011