

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ

8(204)
2013

ТЕОРЕТИЧЕСКИЙ И ПРИКЛАДНОЙ НАУЧНО-ТЕХНИЧЕСКИЙ ЖУРНАЛ

Издается с ноября 1995 г.

УЧРЕДИТЕЛЬ
Издательство "Новые технологии"

СОДЕРЖАНИЕ

МОДЕЛИРОВАНИЕ И ОПТИМИЗАЦИЯ

- Струченков В. И. Математические модели и методы оптимизации в системах проектирования трасс автомобильных дорог 2
Мусихина Е. А., Дарманская Т. А. Исследование рельефа и построение триангуляционной модели 8
Чувашева Е. С., Чувашев С. Н. Выбор рациональных характеристик высокоскоростных летательных аппаратов разных масштабов на основе комплексной математической модели 12

ВЫЧИСЛИТЕЛЬНЫЕ СИСТЕМЫ И СЕТИ

- Шерстюков О. Н., Рябченко Е. Ю., Мартычук С. Л. Радиомодем LPD-диапазона для задач телеметрии 17
Иванова Е. М. Сравнительная оценка производительности вычислительных систем . 22
Бобков С. Г., Чибисов П. А. Повышение качества тестирования высокопроизводительных микропроцессоров методами встречного тестирования с анализом функционального тестового покрытия выделенных приложений. 26

СЕТИ И СИСТЕМЫ СВЯЗИ

- Агаев Б. С., Пашаев Ф. Г. Метод оценки качества речи в корпоративных VoIP-сетях. 34

СИСТЕМЫ АВТОМАТИЗИРОВАННОГО ПРОЕКТИРОВАНИЯ

- Паршикова Е. А. Возможные подходы к автоматизации проектирования систем поддержки принятия управленческих решений 40
Бульонков М. А., Филаткина Н. Н. Ситуационный анализ в системе транспортного прогнозирования МИКС-ПРОСТОР 45

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ В МЕДИЦИНЕ

- Садыхов С. С., Буланова Ю. А., Захарова Е. А. Алгоритм выделения области кисты на малоконтрастных маммограммах 53

ГЕОИНФОРМАЦИОННЫЕ СИСТЕМЫ И СИСТЕМЫ ПРИРОДОПОЛЬЗОВАНИЯ

- Солодовниченко М. А., Замятин А. В. Алгоритм распространения лесного пожара с использованием модели Ротермела и клеточных автоматов 58
Бекирова Л. Р. Вопросы минимизации влияния изменчивости атмосферного аэрозоля на информативность спутниковых мультиспектральных изображений . . 64

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ В ОБРАЗОВАНИИ

- Залещанский Б. Д., Свиридов А. П., Шалобина О. А., Шалобина Е. А. Вероятностно-статистические стратегии обеспечения качества подготовки персонала социотехнических систем. 67
Contents 71
Приложение Гринченко С. Н., Щапова Ю. Л. Информационные технологии в истории Человечества

Главный редактор:

СТЕМПКОВСКИЙ А. Л.,
акад. РАН

Зам. главного редактора:

ДИМИТРИЕНКО Ю. И., д. ф.-м. н.
ФИЛИМОНОВ Н. Б., д. т. н.

Редакционный совет:

БЫЧКОВ И. В., акад. РАН
ЖУРАВЛЕВ Ю. И., акад. РАН
КУЛЕШОВ А. П., акад. РАН
ПОПКОВ Ю. С., чл.-корр. РАН
РУСАКОВ С. Г., чл.-корр. РАН
РЯБОВ Г. Г., чл.-корр. РАН
СОЙФЕР В. А., чл.-корр. РАН
СОКОЛОВ И. А., акад. РАН
СВЕТИН Н. В., д. ф.-м. н.
ЧАПЛЫГИН Ю. А., чл.-корр. РАН
ШАХНОВ В. А., чл.-корр. РАН
ШОКИН Ю. И., акад. РАН
ЮСУПОВ Р. М., чл.-корр. РАН

Редакционная коллегия:

АВДОШИН С. М., к. т. н.
АНТОНОВ Б. И.
БАРСКИЙ А. Б., д. т. н.
ВАСЕНИН В. А., д. ф.-м. н.
ГАЛУШКИН А. И., д. т. н.
ДОМРАЧЕВ В. Г., д. т. н.
ЗАГИДУЛЛИН Р. Ш., к. т. н.
ЗАРУБИН В. С., д. т. н.
ИВАННИКОВ А. Д., д. т. н.
ИСАЕНКО Р. О., к. т. н.
КАРПЕНКО А. П., д. ф.-м. н.
КОЛИН К. К., д. т. н.
КУЛАГИН В. П., д. т. н.
КУРЕЙЧИК В. М., д. т. н.
КУХАРЕНКО Б. Г., к. ф.-м. н.
ЛЬВОВИЧ Я. Е., д. т. н.
МАЛЬЦЕВ П. П., д. т. н.
МИХАЙЛОВ Б. М., д. т. н.
НЕЧАЕВ В. В., к. т. н.
ПАВЛОВ В. В., д. т. н.
СОКОЛОВ Б. В., д. т. н.
УСКОВ В. Л., к. т. н.
ФОМИЧЕВ В. А., д. т. н.
ЧЕРМОШЕНЦЕВ С. Ф., д. т. н.
ШИЛОВ В. В., к. т. н.

Редакция:

БЕЗМЕНОВА М. Ю.
ГРИГОРИН-РЯБОВА Е. В.
ЛЫСЕНКО А. В.
ЧУГУНОВА А. В.

Информация о журнале доступна по сети Internet по адресу <http://novtex.ru/IT>.
Журнал включен в систему Российского индекса научного цитирования.
Журнал входит в Перечень научных журналов, в которых по рекомендации ВАК РФ должны быть опубликованы научные результаты диссертаций на соискание ученой степени доктора и кандидата наук.

УДК 519.857

В. И. Струченков, д-р техн. наук, проф.,
e-mail: str1942@mail.ru,
Московский государственный технический
университет радиотехники, электроники
и автоматики (МГТУ МИРЭА)

Математические модели и методы оптимизации в системах проектирования трасс автомобильных дорог¹

Рассматривается задача оптимизации положения трассы новой автомобильной дороги в заданной области варьирования. Предлагаются новые математические модели и алгоритмы проектирования продольного профиля по заданным вариантам плана трассы. Приводится способ сведения данной задачи к решенной ранее аналогичной задаче проектирования продольного профиля новой железной дороги. Проектирование продольного профиля рассматривается как элемент многоэтапного проектного процесса во взаимосвязи с другими проектными задачами.

Ключевые слова: трасса, план, продольный профиль, нелинейное программирование, целевая функция, приведенный градиент

Введение

Целесообразность оптимизации проектов строительства таких дорогостоящих сооружений, как железные и автомобильные дороги с использованием математических моделей, алгоритмов и компьютерных программ была установлена около 50 лет тому назад [1–3].

В частности, в условиях пересеченного рельефа и сложной геологии затраты на строительство и последующую эксплуатацию могут быть существенно снижены при оптимальном расположении трассы проектируемой дороги на местности. Однако сложность формализации творческого процесса выработки проектного решения в адекватных математических моделях не позволила создать математически корректные алгоритмы оптимизации трассы как пространственной кривой.

¹ Данная статья является продолжением статьи "Математические модели и методы оптимизации в системах проектирования трасс новых железных дорог". См. журнал "Информационные технологии", № 7, 2013. С. 10–17.

Не оправдались большие надежды на использование для этой цели динамического программирования по тем же причинам, что и при разработке методов автоматизированного проектирования трасс новых железных дорог [4, 5].

Существенные успехи были достигнуты при решении частной задачи: проектирование оптимального продольного профиля при заданном положении трассы в плане [6]. В 80-х годах прошлого века широкое практическое применение получила система "Профиль-2А" для проектирования новых дорог и ее аналог система "Профиль-2Р" для проектирования реконструкции. Проектная задача была формализована как задача нелинейного программирования с линейной системой ограничений. Задача была решена с помощью модификации известного в математике метода проекции градиента с использованием специфических особенностей системы ограничений [6].

Решение этой частной задачи позволяет не только объективно сравнивать варианты плана трассы, назначаемые проектировщиками, но имеет и самостоятельное значение, так как при проектировании дорог в обжитых районах план трассы, как правило, определяется условиями землепользования.

К сожалению, при разработке современных САПР линейных сооружений на основе новых технических средств и соответственно при наличии новых вычислительных возможностей положительный опыт компьютерного проектирования продольного профиля автодорог был полностью утрачен. Даже в наиболее развитых САПР, таких как CAD-1 [7] или ее российских аналогов Toromaic Robur [8] и Geonics [9] положение трассы и в плане, и в продольном профиле назначается вручную. Проектирование плана и продольного профиля рассматривается как геометрическая задача в отрыве от решения других проектных задач, таких как проектирование поперечных профилей земляного полотна, распределение земляных масс, проектирование водопропускных и других искусственных сооружений и др.

В качестве дополнения к проектированию продольного профиля вручную в упомянутых и в других системах предлагаются простейшие математические модели и алгоритмы оптимизации, в основном эвристические. В частности, в CAD-1 и Toromaic Robur наряду с проектированием продольного профиля вручную предлагается поиск наилучшей среднеквадратической аппроксимации профиля земли с помощью динамического программирования.

Очевидно, что при этом не учитываются конструкции поперечных профилей земляного полотна, геология, наличие искусственных сооружений и др.

Известно, что в одних и тех же условиях, располагая одной и той же информацией, различные специалисты предлагают различные варианты проектных решений. Рассмотрение ограниченного числа интуитивно назначаемых вариантов не гарантирует близость к оптимуму конечного результата такого процесса. В то же время известно, что относительно небольшие изменения положения трассы на местности могут приводить к существенным изменениям затрат на строительство и эксплуатацию дороги [10, 11].

Следовательно, проблема разработки адекватных математических моделей и математически корректных алгоритмов оптимизации трасс новых дорог остается актуальной. Это является наиболее перспективным направлением совершенствования САПР линейных сооружений.

Новые технические возможности позволяют перейти к разработке интеллектуальных систем проектирования трасс линейных сооружений, в которых компьютер используется не только для решения расчетных, графических и оформительских задач, но и для выработки проектных решений на основе комплексных математических моделей и математически корректных алгоритмов оптимизации.

Целью настоящей статьи является изложение усовершенствованных моделей целевой функции и ограничений, а также алгоритмов нелинейного программирования для проектирования оптимального продольного профиля новых автомобильных дорог в расчете на вычислительные возможности современных общедоступных персональных компьютеров.

Содержательная постановка задачи

Трасса дороги — это гладкая трехмерная кривая, состоящая из элементов заданного вида и удовлетворяющая целому ряду ограничений.

Традиционно искомая трехмерная кривая представляется в виде двух плоских кривых: план и продольный профиль. *План* — это проекция трассы на координатную плоскость XOY , а *продольный профиль* — это зависимость аппликаты Z от текущей длины s в плане. Задача проектирования оптимальной трассы превращается в две взаимосвязанные задачи: проектирование плана и проектирование продольного профиля.

На положение трассы влияют рельеф земли, геологические, гидрологические, климатические и другие условия. В настоящее время это влияние пока не формализовано в адекватных математических моделях.

Продольный профиль состоит из элементов определенного вида, на параметры которых накладываются различные ограничения. Могут использоваться отрезки окружностей, сопрягаемые отрезками прямых. Возможно проектирование параболиче-

скими элементами, включая, как частный случай, отрезки прямых.

Требуется найти такую проектную линию, которая удовлетворяет целому ряду технических ограничений и обеспечивает минимум заданного критерия оптимальности. При проектировании продольного профиля в условиях пересеченного рельефа таким критерием могут быть объемы земляных работ или строительные затраты.

Число элементов проектной линии неизвестно, что является серьезным препятствием при формализации задачи. Поэтому ее решение разбивается на ряд этапов.

Сначала ставится цель определить число элементов и их примерное расположение, т. е. определить размерность задачи и построить начальное приближение, а затем решить оптимизационную задачу полученной размерности.

Каждый из этих этапов в общем случае также требует неоднократного повторения расчетов по мере уточнения исходных данных и решения смежных проектных задач. Естественно, что возникают иерархия математических моделей целевой функции и ограничений, а также необходимость использования нескольких оптимизационных алгоритмов.

Формализация задачи и построение начального приближения

На первом этапе проектная линия рассматривается в виде ломаной $Z_j(s_j)$, абсциссы вершин которой s_j совпадают с абсциссами вершин профиля земли $H_j(s_j)$.

Переменными являются $h_j(s_j) = Z_j(s_j) - H_j(s_j)$. Далее $Z_j(s_j)$ называются проектными отметками, а h_j — рабочими отметками.

Ограничения на минимальные длины элементов на данном этапе не учитываются. Ограничения на продольные уклоны (дискретный аналог ограничения на первую производную) и разности смежных уклонов (дискретный аналог ограничения на кривизну), а также ограничения на проектные и рабочие отметки в отдельных точках приводят к системе линейных неравенств, которая полностью аналогична системе ограничений, возникающей при проектировании продольного профиля новых железных дорог [5].

В качестве критерия оптимальности принят суммарный объем земляных работ. Однако его вычисление возможно лишь при известных конструкциях поперечных профилей земляного полотна в насыпях и в выемках, что в свою очередь требует наличия данных о рабочих отметках, т. е. наличия проектной линии. Поэтому конструкции проектных поперечных профилей для первого расчета принимаются в форме трапеций. Ширина поперечника в насыпях B_n определяется заданным числом полос движения, в выемках при задании ширины B_v дополнительно используются данные о типовых кюветах.

Если нет данных о свойствах грунтов, то показатели крутизны откосов в насыпях m_n и в выемках m_b задаются ориентировочно, например по 1,5.

Объем земляных работ V_j на j -м элементе длиной L_j через площадь поперечника $F(h(l))$, где h — рабочая отметка, а l — расстояние от начала элемента, представляется формулой

$$V_j = \int_0^{L_j} F(h(l)) dl. \quad (1)$$

Для принятого на данном этапе представления продольного профиля

$$h(l) = h_j + l(h_{j+1} - h_j)/L_j. \quad (2)$$

Площадь $F(h)$ для насыпей $F_n(h) = B_n h + m_n h^2$, а для выемок $F_b(h) = -B_b h + m_b h^2$.

График $F(h)$ представлен на рис. 1. Для непрерывности производной в окрестности нуля вписывается третья парабола [12].

Производные по неизвестным h_j вычисляются с использованием формул (1), (2). Если обозначить параметры параболы, соответствующей h_j через a , b и c , то

$$F(h) = ah^2 + bh + c$$

и

$$\begin{aligned} \partial V_j / \partial h_j &= \int_0^{L_j} (2a(h_j + l(h_{j+1} - h_j)/L_j) + b) \partial h / \partial h_j dl = \\ &= (1/3a(h_{j+1} + 2h_j) + b/2)L_j; \end{aligned} \quad (3)$$

$$\begin{aligned} \partial V_j / \partial h_{j+1} &= \int_0^{L_j} (2a(h_j + l(h_{j+1} - h_j)/L_j) + b) \partial h / \partial h_{j+1} dl = \\ &= (1/3a(h_j + 2h_{j+1}) + b/2)L_j. \end{aligned} \quad (4)$$

Вместо использовавшегося ранее алгоритма проекции градиента используется алгоритм приведенного градиента и практически та же программа, что и при проектировании новых железных дорог [5].

Полученная проектная линия в виде ломаной рассматривается как ось зоны поиска при дальнейших уточнениях и используется далее для совмест-

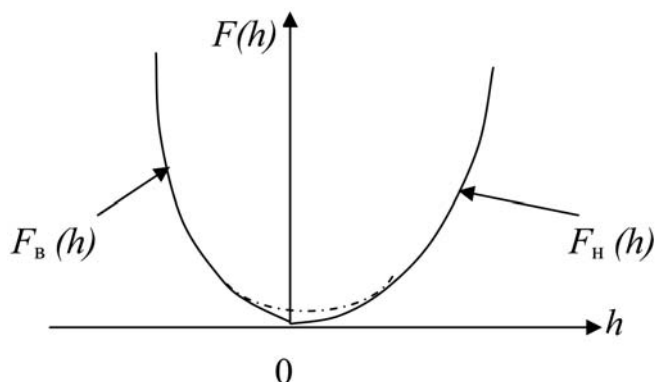


Рис. 1. Зависимость $F(h)$ для простых конструкций поперечников

ного проектирования продольного и поперечных профилей [5].

Реальные конструкции поперечных профилей земляного полотна зависят от геологии и поперечных профилей земли. В системе проектирования используются типовые проектные конструкции, для которых на каждом переломе продольного профиля земли в зоне поиска с заданным шагом Δ изменения рабочей отметки проектируются поперечные профили и вычисляются площади насыпей и выемок. Приближенные зависимости $F_i(h)$ получаются параболической аппроксимацией полученных значений отдельно для насыпей и выемок и используются для продолжения оптимизации. Может оказаться, что в назначенной зоне поиска не для всех значений рабочих отметок есть насыпь (или выемка). Тогда график соответствующей зависимости $F_i(h)$ состоит из горизонтального и параболического участков (рис. 2). При необходимости расчет повторяется при сужении зоны поиска и уменьшении шага Δ .

На каждом переломе профиля земли получают свои зависимости для площадей насыпей и выемок, что не позволяет использовать формулы (3), (4) для вычисления производных. Если $a_j h^2 + b_j h + c_j$ — зависимость площади (например, насыпи) от рабочей отметки на j -м поперечнике, то предполагая, что в пределах от j -го до $j+1$ -го поперечника площадь изменяется линейно, получаем

$$V_{hj} = (a_j h_j^2 + b_j h_j + c_j + a_{j+1} h_{j+1}^2 + b_{j+1} h_{j+1} + c_{j+1}) L_j / 2.$$

Этот результат соответствует приближенному вычислению интеграла для объема (1) по формуле трапеций.

Если считать, что линейно изменяется не площадь, а коэффициенты аппроксимирующей зависимости, то получаем более точный результат, вычисляя интеграл для объема (1) аналитически:

$$V_{hj} = L_j \{ [h_j^2 (3a_j + a_{j+1}) + 2h_j h_{j+1} (a_j + a_{j+1}) + h_{j+1}^2 (a_j + 3a_{j+1})] / 6 + [(b_j + b_{j+1})(h_j + h_{j+1}) + b_j h_j + b_{j+1} h_{j+1}] / 3 + c_j + c_{j+1} \} / 2. \quad (5)$$

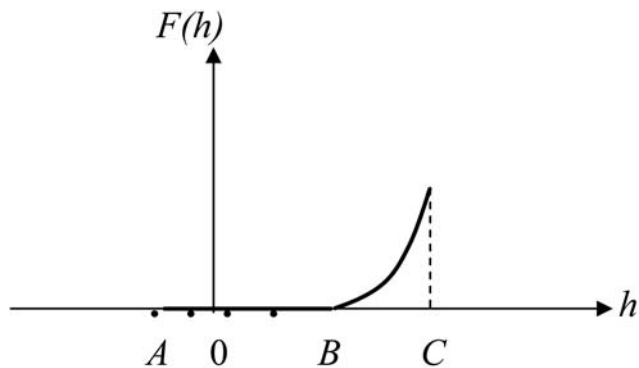


Рис. 2. Аппроксимация: AC — зона поиска, BC — зона аппроксимации

Аналогичная формула получается для объема выемки, что дает возможность вычисления производных по неизвестным рабочим отметкам h_j .

Более того, при наличии нескольких слоев грунта в насыпях и выемках можно вычислить площади и построить соответствующие зависимости (5) для каждого слоя отдельно на каждом поперечнике. Это позволяет перейти от оптимизации по объемам земляных работ к оптимизации по строительным затратам, если грунты выемок не используются для сооружения насыпей. Достаточно площадь каждого слоя умножить на соответствующие расценки по видам земляных работ.

Если же грунты выемок используются для сооружения насыпей (хотя бы некоторые слои), то задача существенно осложняется, так как появляется дополнительная взаимосвязь положения элементов проектной линии в насыпях и в выемках.

По результатам вычисления объемов земляных работ можно выполнить предварительное распределение земляных масс и выбор строительных механизмов. Соответствующие задачи линейного программирования были решены в 60-х годах прошлого века [13].

Для оптимизации по строительным затратам грунты подразделяются на четыре вида в соответствии с возможностями их использования для сооружения насыпей:

- непригодные;
- обыкновенные;
- дренирующие нескальные;
- скальные.

Площади и объемы вычисляются для каждого вида грунтов отдельно.

Дополнительно для грунтов каждого вида задаются в расчете на 1 м^3 :

- q_1 — затраты на сооружение насыпи из выемки;
- q_2 — затраты на сооружение насыпи из грунта карьеров (резервов);
- q_3 — разработки непригодного или излишнего грунта выемок в отвал.

Эти данные могут различаться на различных участках проектируемого профиля. Затраты на сооружение земляного полотна K вычисляются через объемы насыпей v_H и выемок v_B в зависимости от соотношения объемов:

- если $v_H > v_B$, то $K = q_1 v_B + q_2 (v_H - v_B) = q_2 v_H + (q_1 - q_2) v_B$;
- если $v_H \leq v_B$, то $K = q_1 v_H + q_3 (v_B - v_H) = (q_1 - q_3) v_H + q_3 v_B$.

Наличие непригодного грунта учитывается отдельно.

Характерно, что коэффициенты при v_H , v_B (приведенные единичные стоимости) меняются в процессе оптимизации от итерации к итерации, если меняется соотношение объемов. Поэтому на каждой итерации вычисляются соответствующие площади и объемы в целом по участку, на котором насыпи и выемки сооружаются совместно, затем оп-

ределяются приведенные единичные стоимости для грунтов каждого вида и далее градиент целевой функции. Таким образом, учитывается взаимосвязь положения проектной линии в насыпях и выемках, сооружаемых совместно. Именно эта взаимосвязь не позволяет при оптимизации продольного профиля по строительным затратам использовать динамическое программирование, но учитывается в целевой функции при использовании нелинейного программирования [5]. Если задать единичные стоимости в расчете на какое-либо соотношение объемов (например, насыпей больше, чем выемок), то в условиях пересеченного рельефа получается проектная линия с обратным соотношением объемов.

В строительную стоимость включаются и затраты на искусственные сооружения, для чего задаются зависимости их стоимостей от рабочих отметок. Эти зависимости уточняются в процессе проектирования, так же как и типы искусственных сооружений. При этом может потребоваться повторение расчетов по оптимизации профиля с уточненными данными.

Для оптимизации по строительным затратам используется та же компьютерная программа, так как изменяются только правила вычисления градиента.

Если проектирование продольного профиля выполняется для сравнения вариантов трассы, то после вычисления тех видов затрат, которые практически не изменяются при изменениях проектной линии продольного профиля (например, затрат, пропорциональных длине трассы), и сравнения вариантов по суммарным строительным затратам уже на этом этапе часть вариантов трассы может быть исключена из дальнейшего рассмотрения.

Для уточнения положения проектной линии продольного профиля, полученной в виде ломаной, она должна быть преобразована в последовательность элементов требуемого вида с соблюдением всех норм и ограничений. В системе "Профиль-2А" для этой цели использовался эвристический алгоритм последовательной локальной аппроксимации с возвратами. Возможности современных персональных компьютеров позволили реализовать метод динамического программирования как для преобразования в последовательность круговых кривых, сопрягаемых отрезками прямых [14], так и для преобразования в последовательность парабол второй степени [15].

Если грунты выемок не используются для сооружения насыпей, то целевая функция соответствует строительным затратам, и решение, полученное по алгоритму динамического программирования [14, 15], может быть принято как окончательное. Его дальнейшая оптимизация может устранить только влияние накопления ошибок, обусловленных дискретностью поиска. Если же насыпи сооружаются из грунтов выемок, то целевая функция соответствует объемам земляных работ, и оптимизация полученной линии по строительным затратам необходима. При этом отклонения от аппроксимированной ломаной, как правило, не превышают 0,5 м.

Оптимизация проектной линии

Независимо от принятого представления проектной линии задача оптимизации ее положения может быть сведена к оптимизации ломаной, составленной из элементов требуемой длины, т. е. к задаче, решенной при проектировании продольного профиля новых железных дорог.

Оптимизация при наличии круговых кривых. Поскольку смежные круговые кривые имеют общую касательную, то проектная линия полностью определяется координатами вершин углов поворота и радиусами вписанных окружностей (рис. 3).

Если считать неизменными длины проектных элементов S_j и выбрать в качестве неизвестных проектные отметки вершин углов поворота и радиусы вписанных кривых, то через эти "новые" неизвестные можно выразить "старые" неизвестные, т. е. рабочие отметки в точках перелома профиля земли ($B'B''$ на рис. 3). При наличии таких выражений вычисление градиента сводится к простому пересчету производных целевой функции [4, 5].

Действительно, отметки точек (A, D, B) на прямолинейных элементах линейно зависят от отметок вершин ломаной. В силу малости уклонов i_j элементов их длины практически равны разности абсцисс вершин углов поворота, а угол поворота равен разности уклонов. Это позволяет с достаточной точностью выразить отклонения точек кривой от

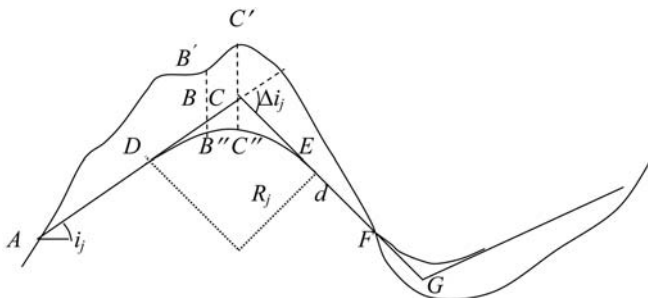


Рис. 3. К пересчету производных при наличии окружностей

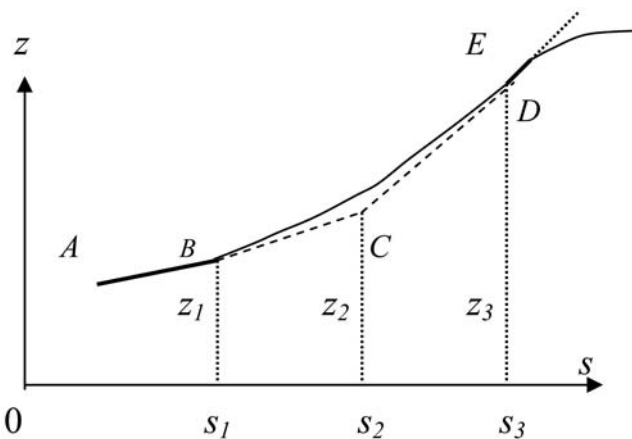


Рис. 4. К пересчету производных при наличии парабол

соответствующих точек прямых (CC'', BB'' на рис. 3), т. е. поправки к рабочим отметкам, вычисленным по прямолинейным элементам ("стрелы"):

$$CC'' = \delta_j = R_j \Delta i_j^2 / 8;$$

$$BB'' = \delta_B = \delta_j - t_B \Delta i_j / 2 + t_B^2 / (2R_j), \quad (1)$$

где $t_B = s_C - s_B$ — разность абсцисс; $\Delta i_j = i_{j+1} - i_j$, а $i_j = (Z_{j+1} - Z_j) / S_j$. Здесь Z_j — неизвестные проектные отметки вершин углов поворота, а S_j — длины элементов (AC, CG на рис. 3).

Поскольку в условиях пересеченного рельефа, как правило, выпуклые кривые проектируются в выемках, а вогнутые в насыпях, то радиусы R_j принимают минимально допустимые значения, так как иначе можно, уменьшив радиусы, уменьшить и рабочие отметки и, следовательно, объемы земляных работ.

На данном этапе возможны только незначительные изменения проектной линии и соответственно малые изменения радиусов R_j , полученных при преобразовании оптимальной ломаной линии с помощью алгоритма динамического программирования в последовательность прямых и круговых кривых. Поэтому радиусы после округления до ближайших целых значений фиксируются или устанавливается диапазон их изменения от $R_{j \min}$ до $R_{j \max} = R_j + 10$ м.

Вместо ограничений на разность уклонов есть ограничения на минимальные длины кривых $R_j \Delta i_j \geq L_{j \min}$ и требование исключить наложение кривых, т. е. суммы $CE + FG$ на рис. 3 не должны превосходить длины элементов: $R_j \Delta i_j / 2 + R_{j+1} \Delta i_{j+1} / 2 \leq S_j$.

В силу малых изменений проектной линии это условие можно упростить и исключить взаимосвязь смежных $R_j \Delta i_j$. Для этого имеющийся минимальный "запас длины" слева и справа d распределяется поровну между смежными кривыми и вычисляются их максимальные длины $L_{j \max}$. (На рис. 3 $L_{j \max} / 2 = CE + d/2$). В итоге $L_{j \min} \leq R_j \Delta i_j \leq L_{j \max}$. Переходя от переменного радиуса к переменной кривизне, получаем двустороннее линейное неравенство. Если на некоторой итерации это ограничение активно, то в дополнение к ранее рассматривавшимся базисным векторам (сдвиг и повороты) [4, 5] есть базисный вектор, соответствующий изменению разности уклонов с компенсацией этого изменения за счет изменения кривизны.

Оптимизация при наличии парабол. Этот случай отличается от рассмотренного тем, что вместо вычисления рабочих отметок на касательных, а затем поправок к ним ("стрел") вычисляются проектные отметки по параболом в точках перелома профиля земли и далее рабочие отметки (рис. 4).

Рассмотрим параболу BD (рис. 4). Точка A — конец предыдущей параболы или начальная точка профиля, а точка E — начало последующей параболы.

В качестве неизвестных снова примем ординаты вершин углов поворота (точка C на рис. 4). Легко показать, что $s_2 = (s_1 + s_3)/2$ и соответственно

$$z_2 = z_1 + (s_3 - s_1)(2a_j s_1 + b_j)/2.$$

Здесь a_j и b_j — параметры параболы $z = a_j s^2 + b_j s + c_j$ и соответственно $2a_j s_1 + b_j$ — уклон в точке B . Следовательно, получив в результате преобразования оптимальной ломаной координаты концов парабол и прямых, мы можем вычислить координаты вершин углов поворота, т. е. начальное приближение для оптимизации элементов.

Далее, $1/(2a_j) = R_j$, где R_j — радиус кривизны в вершине параболы. По аналогии с оптимизацией круговых элементов радиусы (соответственно параметры a_j парабол) можно зафиксировать или установить для каждого из них диапазон изменения и включить в число неизвестных.

Через неизвестные ординаты вершин углов поворота линейно выражаются уклоны и ординаты точек касательных, но для вычисления ординат точек парабол на каждой итерации придется вычислять координаты концов парабол (s_1, z_1, s_3, z_3 на рис. 4) через изменившиеся координаты вершин углов поворота.

Уклон j -го элемента

$$i_j = 2a_j s_1 + b_j; i_{j+1} = 2a_j s_3 + b_j.$$

Следовательно, $\Delta i_j = 2a_j(s_3 - s_1)$. Но $s_3 - s_1 = 2(s_2 - s_1)$, s_2 известен и $s_1 = s_2 - \Delta i_j/(4a_j)$, а $s_3 = s_2 + \Delta i_j/(4a_j)$; $b_j = i_j - 2a_j s_1$; $c_j = z_1 - a_j s_1^2 - b_j s_1$; z_1 выражается через координаты вершин углов поворота и найденное значение s_1 с помощью линейной интерполяции. Найденных величин (параметры параболы и абсциссы ее концов) достаточно для вычисления рабочих отметок и градиента целевой функции по "старым", а затем и по "новым" переменным.

Поскольку система ограничений имеет ту же структуру, то можно использовать тот же алгоритм приведенного градиента, что и при проектировании новых железных дорог [4, 5].

Проектирование продольного профиля в равнинных условиях

Отметим, что при проектировании дорог в равнинных условиях практически нет задачи оптимизации продольного профиля по строительным затратам, так как проектируется насыпь минимальной высоты h_p (руководящая отметка), определяемая из условий предотвращения снеготранспорта. Локальный участок с малыми колебаниями отметок профиля земли может встретиться и при проектировании в условиях пересеченного рельефа. При оптимизации по строительным затратам на таком участке программа дает рабочие отметки, близкие к нулю, что недопустимо по условиям предотвращения снеготранспорта. В пределах такого участка можно было бы задать дополнительное ограничение по минимальной высоте насыпи. Однако

строгое выполнение этого ограничения во всех точках перелома профиля земли приводит к излишним объемам земляных работ из-за колебаний отметок профиля земли. Если при нарушении этого ограничения, например при $h_p = 2$ м, получаем насыпь высотой $h = 1,8$ м и длиной 20 м, но на значительной длине снижаем высоту насыпи на смежных участках, то такое локальное нарушение обычно допускается (точка A на рис. 5).

Длина такого относительно ровного участка может составлять несколько сотен метров и проектировать его отдельно нет смысла. Поэтому на таких участках изменяется целевая функция. Ветвь функции $F(h)$, соответствующая насыпям $F_H(h)$ (см. рис. 1) при $h > h_p$ сохраняется, а при $h < h_p$ модифицируется, как показано на рис. 6. Эта ветвь используется при вычислении градиента при $h > h_p + \Delta$, в противном случае используется аппроксимирующая парабола и дополнительная парабола $F(h) = F_H(h_p) + k(h - h_p)^2$. Здесь $k > 3$ — коэффициент, заданием которого реализуется возможность проектирования на коротких участках насыпи ниже h_p . Отметим, что при вычислении объемов работ используется только ветвь $F_H(h)$.

При проектировании равнинных участков значительной протяженности можно использовать эту же модель и алгоритм динамического программирования, задав соответствующую зону поиска относительно h_p [15]. Полученное решение в таких условиях является окончательным. При этом нет необходимости в построении начального приближения в виде оптимальной ломаной, так как малая ширина зоны поиска не создает вычислительных трудностей, в отличие от проектирования в условиях пересеченного рельефа.

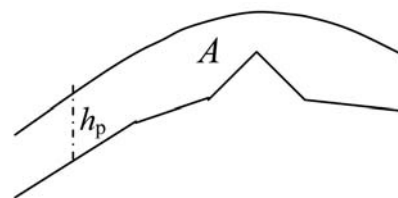


Рис. 5. Проектирование на равнинном участке

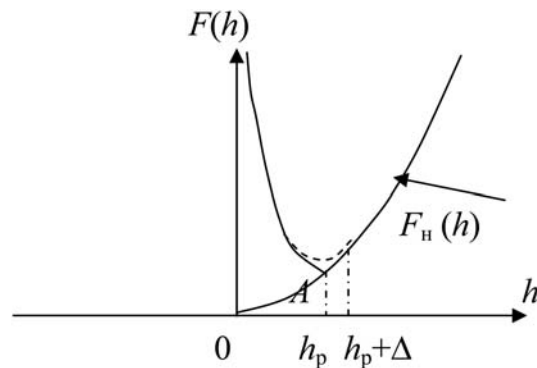


Рис. 6. Моделирующая функция при проектировании на равнинных участках

Заключение

Комплексное применение математических методов оптимизации с использованием предложенных математических моделей позволяет осуществлять компьютерное проектирование продольного профиля автомобильных дорог как для целей сравнения заданных вариантов плана трассы, так и для разработки проекта по выбранному варианту плана трассы. Применение таких методов требует существенного изменения технологии проектного процесса, разработки развитых интеллектуальных систем проектирования нового поколения.

Целесообразность применения разного рода эвристических алгоритмов компьютерного проектирования сомнительна, так как современные вычислительные возможности позволяют использовать математически корректные алгоритмы.

Список литературы

1. Ляховский В. Н., Михалевич В. С., Быков В. И. Определение на ЭВМ наиболее выгоднейшего положения красной линии продольного профиля на вольном ходу // Транспортное строительство. 1962. № 4. С. 41—43.
2. Михалевич В. С., Шор Н. З. Математические основы решения задачи выбора оптимального очертания продольного профиля // Труды Всесоюзного НИИ транспортного строительства. 1964. Вып. 51. С. 12—22.
3. Хавкин К. А. Алгоритм проектирования продольного профиля автомобильных дорог // Автомобильные дороги и дорожное строительство. Киев: КАДИ, 1965. С. 49—57.
4. Struchenkov V. I. Mathematical Models and Optimization in Line Structure Routing: Survey and Advanced Results // International Journal Communication, Network and System Sciences. Special Issue: Models and Algorithms for Application. 2012. № 5. P. 631—638.
5. Струченков В. И. Математические модели и методы оптимизации в системах проектирования трасс новых железных дорог // Информационные технологии. № 7. 2013. С. 10—17.
6. Струченков В. И., Карих Ю. С., Шварц П. С. Математические методы оптимизации в системе автоматизированного проектирования дорог // Автомобильные дороги. 1980. № 12. С. 24—25.
7. CARD/1. URL: <http://www.card-1.com/en/home/>
8. Topomatic Robur. URL: <http://www.topomatic.ru>.
9. Курилко Ю., Чешева В. Geonics ЖЕЛДОП — САПР // CADmaster 2007. № 1 (36).
10. Shafahi Y., Shahbazi M. J. Optimum railway alignment. URL: http://www.uic.org/cdrom/2001/wcrr2001/pdf/sp/2_1_1/210.pdf
11. Jha M. K., Schonfeld P. M., Yong J.-C. and Kim E. Intelligent Road Design. WIT Press, Southampton, 2006.
12. Струченков В. И. Методы оптимизации в прикладных задачах. М.: Солон-Пресс, 2009.
13. Нефёдов П. П., Лопухов А. Е. Современный способ решения задачи распределения земляных масс // Транспортное строительство. 1964. № 4. С. 39—40.
14. Струченков В. И., Козлов А. Н., Егунов А. С. Динамическое программирование в проектировании трасс линейных сооружений // Информационные технологии. 2011. № 8 (180). С. 33—37.
15. Струченков В. И., Козлов А. Н., Егунов А. С. Кусочно-параболическая аппроксимация плоских кривых при наличии ограничений специального вида // Информационные технологии. 2011. № 7 (179). С. 44—49.

УДК 519.688:515

Е. А. Мусихина, канд. техн. наук, проф.,
Т. А. Дарманская, канд. техн. наук, доц.
Иркутский государственный
технический университет,
e-mail: tasha1982@yandex.ru

Исследование рельефа и построение триангуляционной модели

Исследования фрактальных свойств рельефа Иркутской области показали, что природная система неустойчива и не находится в состоянии текущего равновесия. Использован синтез информационных технологий Golden Software Surfer и Autodesk Civil 3D при построении модели участка поверхности в целях выявления характерных особенностей рельефа предполагаемой территории застройки. Обосновано построение TIN поверхности для моделирования сильно расчлененного рельефа.

Ключевые слова: фрактальные свойства рельефа, синергетика, цифровая модель, расчлененность рельефа, массив данных, триангуляция

Общеизвестно, что одной из важнейших задач при проектировании зданий и сооружений является определение характеристик рельефа предполагаемой территории застройки. Особое внимание следует уделить расчлененности рельефа местности. Использование современных информационных технологий позволяет провести моделирование участка застройки в целях выявления его характерных особенностей.

Город, являющийся специфической пространственной системой, которая формируется в процессе развития общества и считается одним из высших проявлений цивилизации, активно воздействует на природную среду, противопоставляя себя ей. Характеризуются городские поселения высокой компактностью, плотностью освоения и коммуникационной насыщенностью. Обычно город развивается в соответствии со стратегией социально-экономического развития, связанной с долгосрочным и краткосрочным планами развития городской территории. Стратегия социально-экономического развития предполагает создание информационной системы принятия решений, базовой составляющей которой является системное моделирование городских поселений. Анализ и моделирование городских систем следует выполнять с учетом сле-

дующих характерных особенностей: город является сложной самоорганизующейся слабоструктурированной нелинейной социальной системой, при управлении которой часто возникает конфликт между целями долгосрочного планирования и краткосрочными решениями; город — целенаправленная и многоцелевая система. Более того, город как объект моделирования характеризуется слабостью теоретических знаний, отсутствием теории развития городов, качественным характером знаний и высоким уровнем неопределенности исходной информации как внутренней, так и внешней. Поэтому в качестве метода моделирования взаимоотношений

Таблица 1

Показатели фрактальности области воздействия

lnR (натуральный логарифм размеров квадратной ячейки)	lnN, % (натуральный логарифм числа непустых ячеек от общего их числа)
1,6	4,34
2,3	4,43
3	4,61

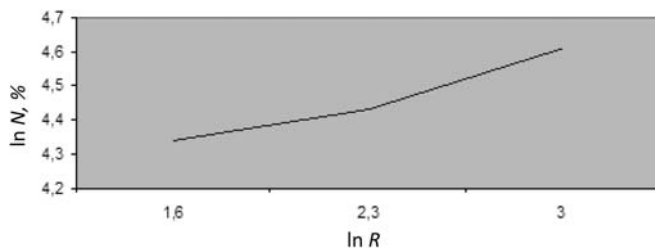


Рис. 1. Фрактальная размерность области воздействия на окружающую среду

Таблица 2

Показатели структурной организации области воздействия на ОС

Показатель	Уровни организации			
	I	II	III	IV
Параметр ячейки структурной матрицы в плане ¹ , км	$\frac{7}{5-9}$	$\frac{25}{17-33}$	$\frac{58}{52-68}$	$\frac{150}{100-150}$
Скейлинговый коэффициент ²	3,1	2,3	2,6	

Примечание: 1) числитель — среднее значение, знаменатель — пределы изменения значений (min-max). Точность измерения соответствует масштабу карты 1:1000000. 2) скейлинговый коэффициент (Ск) — масштабный коэффициент подобия.

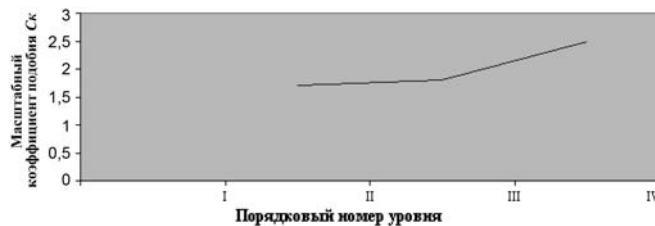


Рис. 2. Кривая степени организации природной системы с учетом пространственно-временных параметров

антропогенной среды с природной предполагается использовать синергетический подход с его принципами. Сочетание фрактального анализа и информационных технологий повысит достоверность предпроектного исследования территории застройки.

Исследование фрактальных свойств рельефа, проведенное традиционным методом (простейшим способом определения размерности), путем набрасывания сетки с последовательным удвоением размеров ее квадратной ячейки, показало, что природная система Иркутской области неустойчива и не находится в состоянии текущего равновесия, поскольку процесс ее формирования продолжается. Что, кстати, подтверждается высокой сейсмической активностью и характеризуется фрактальной размерностью, имеющей значение $D = 1,34$ (тангенс угла наклона осредненного графика равен 0,34) (табл. 1, рис. 1), и скейлинговым коэффициентом (табл. 2, рис. 2), характеризующим масштабное подобие системы, значение которого превышает устойчивую норму (более 2) [1].

В научной работе [1] для оцифровки, векторизации и построения различных карт и планов как двумерного, так и трехмерного вида использовались возможности графического редактора Golden Software Surfer, который позволяет визуализировать цифровую модель рельефа местности (ЦМР) путем обработки снятых координат двух—трех изолиний и нескольких пикетных отметок (рис. 3) [2]. Восприятие файлов формата .bln редактором электронных таблиц MS Excel позволяет оперативно обрабатывать значительные массивы данных, существенно сокращая время, требуемое для построения общей сетки цифровых данных, которая необходима для построения ЦМР.

Создание общего координатного массива, полученного с помощью Golden Software Surfer, позволяет перейти к более широкому спектру использования информационных технологий при исследо-

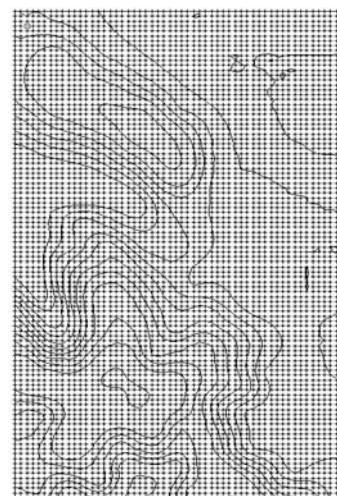


Рис. 3. Оцифровка рельефа местности, выполненная в Golden Software Surfer

вании свойств рельефа предполагаемой территории застройки.

Autodesk Civil 3D является простой в использовании программой для выполнения проектов и построения чертежей, позволяющей решать широкий круг задач гражданского строительства. Создание точек рельефа по полученным с помощью графического редактора Golden Software Surfer координатам показано на рис. 4 (см. вторую сторону обложки).

Создание цифровой модели рельефа.

В AutoCAD CIVIL-поверхности, в зависимости от назначения, делятся на две группы и являются объектами с именами:

- AECC TIN SURFACE — TIN-поверхности;
- AECC GRID SURFACE — сетчатые поверхности [3].

Поверхность — это трехмерное геометрическое представление участка земной поверхности (рис. 5) или разность (композиция) двух областей поверхностей (в случае поверхностей для вычисления объема). Поверхность формируется из треугольников или сеток, которые создаются в AutoCAD Civil 3D в результате соединения точек, составляющих данные поверхности. Линии TIN служат сторонами треугольников, образующих триангуляционную сеть поверхности. Линии TIN создаются в AutoCAD Civil 3D посредством соединения точек поверхности, расстояние между которыми является наименьшим. Отметка каждой точки на поверхности определяется посредством интерполяции значений отметки в вершинах треугольника, в котором рас-

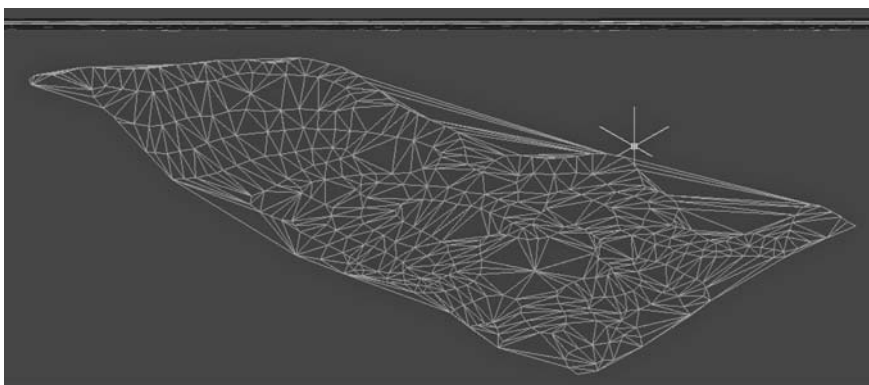


Рис. 5. Поверхность TIN

положена данная точка. Эти поверхности являются наиболее точными и с большим спектром возможностей по изменению и редактированию.

В окне "Свойства поверхности" отображаются все параметры: число точек, минимальная или максимальная отметка, площадь поверхности, интервал уклона и т. д. (рис. 6).

На рис. 7 показана поверхность в стиле Contours 5' and 25' с указанием меток. Этот стиль необходим для детальной привязки здания, что позволяет увеличивать или уменьшать шаг горизонталей в случае использования другого стиля, например Contours 2' and 10'. Метки присваиваются многим объектам и их содержание обновляется каждый раз при изменении объекта, что позволяет контролировать результат и максимально приближаться в заданному. Последовательность действий при редактировании стиля метки: выделите курсором мыши требуемый стиль в окне "Область инструментов" на вкладке

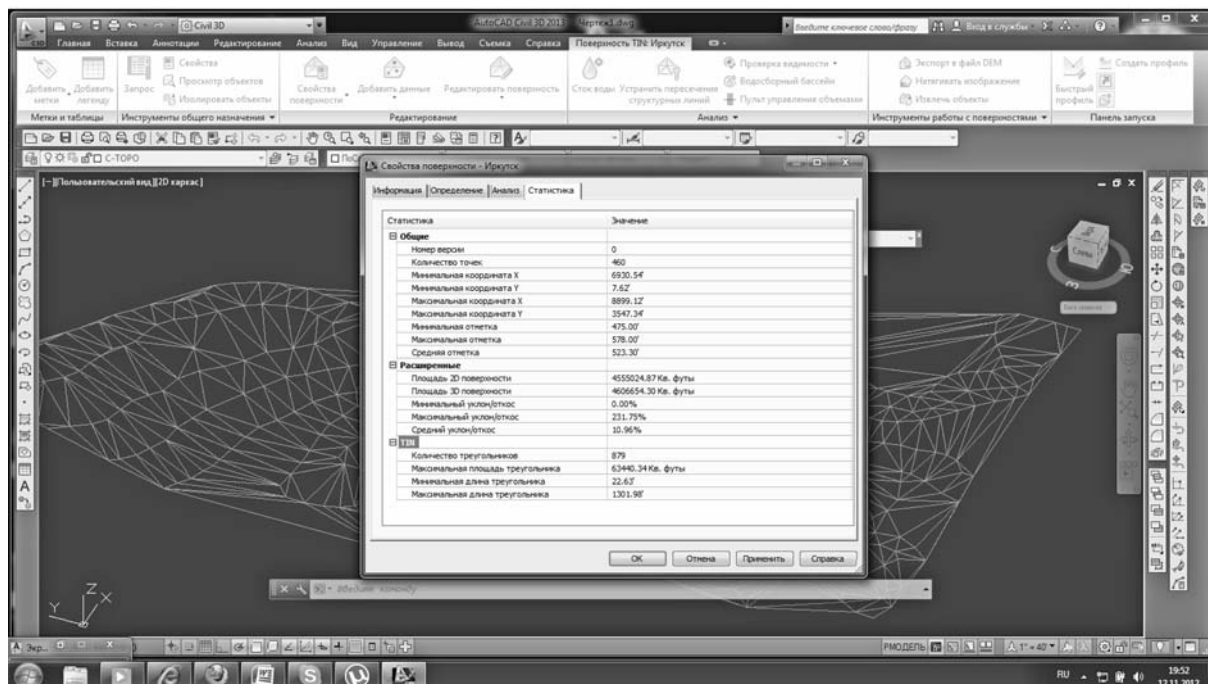


Рис. 6. Характеристика свойств цифровой поверхности

"Параметры", нажмите правую кнопку мыши и выберите "Редактировать". Создание и сохранение наборов меток для трасс, профилей и сечений позволит применять в одной операции метки нескольких типов. Например, набор меток для трасс может содержать метки для основных пикетов, вспомогательных пикетов и геометрических точек.

На триангуляцию поверхности оказывают влияние данные структурных линий (получаемые из структурных линий, горизонталей и границ) [4, 5]. При наличии ребра структурной линии между точками программа предусматривает соединение этих точек ребром триангуляционной сети (TIN) даже в том случае, если это приводит к нарушению свойства Делоне.

Приведенная выше поверхность TIN (см. рис. 5) требует некоторых доработок, поскольку при построении методом триангуляции программа достраивает вершины треугольников по ближайшим точкам, что искажает истинную картину рельефа. Результат обрезки по внешнему контуру представлен на рис. 8.

На рис. 9 представлена 3D-модель поверхности с котлованом для заложения фундамента, наглядно показывающая, что необходимы работы по выявлению значения угла наклона рельефа, направления потока воды в целях определения мест локализации водосбора.

Разность в числовых отметках (30 м) и угол склона (достигает 20 %) указывают на значительный уклон, что необходимо учитывать при выполнении анализа поверхности в целях подготовки площадки под строительство (рис. 10), при задачах определения объема изъятия и/или подсыпки грунта. В изучение рельефа также входит выявление мест локализации водосборов (рис. 11, см. вторую сторону обложки). В программе AutoCAD Civil при изучении водосборов определяются площадь и тип поверхности. Установлено, что на данном участке большие площади занимает поверхность уклона (*boundarysegment*), например, точка 25 имеет тип *boundarysegment* с площадью 782 м², при общей площади поверхности 4606554,30 м².

На рис. 12 (см. вторую сторону обложки) представлена модель направления стока поверхностных вод, что является весьма существенным фактором, если рельеф изучается в целях определения участка под застройку. Котлован строящегося дома представлен прямоугольной выемкой, показывающей необходимость дополнительных мер для защиты от стоков дождевых вод на период строительства и при дальнейшей эксплуатации здания.

Поверхности TIN характеризуются достаточно широким спектром применения. В данном случае

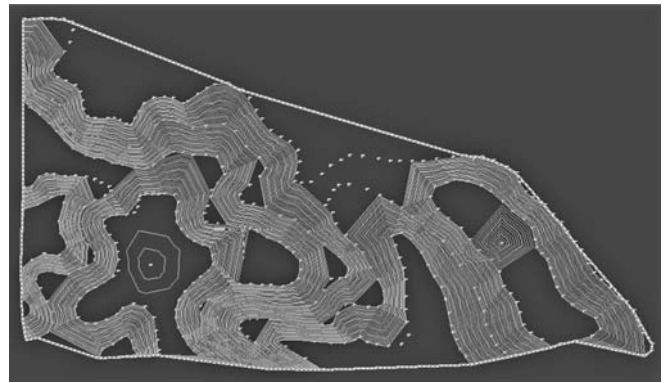


Рис. 7. Цифровая поверхность рельефа

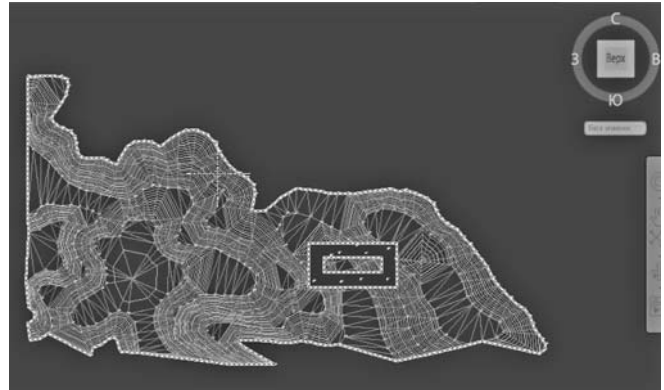


Рис. 8. Полученная в результате обработки модель поверхности

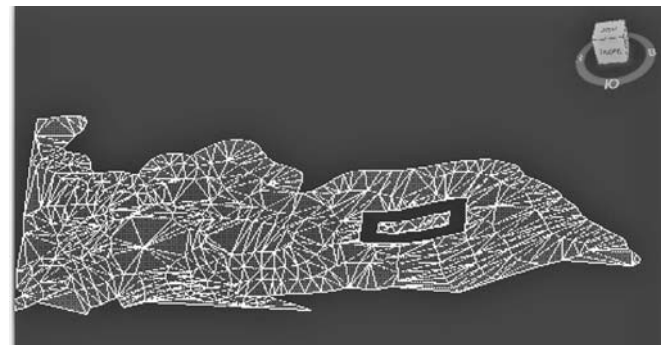


Рис. 9. 3D-модель поверхности с котлованом для заложения фундамента

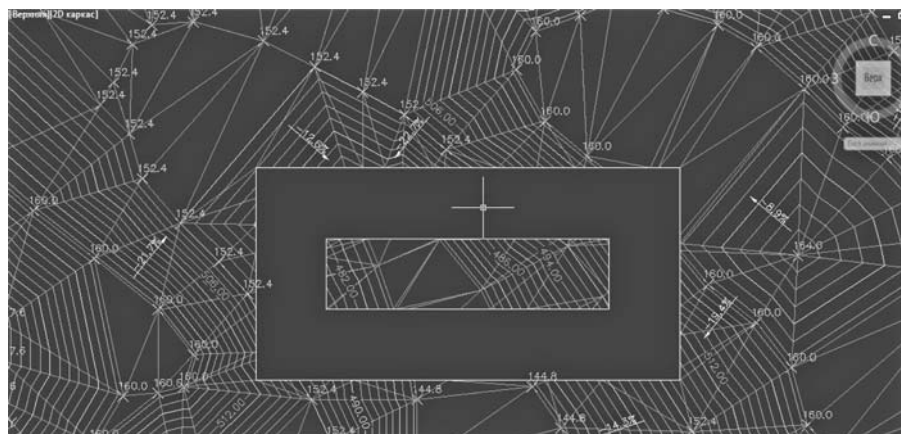


Рис. 10. Моделирование рельефного перепада

моделирование поверхностей TIN проводится для следующих целей:

- моделирование поверхностей с высокой степенью неоднородности рельефа, характеризующихся неравномерным распределением выборочных данных, отражающим влияние потоков, дорог и водоемов;
- исследование локализованных областей (крупномасштабные карты).

Построение поверхностей TIN требует большего временного интервала и соответственно увеличения накопительного ресурса по сравнению с сетчатыми поверхностями и именно поэтому их моделирование не всегда эффективно. Однако использование синтеза информационных технологий Golden Software Surfer и Autodesk Civil 3D в сочетании с фрактальным анализом свойств рельефа при построении моделей участка поверхности со сложным и сильно

расчлененным рельефом, предполагает надежность выбора планировочного решения при проектировании объекта строительства [6].

Список литературы

1. Мусихина Е. А. Методологический аспект технологии комплексной оценки экологической емкости территорий. М.: Академия Естествознания, 2009. 137 с.
2. Мусихина Е. А. Пособие к выполнению лабораторных работ в Golden Software Surfer. Иркутск: Изд-во ИрГТУ, 2007. 40 с.
3. Пелевина И. А. Самоучитель AutoCAD Civil 3D 2010. СПб.: БВХ-Петербург, 2010. 512 с.
4. Кошкарёв А. В., Тикунов В. С. Геоинформатика. М.: Картогеоцентр-Геозидат, 1993. 213 с.
5. Костюк Ю. Л. Представление рельефа земной поверхности в геоинформационных системах // Геоинформатика—2000: Труды Межд. науч.-практ. конф. Томск: Изд-во Том. ун-та, 2000. С. 12—17.
6. Мусихина Е. А., Хохрин Е. В., Сунатори О. М. Развитие урбанизированных территорий: сценарное моделирование. Saarbrücken: LAP LAMBERT Academic Publishing GmbH&Co, 2012. 95 с.

УДК 629.73.533.6

Е. С. Чувашева, ассистент,
С. Н. Чувашев, д-р физ.-мат. наук, проф.
"МАТИ"— РГТУ им. К. Э. Циолковского
e-mail: chuvashevalena@gmail.com, snchuv@mail.ru

Выбор рациональных характеристик высокоскоростных летательных аппаратов разных масштабов на основе комплексной математической модели

С помощью разработанной прикладной программы, реализующей комплексную модель, описывающую разнородные физические процессы в высокоскоростных летательных аппаратах, рассчитываются аппараты различных масштабов и конструкций и разные режимы полета, указываются узкие места и возможности их устранения. Такой анализ технических решений предназначен для автоматизации и эскизного проектирования.

Ключевые слова: высокоскоростные летательные аппараты, теплообмен, динамика полета, комплексные модели

Введение

При разработке эскизных проектов новых технически сложных изделий в последнее время все чаще применяют математическое моделирование. Особенно эффективно оно в случаях принципно

ально новых аппаратов, для которых затруднены или практически невозможны традиционные эвристические методы эскизного проектирования на основе опыта эксплуатации ближайших аналогов.

Это в полной мере относится к новому классу летательных аппаратов, предназначенных для длительного (минуты—десятки минут) движения в верхней атмосфере с высокими скоростями (выше 1 км/с) с помощью прямоточных воздушно-реактивных двигателей. Они отличаются от самолетов повышенными скоростями, а от ракетной техники — экономичностью (для них не требуется запаса окислителя). Среди необычных особенностей таких аппаратов можно отметить остроту вопроса поддержания теплового баланса: на аппарат со стороны окружающей среды постоянно поступают тепловые потоки вследствие аэродинамического нагрева, которые могут сниматься либо тепловым излучением (что предполагает высокие температуры поверхности), либо за счет расходования запасенного бортового хладагента.

1. Комплексная математическая модель

Для выполнения различных задач потребуются различные характеристики высокоскоростных летательных аппаратов, для которых оптимальными будут различные технические решения. Для их предварительного анализа нами создана комплексная математическая модель высокоскоростного летательного аппарата [1]. В ней описывается трехмерная динамика полета [2] с учетом заданных законов управления, в том числе ручного. Возможно задание геометрии ЛА, составленной из характерного

набора поверхностей — плоские, конические, цилиндрические, тороидальные... Для описания сверхзвуковой аэродинамики обтекания ЛА использовали аналитические соотношения для косых ударных волн и течений Прандтля—Майера [3], полуэмпирические соотношения для тепловых потоков и сил трения в ламинарных и турбулентных пограничных слоях [4, 5], для окрестностей критических точек и линий — формулы [6—8], для ламинарно-турбулентного перехода — соотношения [9]. Распределения основных параметров в ядре течения в двигателе рассчитывали с учетом зависимости от продольной координаты по балансовым соотношениям, торможение и нагрев на стенках — по работам [3, 5], горение — по [10—12], предполагалось сгорание углеводородов до CO_2 и H_2O . Рассчитывается нестационарный нагрев внутренних структур аппарата. Учитывается аэродинамический нагрев, радиационное охлаждение, передача теплового потока по "тепловым мостикам" (креплениям, изоляции, трубам и т. п.).

2. Примеры выбора рациональных параметров. Аппарат большого масштаба

В данной работе приведены примеры компьютерного моделирования с помощью этой комплексной модели для выбора конструкций и характеристик высокоскоростных летательных аппаратов на примере выполнения задачи доставки груза фиксированной массы (300 кг) на максимальное расстояние.

В рассмотренных ЛА нижняя часть носовой поверхности выбрана конической, а верхняя — тороидальной, образующиеся при их пересечении острые кромки закруглены с радиусом 1 см. Воздухозаборник состоит из участков трех соосных конусов, расположенных друг за другом. Обечайка двигателя и верхняя часть корпуса над двигателем — участки цилиндрических поверхностей. Также имеются управляющие аэродинамические поверхности, расположенные в районе двигателя. При расчете охлаждения стенок двигателя считалось, что на испарение тратится 2,32 МДж/кг, а на эндотермическую химическую реакцию — 3,60 МДж/кг. Внутренняя структура аппарата (рис. 1) содержит "внешний слой" оборудования и приборов b (в носовой части и над двигателем его толщина 10 см, в области воздухозаборника 20 см, $\rho_b C_{pb} = 2,75 \text{ МДж}/(\text{м}^3 \cdot \text{К})$, ρ — плотность, C_p — теплоемкость), на который падает излучение со стенки w (толщиной 0,5 см, с $\rho_w C_{pw} = 5,5 \text{ МДж}/(\text{м}^3 \cdot \text{К})$), и заэкранированными этим слоем b "внутренними частями" d (толщиной 3 см, $\rho_d C_{pd} = 2,75 \text{ МДж}/(\text{м}^3 \cdot \text{К})$). Для "тепловых мостиков" общее сечение, равно 5 % от площади теплообмена при $\lambda/\Delta = 0,052 \text{ Вт}/(\text{м}^2 \cdot \text{К})$, (λ — теплопроводность, Δ — длина мостика). Для обеспечения оптимальной работы двигателя применяется закон управления, при котором полет происходит

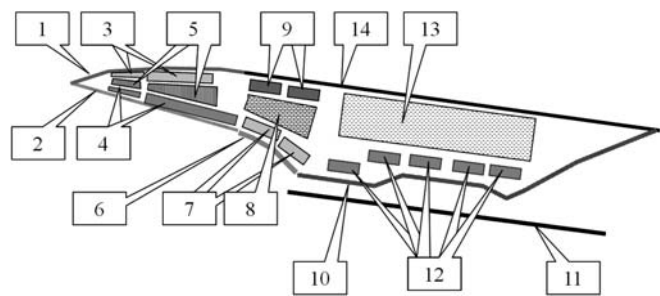


Рис. 1. Структура ЛА (к расчету тепловых процессов при полете):

1, 2 — верхняя и нижняя стенки передней части ЛА; 3, 4 — устройства, примыкающие к верхней и нижней стенкам передней части ЛА; 5 — устройства в глубине передней части ЛА; 6 — стенка воздухозаборника; 7 — устройства, примыкающие к стенке воздухозаборника; 8 — устройства в глубине над воздухозаборником; 9 — устройства, примыкающие к верхней стенке основной части фюзеляжа ЛА; 10, 11 — верхняя и нижняя стенки двигателя; 12 — устройства, примыкающие к верхней стенке двигателя; 13 — бак с основным горючим; 14 — верхняя стенка основной части фюзеляжа ЛА

с постоянным углом атаки 6° и постоянным коэффициентом избытка воздуха, равным 1,5. Некоторые расчеты с исследованием режимов полета проведены для массы конструкции ЛА 506 кг, начальной массы топлива 1288 кг, длины ЛА 6,1 м, диаметра 0,9 м. В работе [1] было показано, что при полете такого ЛА перегревается бак с горючим и устройства под ним, поэтому они заизолированы от излучения со стенок двигателя двумя слоями фольги с коэффициентом отражения каждого слоя 97 %.

В работе [1] полет такого же ЛА при вышеописанных законах управления происходил по периодической траектории с заметными колебаниями высоты полета и других связанных с ней характеристик. Проведенные в данной работе расчеты показали, что амплитуда этих колебаний сильно зависит от скорости и высоты полета, при которых включается прямоточный двигатель. На рис. 2, 3 представлены параметры режима, при котором колебания практически отсутствуют.

Сравнение результатов с данными по режиму [1] показало, что давление в двигателе при таком режиме меняется в меньшем диапазоне и не превышает 16 атм. (в работе [1] — 20 атм.), что несколько снижает требования к жаропрочности стенок. Дальность полета всего на 7 % ниже. Большие положительные отличия в температурных режимах устройств над двигателем и бака с горючим связаны с мерами по изоляции от радиационных потоков со стороны двигателя. Уровни температуры остальных устройств, скорость и высота полета остаются на том же уровне. Таким образом, существенный выигрыш от применения периодических "рикошетирующих" траекторий [13] получается не всегда.

Описанные режимы не оптимальны с точки зрения стартовой массы комплекса: ЛА + первая (ракетная) ступень, которая разгоняет ЛА и поднимает в разреженные слои атмосферы. Эта ступень

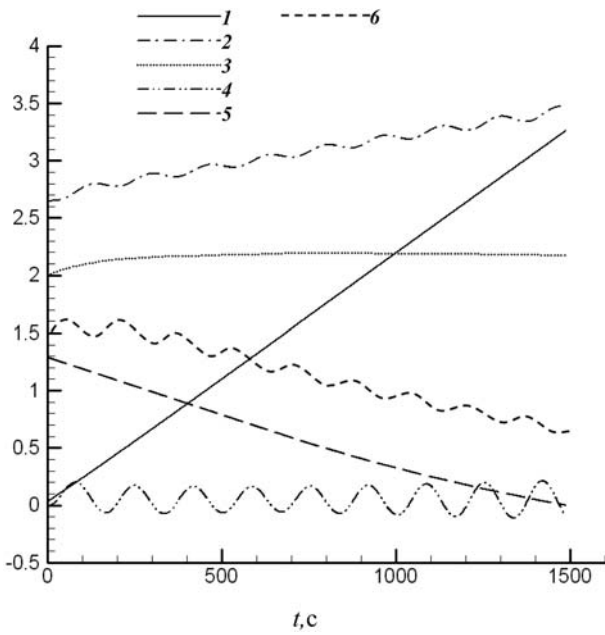


Рис. 2. Временные зависимости параметров полета: 1 — удаление L от точки старта прямооточного воздушно-реактивного двигателя, 10^6 м; 2 — высота полета Z , 10^4 м; 3, 4 — горизонтальная и вертикальная составляющие скорости, 10^3 м/с; 5 — масса топлива, 10^3 кг; 6 — максимальное давление в двигателе, 10^6 Па

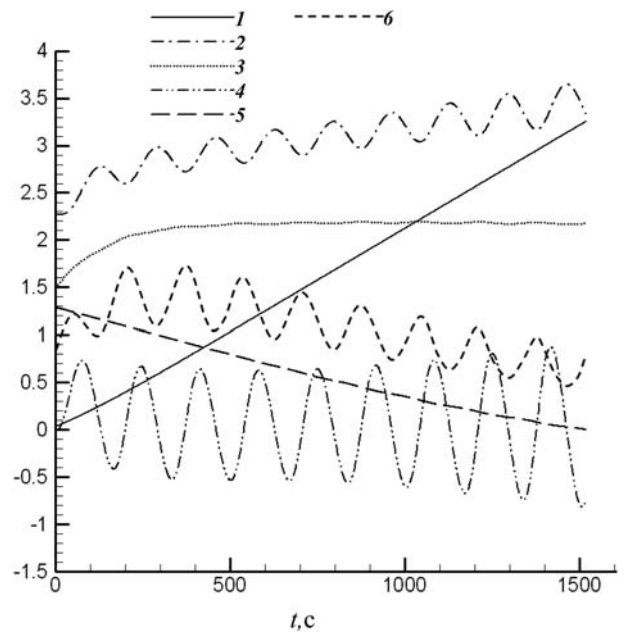


Рис. 4. Временные зависимости параметров полета (запуск с уменьшенными начальными значениями высоты и скорости). Обозначения см. рис. 2

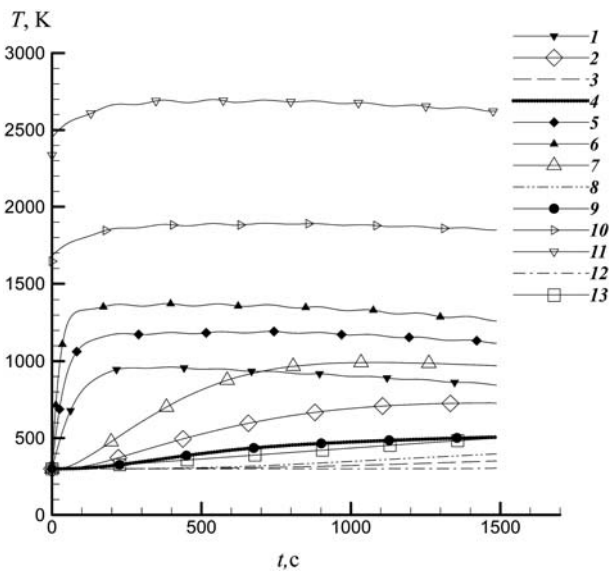


Рис. 3. Временные зависимости изменения температуры устройств ЛА: 1 — средняя температура нижней стенки передней части ЛА; 2 — средняя температура устройств, примыкающих к нижней стенке передней части ЛА; 3 — средняя температура устройств, находящихся в центре корпуса в передней части; 4 — средняя температура устройств, примыкающих к верхней стенке передней части ЛА; 5 — средняя температура воздухозаборника; 6 — максимальная температура воздухозаборника; 7 — средняя температура устройств, примыкающих к воздухозаборнику; 8 — средняя температура устройств, находящихся в центре корпуса в области воздухозаборника; 9 — средняя температура устройств, примыкающих к верхней стенке в области воздухозаборника; 10 — средняя температура в двигателе; 11 — максимальная температура в двигателе; 12 — средняя температура бака; 13 — средняя температура приборов под баком

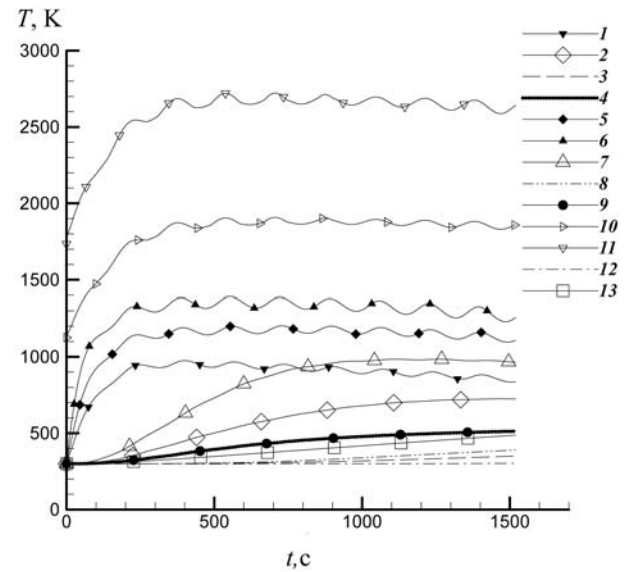


Рис. 5. Временные зависимости изменения температуры устройств ЛА при полете (запуск с уменьшенными начальными значениями высоты и скорости). Обозначения см. рис. 3

в несколько раз больше ЛА, и чем меньше высота и скорость, до которых она работает, тем меньше могут быть ее размеры. С этой точки зрения более эффективен, например, режим, при котором включение прямооточного двигателя происходит на меньшей высоте (22,7 км) с меньшей скоростью (1,5 км/с, рис. 4, 5). ЛА после этого набирает скорость и высоту за счет работы прямооточного двигателя. Экономия стартовой массы связана с тем, что при таком

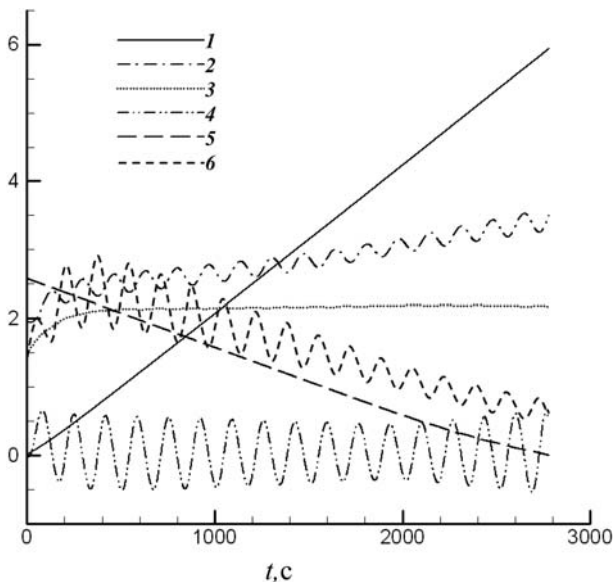


Рис. 6. Временные зависимости параметров полета (масса топлива увеличена вдвое). Обозначения см. рис. 2

разгоне и подъеме не требуется расходовать окислитель. При этом дальность полета, температурные режимы устройств, требования по жаростойкости — примерно такие же, как при предыдущем режиме. Следует отметить, что расчеты показали, что дальше понижать начальную скорость нельзя, иначе поток на входе в прямоточный воздушно-реактивный двигатель становится дозвуковым, и эффективный режим работы последнего не достигается, тяга отсутствует, и аппарат падает. Вместе с тем, необходимо обеспечить достаточную начальную высоту, чтобы не превысить допустимые значения давления

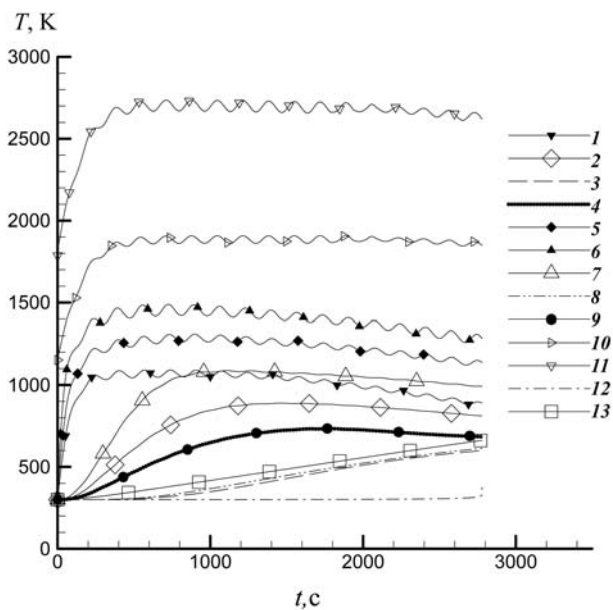


Рис. 7. Временные зависимости изменения температуры устройств ЛА при полете (масса топлива увеличена вдвое). Обозначения см. рис. 3

в двигателе, зависящего от плотности потока поступающего воздуха.

При увеличении массы топлива пропорционально ей растет дальность полета (рис. 6, 7). При этом режиме высота со временем несколько увеличивается, что приводит к снижению плотности гиперзвукового потока, снижению аэродинамических тепловых потоков и температур стенок и прилегающих к ним устройств, которые успевают прийти с ними в тепловое равновесие. Внутренние части ЛА успевают прогреваться до несколько более высоких температур, их изоляция от тепловых потоков может оказаться недостаточной. Пиковый уровень давления приближается к 30 атм., что связано с динамикой ЛА с начальной массой, возросшей за счет дополнительного горючего.

3. Примеры выбора рациональных параметров. Аппараты малого масштаба

Разработанный программный комплекс применим для различных геометрий и масштабов ЛА. Для примера на рис. 8—9 представлены результаты расчетов для ЛА, линейные размеры которого уменьшены в 2,5 раза относительно первого варианта. При этом масса топлива при запуске — 82 кг, коэффициент избытка воздуха равен 2, масса полезного груза и остальные параметры такие же, как и у базового варианта конструкции. При таком уменьшении размеров аппарата существенно (до 700 км) уменьшается дальность полета, однако проблема охлаждения оказывается не так актуальна: температура устройств ЛА ниже за счет инерционности тепловых процессов, а также меньшей скорости

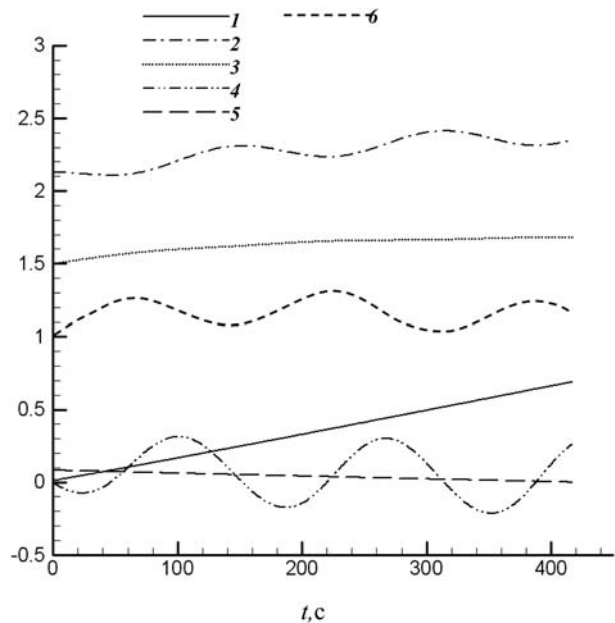


Рис. 8. Временные зависимости параметров полета (размеры ЛА уменьшены в 2,5 раза). Обозначения см. рис. 2

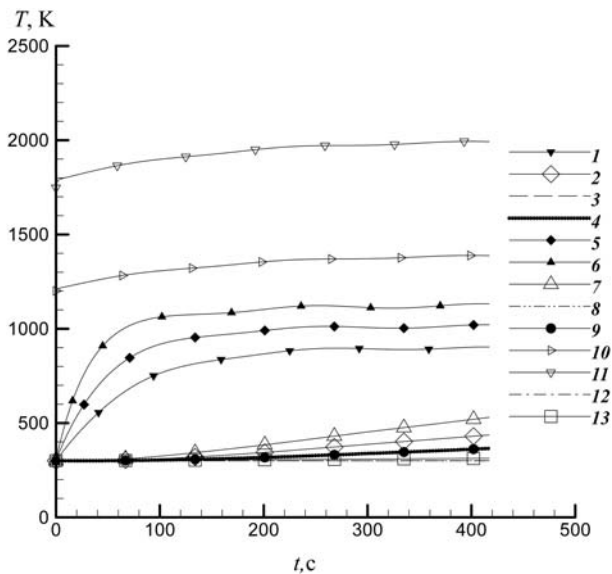


Рис. 9. Временные зависимости изменения температуры устройств ЛА при полете (размеры ЛА уменьшены в 2,5 раза). Обозначения см. рис. 3

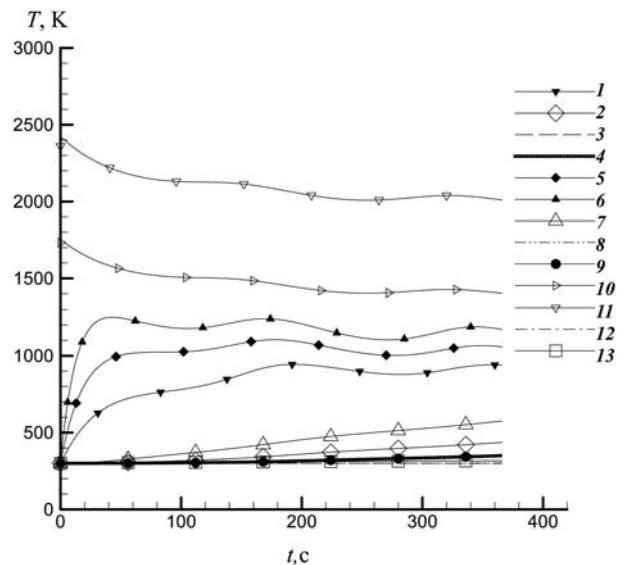


Рис. 11. Временные зависимости температуры устройств ЛА при полете (ракетная ступень больше необходимой). Обозначения см. рис. 3

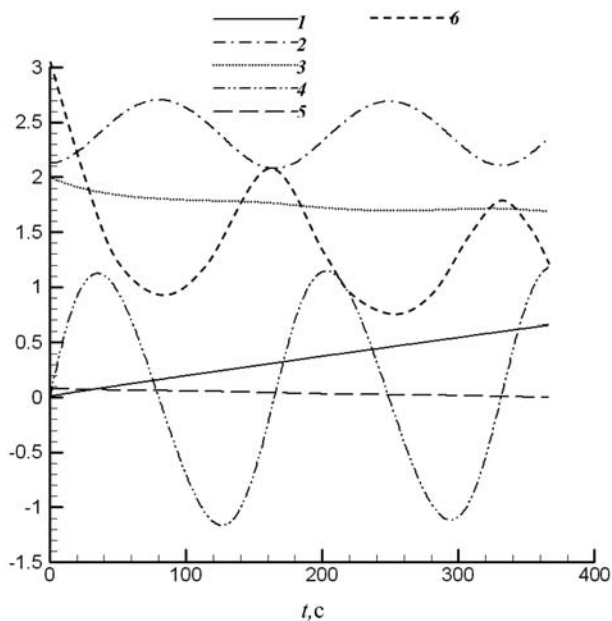


Рис. 10. Временные зависимости параметров полета (ракетная ступень больше необходимой). Обозначения см. рис. 2

набегающего потока. Давление в двигателе не превышает 13 атм. В ходе полета ЛА набирает скорость до $v_k \approx 1,5$ км/с и поднимается на несколько километров.

Если ракетная ступень оказывается больше необходимой (рис. 10, 11) и выводит тот же аппарат на ту же высоту с большей начальной скоростью, то в процессе полета скорость все равно спадает до v_k , т. е. выигрыша по дальности и скорости не получается, однако это приводит к значительному повышению пикового давления в двигателе (до 30 атм.).

Итак, с помощью описанной методики можно указать на узкие места, определить возможные тех-

нические решения и проверить их эффективность. Таким образом, разработанная комплексная математическая модель может помочь в эскизном проектировании вновь создаваемых высокоскоростных летательных аппаратов.

Список литературы

1. Чувашева Е. С., Чувашев С. Н., Зорина И. Г. Комплексная математическая модель для концептуального проектирования высокоскоростных летательных аппаратов // Информационные технологии". 2012. №11. С. 10–44.
2. Бюшгенс Г. С., Студнев Р. В. Динамика самолета. Пространственное движение. М.: Машиностроение, 1983. 320 с.
3. Христианович С. А., Гальперин В. Г., Миллионщиков М. Д., Симонов Л. А. Прикладная газовая динамика. М.: ЦАГИ, 1948. 150 с.
4. Аржаников Н. С., Садекова Г. С. Аэродинамика больших скоростей. М.: Высшая школа. 1965. 562 с.
5. Кутателадзе С. С. Теплопередача и гидродинамическое сопротивление. М.: Энергоатомиздат, 1990. 367 с.
6. Черный Г. Г. Влияние малого затупления на гиперзвуковое обтекание тел // Газовая динамика. Избранное / Под ред. А. Н. Крайко. М.: Физматгиз, 2005. С. 292–308.
7. Beckwith I. E., Gallagher J. J. Local Heat Transfer and Recovery Temperatures on a Yawed Cylinder at a Mach Number of 4.15 and High Reynolds Numbers. NASA TR R-104, 1961.
8. Fay J. A., Riddell F. R. Theory of Stagnation Point Heat Transfer in Dissociated Air // Journal of the Aeronautical Sciences. 1958. N 2 (25). P. 73–85.
9. Quinn R. D., Gong L. A Method for Calculating Transient Surface Temperatures and Surface Heating Rates for High-Speed Aircraft. Edwards, California: Dryden Flight Research Center, (NASA/TP-2000-209034). 2000.
10. Алесасов В. Е., Дрегалин А. Ф., Тишин А. П. Теория ракетных двигателей. М.: Машиностроение, 1980. 535 с.
11. Зельдович Я. Б., Баренблат Г. Н., Либрович Е. И., Махвиладзе Г. М. Математическая теория горения и взрыва. М.: Наука, 1980. 478 с.
12. Физические величины. Справочник / Под ред. И. С. Григорьева, Е. З. Мейлихова. М.: Энергоатомиздат, 1991. 1232 с.
13. Фомин В. М., Аульченко С. М., Звезгинцев В. И. Полет гиперзвукового летательного аппарата с прямоточным воздушно-реактивным двигателем по рикошетирующей траектории // Прикладная механика и техническая физика. 2010. Т. 4 (51). С. 85–94.

УДК 621.396.721, 621.396.73

О. Н. Шерстюков,

д-р физ.-мат. наук, проф., зав. каф.,

Е. Ю. Рябченко, канд. физ.-мат. наук, доц.,

С. Л. Мартыничук, аспирант

Казанский (Приволжский)

федеральный университет

e-mail: marser1986@rambler.ru

Радиомодем LPD-диапазона для задач телеметрии

Предложен метод построения радиомодема LPD-диапазона (433 МГц), позволяющего реализовать систему беспроводной телеметрии в условиях значительного затухания сигнала для частот диапазона 2,4 ГГц, когда затруднено применение беспроводных технологий Wi-Fi, Bluetooth и ZigBee (802.15.4). Предлагаемый радиомодем основан на микросхеме приемо-передатчика SX1233. Для радиомодема разработан протокол канального уровня, описана математическая модель протокола.

Ключевые слова: радиомодем, радиосвязь, LPD, беспроводная телеметрия, протокол канального уровня

Введение

В настоящее время основным средством сбора данных в сложных ландшафтных условиях является кабельная сеть. Сегодня с революционным развитием элементной базы микроэлектроники и локальных цифровых радиосетей появляется возможность создания комплексов сбора данных на основе беспроводных технологий. При выборе технологии построения радиоканала важно придерживаться нелицензируемых диапазонов радиочастот, называемых также ISM (от англ. *Industrial, Scientific and Medical*), что позволит исключить необходимость получения лицензии от ГКРЧ и иных ведомств, а в ряде случаев и регистрации радиочастотных компонентов комплекса.

На сегодняшний день существует ряд стандартов для построения беспроводных систем телеметрии:

- IEEE 802.11 (Wi-Fi), диапазоны 2,412...2,472 ГГц и 5,150...5,825 ГГц;
- IEEE 802.15.1 (Bluetooth), диапазон 2,4015...2,4805 ГГц;
- IEEE 802.15.4 с технологией ZigBee, диапазоны 868,0...868,6 МГц, 902,0...928,0 МГц и 2,412...2,472 ГГц [1, 2, 3].

Радиочастоты, используемые данными технологиями, характеризуются высоким затуханием сигнала. Эффективный коэффициент затухания для радиоволн с частотой 2,4 ГГц при распространении через лесной массив приблизительно равен 0,48 дБ/м, тогда как для диапазона Low Power Device (433 МГц) указанный коэффициент лежит в пределах 0,08...0,17 дБ/м [4]. Для развертывания систем телеметрии в сложных ландшафтных условиях рационально использовать радиоволны более низких частот.

Согласно российскому законодательству диапазон Low Power Device (LPD) разрешено использовать без получения лицензии при выходной мощности передатчика не более 10 мВт. Предусмотрено 69 каналов на частотах 433,075...433,775 МГц с шагом 25 кГц, класс излучения 16K0F3E [5]. Однако в этом диапазоне отсутствуют принятые стандарты цифровых сетей передачи данных. В данной работе приводится способ построения радиомодема, на основе которого будет возможна реализация сети сбора данных, предназначенной для работы в сложных ландшафтных условиях.

1. Построение радиомодема на основе трансивера SX1233

Основой радиомодема (рис. 1) является микросхема трансивера SX1233 компании *Semtech* [6]. Формируемый трансивером сигнал поступает на вход LC-фильтра, предназначенного для подавления внеполосных помех. Через эту же цепь к трансиверу поступает входной сигнал с антенны. Работой трансивера управляет микроконтроллер C8051F350 (производитель *SiLabs*). Для управления радиомодемом используется универсальный асинхронный интерфейс микроконтроллера (UART), который



Рис. 1. Функциональная схема радиомодема

с помощью кабеля FTDI-3V3 может быть соединен с USB-портом компьютера. Кабель FTDI-3V3 содержит преобразователь интерфейсов USB и RS-232, что позволяет подключить интерфейс UART к порту USB. В компьютере после установки соответствующих драйверов данный преобразователь будет определяться как виртуальный COM-порт, более простой в программировании по сравнению с USB-портом [7].

Микросхема трансивера SX1233 позволяет реализовать передачу и прием радиосигнала с использованием различных типов модуляции и несущих частот. Трансивер имеет следующие характеристики:

- диапазон частот 290...1020 МГц;
- максимальная скорость передачи данных — 600 кбит/с при использовании FSK-модуляции;
- выходная мощность передатчика — 18...17 дБм;
- размер буфера — 66 байт;
- ток, потребляемый в режиме передачи, при выходной мощности 13 дБм составляет 45 мА;
- ток, потребляемый в режиме приема, 17 мА.

Функциональная схема трансивера (рис. 2) содержит блок управления, блок приема с входным усилителем, блок передачи с выходным усилителем мощности и синтезатор частоты.

Управление трансивером осуществляется через интерфейс SPI. Микроконтроллер отправляет трансиверу команды определенного формата, позволяющие получить доступ к адресному пространству управляющих регистров. Для доступа к управляющим регистрам трансивера используется 7-битная адресация, при этом 8-й бит адресного байта определяет режим чтения/записи.

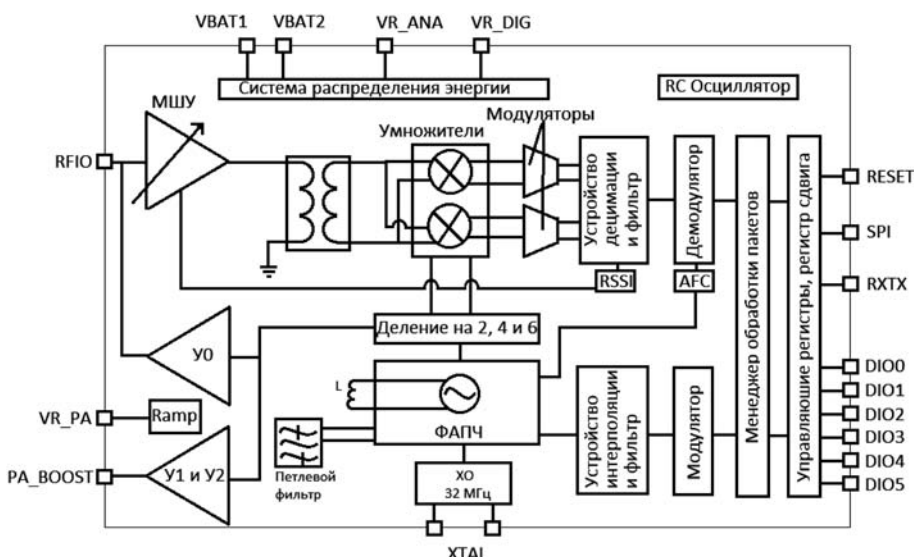


Рис. 2. Функциональная схема трансивера SX1233

Преамбула (1-65535 байт)	Номер сети (1-8 байт)	Адрес (1 байт)	Полезные данные (1-66 байт)	CRC (2 байта)
-----------------------------	--------------------------	-------------------	--------------------------------	------------------

Рис. 3. Структура информационного пакета

SX1233 позволяет реализовать два режима передачи данных: прозрачный и пакетный. Прозрачный режим предоставляет разработчику возможность непосредственно контролировать процесс модуляции и выполнять обработку демодулированного сигнала. При использовании пакетного режима основные операции по обработке сигналов, такие как кодирование, декодирование, формирование передаваемых информационных пакетов, проверка принятых пакетов на наличие искажений трансивер выполняет автоматически. В силу очевидных преимуществ при разработке радиомодема мы использовали пакетный режим передачи данных.

Структура информационного пакета, передаваемого по радиоканалу при работе в пакетном режиме, изображена на рис. 3.

Размеры полей преамбулы, номера сети, полезных данных настраиваются через соответствующие управляющие регистры. Данные записываются в буфер FIFO трансивера, размер которого составляет 66 байт. Для обнаружения ошибок, возникающих при передаче данных, в поле CRC пакета записывается результат вычисления контрольной суммы. Приемник вычисляет контрольную сумму принимаемых данных и сравнивает ее со значением CRC для определения целостности пакета [6].

3. Протокол канального уровня

Использованный авторами протокол канального уровня реализует принцип работы протокола с остановкой и ожиданием. При передаче данных используются два радиомодема — центральный и локальный, в одинаковом исполнении, но различающиеся по функциональному назначению. Определенный объем данных передается от локального модема к центральному.

Протокол канального уровня функционирует по приведенному алгоритму. Информация, которую необходимо передать, делится на пронумерованные блоки одинакового размера. Из полученных блоков информации трансивер локального модема формирует пакеты данных для передачи по радиоканалу, добавляя к ним служебные поля (номер блока данных, адрес получателя, результат вычисления контрольной суммы). Центральный модем отправляет в радиоэфир кадр запроса, содержащий адрес локального модема и номер запрашиваемого блока данных. В ответ на запрос локальный модем отправляет в эфир соот-

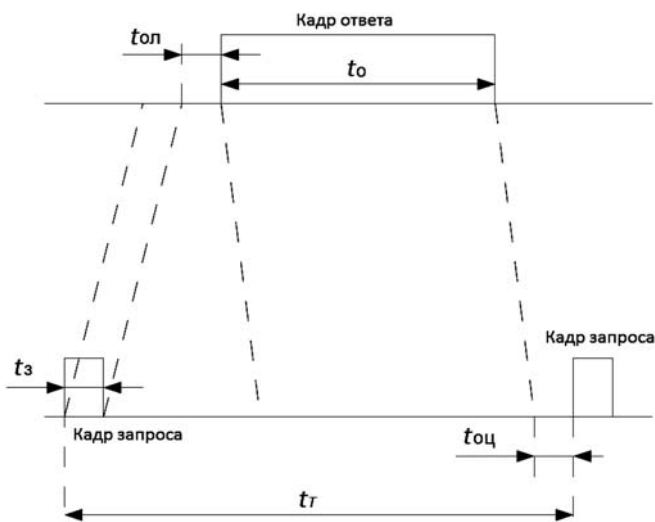


Рис. 4. Временная диаграмма протокола канального уровня

ветствующий пакет данных. При обнаружении искажений информации или отсутствии ответа от локального модема центральный модем выполняет повторный запрос. Передача данных проиллюстрирована рис. 4, где t_3 — длительность кадра запроса; t_o — длительность кадра ответа; $t_{о.ц}$ — время обработки кадра центральным модемом; $t_{о.л}$ — время обработки кадра локальным модемом; t_T — время одного цикла опроса.

Пусть p_3 и p_o — вероятности искажения кадра запроса и кадра ответа соответственно. Тогда вероятность неудачной попытки опроса локального модема $p = p_3 + p_o$. Под опросом здесь понимается получение от локального модема требуемого пакета данных.

Рассчитаем вероятность того, что число повторных запросов, необходимое для успешного опроса локального модема, равно i . Для того чтобы произошел i -й повторный запрос, должны закончиться неудачей первый запрос и $i - 1$ повторных запросов, т. е. всего должно произойти i неудачных попыток опроса локального модема, вероятность этого события равна p^i . Вероятность того, что i -й повторный запрос окажется успешным равна $1 - p$. Таким образом, вероятность $P(i)$ того, что для получения от локального модема требуемого пакета данных центральному модему потребуются выполнить i повторных запросов [8], определяется выражением

$$P(i) = p^i(1 - p). \quad (1)$$

Определим среднее время t_V получения центральным модемом пакета данных от локального модема:

$$t_V = t_T + \sum_{i=1}^{\infty} P(i)it_T. \quad (2)$$

Подставляя выражение (1) в (2), получаем:

$$\begin{aligned} t_V &= t_T + \sum_{i=1}^{\infty} p^i(1 - p)it_T = t_T + t_T(1 - p) \sum_{i=1}^{\infty} ip^i = \\ &= t_T + t_T \left(\sum_{i=1}^{\infty} ip^i - \sum_{i=1}^{\infty} ip^{i+1} \right) = \\ &= t_T + t_T \left(\sum_{i=0}^{\infty} (i + 1)p^{i+1} - \sum_{i=1}^{\infty} ip^{i+1} \right) = \\ &= t_T + t_T \left(p + \sum_{i=1}^{\infty} (i + 1)p^{i+1} - \sum_{i=1}^{\infty} ip^{i+1} \right) = \\ &= t_T + t_T \left(p + \sum_{i=1}^{\infty} p^{i+1}(i + 1 - i) \right) = \\ &= t_T + t_T \left(p + \sum_{i=1}^{\infty} p^{i+1} \right) = t_T + t_T \left(p + \sum_{i=2}^{\infty} p^i \right) = \\ &= t_T + t_T \sum_{i=1}^{\infty} p^i = t_T \left(1 + \sum_{i=1}^{\infty} p^i \right) = t_T \sum_{i=0}^{\infty} p^i. \quad (3) \end{aligned}$$

Подставляя в (3) частный случай формулы биномиального разложения

$$\frac{1}{1 - a} = \sum_{i=0}^{\infty} a^i,$$

получаем

$$t_V = t_T \frac{1}{1 - p}. \quad (4)$$

Согласно формуле (4) среднее время доставки пакета данных от локального модема к центральному модему равно произведению длительности одного цикла опроса на множитель $\frac{1}{1 - p}$, физический смысл данного множителя — математическое ожидание числа запросов, необходимого для доставки требуемого пакета данных от локального модема к центральному модему: $M(k) = \frac{1}{1 - p}$.

Обозначим длину пакета, содержащего поле полезных данных, символом l , тогда скорость передачи данных определяется выражением

$$R = \frac{l}{t_V}.$$

С учетом формулы (4)

$$R = \frac{l}{t_T} (1 - p),$$

откуда получаем выражение для вероятности неудачного опроса:

$$p = 1 - \frac{Rt_T}{l}. \quad (5)$$

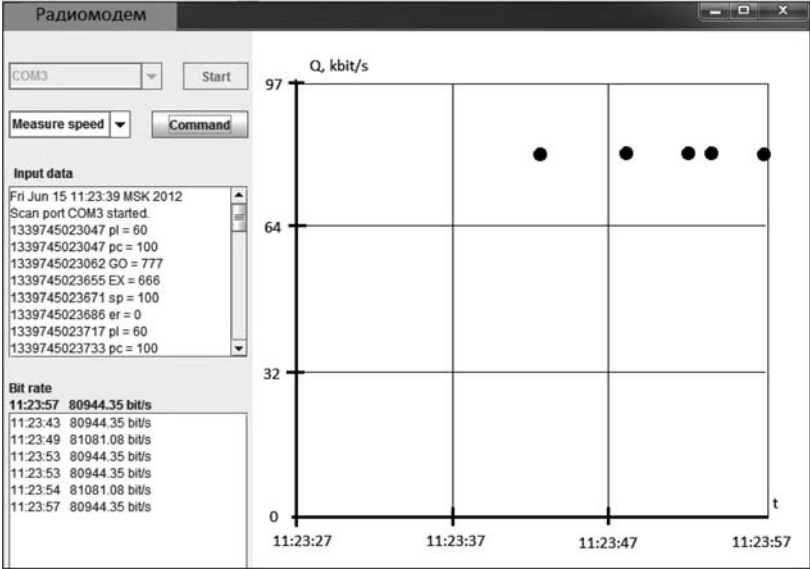


Рис. 5. Программа управления радиомодемом

Определим, как связаны между собой вероятность p_B искажения одного бита и вероятность p неудачного опроса. Опрос локального модема заканчивается неудачей в том случае, если хотя бы один бит кадра запроса или кадра ответа искажен. При вероятности искажения одного бита p_B вероятность p_0 успешного опроса определяется выражением:

$$p_0 = (1 - p_B)^{l_3 + l_0},$$

где l_3 — длина кадра запроса; l_0 — длина кадра ответа (длины кадров выражены в бит). Тогда можно записать:

$$p = 1 - p_0 = 1 - (1 - p_B)^{l_3 + l_0}.$$

Используем формулу биномиального разложения, получим

$$(1 + a)^K = 1 + \sum_{n=1}^{\infty} \frac{K(K-1)(K-2)...(K-n+1)}{n!} a^n.$$

Учитывая, что $p_B \ll 1$, оставляем только первые два члена разложения и получаем

$$p = p_B(l_3 + l_0),$$

откуда следует выражение для вероятности искажения одного бита:

$$p_B = \frac{p}{(l_3 + l_0)}. \tag{6}$$

Полученные формулы (5) и (6) позволяют экспериментально исследовать вероятность неудачного опроса и вероятность искажения бита при передаче данных в различных условиях.

Команды управления радиомодемом

Код команды	Название команды
1	Загрузка текущего времени
2	Чтение текущего времени
3	Загрузка длины пакета
4	Чтение длины пакета
5	Загрузка данных
6	Чтение данных
7	Измерение скорости связи

Преамбула (1 байт)	Код команды (1 байт)	Данные (8 байт)
-----------------------	-------------------------	--------------------

Рис. 6. Формат команд управления радиомодемом

Преамбула (1 байт)	Имя параметра (2 байта)	Значение параметра (8 байт)
-----------------------	----------------------------	--------------------------------

Рис. 7. Формат кадров отчета

4. Протокол обмена данными между компьютером и радиомодемом

Программа управления радиомодемом (рис. 5), написанная на языке Java, осуществляет обмен данными с центральным модемом через интерфейс RS-232. Программа позволяет отправлять радиомодему управляющие команды, а также обрабатывать и визуализировать данные, полученные от радиомодема.

Для управления радиомодемом используется формат команд, изображенный на рис. 6.

Первые два байта содержат преамбулу и код команды, остальные восемь байт — данные. Размер поля данных выбран равным восьми байтам в силу необходимости передавать значение времени в миллисекундах. Список команд приведен в табл. 1.

Первые шесть команд выполняются в соответствии с названием, команда измерения скорости связи запускает сеанс передачи файла размером 6 Кбайт от локального модема к центральному. Отчет о проведенном сеансе связи отправляется в компьютер, формат кадров отчета изображен на рис. 7.

Для каждого сеанса связи предусмотрены следующие параметры:

- ER* — число обнаруженных ошибок;
- PL* — длина пакета;
- PC* — число запрошенных пакетов;
- SP* — число полученных пакетов.

На основе отчета о сеансе связи вычисляется скорость передачи данных.

5. Испытания радиомодема

Были проведены измерения скорости связи между центральным и локальным радиомодемами в условиях прямой видимости при различных зна-

чениях расстояния между устройствами. Четверть-волновые антенны модемов во время эксперимента были закреплены на высоте 1,5 м от земли.

В результате эксперимента была получена зависимость скорости передачи данных от расстояния (рис. 8). Для каждого значения расстояния проводили 50 экспериментов, в которых выполнялась передача файла размером 6 Кбайт.

Математическая модель протокола канального уровня позволяет с помощью формул (5) и (6) вычислить вероятность p_B искажения одного бита для данного значения скорости передачи данных. Зависимость p_B от расстояния изображена на рис. 9.

Табл. 2 объединяет результаты измерения скорости передачи данных, вычисления вероятности p неудачного опроса локального модема, а также вероятности p_B искажения одного бита информации.

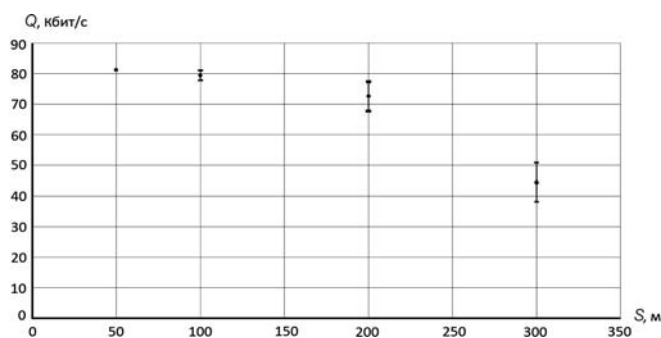


Рис. 8. Зависимость скорости передачи данных от расстояния с указанием доверительного интервала

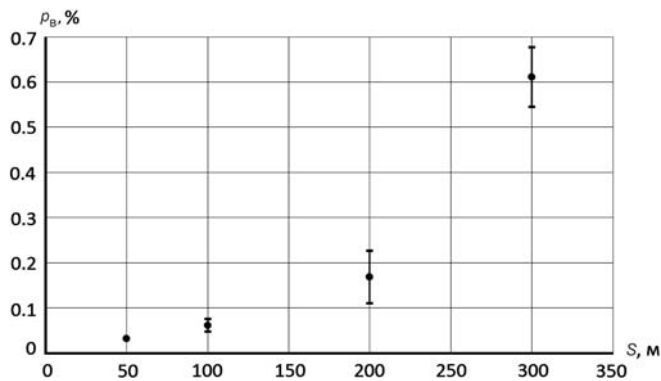


Рис. 9. Зависимость вероятности искажения одного бита информации от расстояния с указанием доверительного интервала

Таблица 2

Результаты измерения скорости передачи данных

S , м	Q , бит/с	p , %	p_B , %
50	81195 ± 237	$2,4 \pm 0,3$	$0,032 \pm 0,003$
100	79326 ± 985	$4,6 \pm 1,2$	$0,061 \pm 0,016$
200	72537 ± 3825	$12,8 \pm 4,6$	$0,168 \pm 0,06$
300	44541 ± 4032	$46,5 \pm 4,9$	$0,611 \pm 0,064$

При обработке результатов измерений было принято значение доверительной вероятности $\alpha = 0,9$, учитывая число измерений для каждого значения расстояния $n = 50$, коэффициент Стьюдента $t_{\alpha n} = 1,6759$. Из табл. 2 видно, что в результате увеличения вероятности искажения бита информации, ширина доверительного интервала увеличивается с расстоянием.

Заключение

В статье приведен обзор технологий построения радиосетей, использующих различные частотные диапазоны, представлена информация о затухании сигналов соответствующих диапазонов в различных условиях.

Предложен метод построения радиомодема на основе микросхемы Semtech SX1233. Этот метод позволяет передавать данные со скоростью до 80 кбит/с в диапазоне LPD (433 МГц). Устройство выполнено в металлическом корпусе, управление радиомодемом осуществляется с помощью компьютера через порт USB.

Описана математическая модель протокола канального уровня, позволяющая найти аналитическую зависимость между вероятностью искажения одного бита информации и скоростью передачи данных. Приведен вывод выражения для определения среднего времени доставки пакета данных от локального модема. Проанализирована связь между вероятностью искажения одного бита информации и вероятностью неудачного опроса.

Проведены испытания радиомодема, приведен анализ результатов испытаний на основе представленной математической модели, значение доверительной вероятности было принято равным 0,9.

Работа выполнена при финансовой поддержке Министерства образования и науки РФ (проект № 2010-218-01-192).

Список литературы

1. **IEEE Standard 802.11** Institute of Electrical and Electronics Engineers, Inc., 2007. — 1184 с.
2. **IEEE Standard 802.15.1** Institute of Electrical and Electronics Engineers, Inc., 2005. — 580 с.
3. **IEEE Standard 802.15.4** Institute of Electrical and Electronics Engineers, Inc., 2003. — 670 с.
4. **Шерстюков О. Н., Рябченко Е. Ю., Гаязутдинов А. Р., Мартыничук С. Л.** Применение беспроводных сетевых технологий в системах сбора сейсморазведочных данных // Георесурсы, 2011. № 6 (42). С. 50—56.
5. **Решение ГКРЧ** от 6 декабря 2004 г. № 04-03-04-001.
6. **SX1233** Datasheet. 83 с. URL: <http://www.semtech.com>.
7. **TTL-232R**. TTL to USB Serial Converter Range of Cables Datasheet. 25 p. URL: <http://www.ftdichip.com>.
8. **Шварц М.** Сети связи: протоколы, моделирование и анализ. В 2-х ч. Ч. 1: пер с англ. М.: Наука. Гл. ред. физ.-мат. лит., 1992. 336 с.

Е. М. Иванова, канд. техн. наук, доц.
 Московский институт
 электроники и математики
 Национального исследовательского университета
 "Высшая школа экономики"
 (МИЭМ НИУ ВШЭ),
 e-mail: theposte@mail.ru

Сравнительная оценка производительности вычислительных систем

Рассматриваются критерии сравнительной оценки различных вычислительных систем. Особое внимание уделяется такому критерию как производительность. Приводится анализ достоинств и недостатков различных методик и единиц измерения производительности.

Ключевые слова: вычислительные системы, быстродействие, производительность, пиковая производительность, предельная производительность, номинальная производительность, вычислительная мощность, рейтинг MIPS, рейтинг FLOPS, LINPACK, SPEC

Введение

Современное развитие вычислительной техники, совершенствование организации компьютеров направленно на повышение их производительности, надежности и достоверности вычислений, предоставления пользователям дополнительных сервисных услуг. Эти параметры в значительной мере обеспечиваются не столько технологическими возможностями (которые небезграничны), сколько средствами распараллеливания вычислений. Все это привело к тому, что практически каждый современный компьютер имеет в своем составе несколько процессоров: центральный и сопроцессоры (видео-процессор, звуковой процессор, процессоры ввода/вывода, коммуникационный процессор, различные контроллеры и т. п.). В свою очередь, центральный процессор (ЦП) может быть многоядерным, т. е. по сути, тоже состоит из нескольких взаимодействующих процессоров. Поэтому при современном определении и использовании термина вычислительная система (ВС) стирается грань между понятиями ЭВМ (компьютер) и ВС.

1. Критерии сравнительной оценки вычислительных систем

Огромное многообразие и количество различных ВС требуют создания критериев и инструментов их сравнительной оценки. Как правило, оцениваются

и сравниваются технико-эксплуатационные характеристики ВС: производительность; отношение производительности к стоимости; емкость и тип разных видов памяти (оперативной, КЭШ, внешних запоминающих устройств); тактовая частота; разрядность процессора (длины обрабатываемых слов) и разрядность шин (количество одновременно передаваемой информации); возможности системы команд; универсальность; программная совместимость; масса; стоимость; габаритные размеры; энергопотребление и др.

Сравнительная оценка различных ВС по большинству критериев не представляет никакой сложности, так как достоверная информация предоставляется разработчиком или продавцом изделия. Но самой большой сложностью, вызывающей наибольшее количество споров и разночтений, является оценка производительности.

Производительность — это важнейший технический параметр любой вычислительной системы, обозначающий скорость обработки информации. И чем больше скорость, тем лучше показатель производительности. Но измерить этот параметр очень сложно. Даже для одного однопроцессорного компьютера эта характеристика не является величиной постоянной, а тем более для сложной вычислительной системы, включающей в свой состав несколько процессоров, ведущих обработку информации с разной скоростью.

В связи с усложнением вычислительных алгоритмов, архитектуры ВС, совершенствованием методик и нюансов оценки производительности компьютеров многочисленные исследования в данной области породили множество различных понятий и характеристик, с использованием которых делаются попытки сравнивать различные вычислительные системы. Это такие характеристики, как быстродействие, производительность (пиковая, предельная, номинальная, эксплуатационная, системная, эффективная, максимальная), вычислительная мощность, вычислительная способность (*capability computing*), вычислительная емкость (*capacity computing*), рейтинг MIPS, рейтинг FLOPS, эффективность, время ответа (*response time*), время выполнения (*execution time*), прошедшее время (*elapsed time*), время ЦП (CPU *time*), среднее число тактов синхронизации на одну команду — CPI (*clock cycles per instruction*). Каждое исследование подразумевает свои особенности методологии оценки различных ВС и даже целых классов таких систем. Но в итоге все пытаются количественно выразить скорее качественное свойство, которое достигается и за счет применяемых технологий, и за счет свойств архитектуры вычислительной системы.

2. Способы оценки производительности вычислительной системы

Еще недавно производительность любого компьютера или системы принято было напрямую связывать с понятием быстродействия (числом операций/инструкций, выполняемых машиной в единицу времени [4]). Самыми распространенными единицами измерения для количественного обозначения производительности (быстродействия) являются:

- MIPS (*mega instruction per second*) — миллион (10^6) инструкций (целочисленных команд) в секунду;
- FLOPS (*floating operations per second*) — число операций над числами с плавающей запятой в секунду*.

Очень часто для характеристики возможностей системы используют критерий пиковой производительности.

Пиковая или предельная производительность — это скорость работы процессора по выполнению собственных инструкций без учета времени обращения к оперативной памяти (ОП) за операндами и без учета операций ввода/вывода. Она определяется тактовой частотой процессора и рассчитывается как среднестатистическое число инструкций (целочисленных операций типа "регистр-регистр"), выполняемых в единицу времени (секунду) без учета их статистического веса в выбранном классе задач:

$$V_{\text{пик}} = \frac{f}{l_{\text{инстр}}^{\text{ср}}},$$

где $V_{\text{пик}}$ — предельная производительность (измеряется в MIPS = 10^6 инстр/с); f — тактовая частота центрального процессора, Гц (такт/с); $l_{\text{инстр}}^{\text{ср}}$ — средняя длительность инструкции (такт/инстр), которая рассчитывается по формуле

$$l_{\text{инстр}}^{\text{ср}} = \frac{T}{n},$$

где T — длительность выполнения программы, измеренная в тактах; n — число машинных команд (или целочисленных операций).

Под пиковой производительностью многопроцессорной вычислительной системы понимают величину, равную произведению пиковой производительности одного процессора на число N таких процессоров в системе:

$$R_{\text{пик}} = V_{\text{пик}} N.$$

При этом предполагается, что все устройства вычислительной системы работают в максимально производительном режиме. Пиковая производительность любого компьютера или системы вычисляется однозначно и просто. Значит, такой критерий удобно использовать, и он до сих пор является

базовым для сравнения высокопроизводительных ВС. Чем больше пиковая производительность, тем мощнее должна быть система, и тем быстрее (теоретически) можно выполнить любую программу, а следовательно, и решить любую задачу.

Но к сожалению, на такую оценку производительности влияют не только характеристики системы (f), но и характеристики программы ($l_{\text{ср}}^{\text{оп}}$). Можно показать, что производительность, измеренная в MIPS (для конкретной программы) — есть отношение числа машинных команд в программе к времени ее выполнения:

$$V_{\text{пик}} = \frac{f}{l_{\text{инстр}}^{\text{ср}}} = \frac{f}{T/n} = \frac{n}{T}.$$

3. Достоинства и недостатки канонических способов измерения производительности

Производительность в MIPS — это величина, обратная к среднему времени выполнения команды, и больше относится к скорости работы процессора (процессоров). Интуитивно кажется, что системы с более быстрыми процессорами имеют более высокий рейтинг MIPS. Измерение производительности в MIPS (или рейтинг MIPS) имеет свои достоинства и недостатки (даже опасность) [6].

Достоинствами такого способа оценки производительности является то, что эта характеристика понятна пользователю, ее легко рассчитать, и как правило, более быстрая машина будет характеризоваться большим рейтингом MIPS. **К недостаткам** методики можно отнести следующее. Рейтинг MIPS зависит от набора команд процессора, что затрудняет сравнение по MIPS компьютеров, имеющих разные системы команд, рейтинг MIPS даже на одном и том же компьютере меняется от программы к программе, рейтинг MIPS для системы, имеющей качественно лучшие характеристики производительности, может быть количественно меньше, чем рейтинг качественно худших систем.

Для иллюстрации последнего случая исследователи рассматривают, например, две системы: в первую систему входит сопроцессор для обработки чисел с плавающей точкой (FPU), во второй системе его нет. При отсутствии сопроцессора операции над числами с плавающей точкой (вещественными числами) реализуются с помощью специализированных подпрограмм, включающих команды целочисленной арифметики. Производительность первой системы безусловно выше, так как для обработки одной вещественной операции в ней используется всего одна команда сопроцессора взамен подпрограммы, включающей до нескольких десятков целочисленных команд в системах второго типа. Но рейтинг MIPS для первой системы будет меньше, так как в программах сокращается число быстрых целочисленных операций (выполняемых за меньшее число тактов) и увеличивается число более медлен-

* Сегодня с ростом производительности используют более крупные величины: GIPS (10^3 MIPS), MFLOPS (10^6 FLOPS), GFLOPS (10^9), TFLOPS (10^{12}), PFLOPS (10^{15})....

ных вещественных операций (выполняемых за большее число тактов). Для таких систем среднее время выполнения команды будет больше, и, как следствие, они имеют более низкий рейтинг MIPS — величину, обратную среднему времени выполнения команды. Подобные странности наблюдаются и при использовании оптимизационных техник для компиляторов, когда в результате оптимизации производительность увеличивается, а рейтинг MIPS уменьшается, так как сокращается число выполняемых в программе команд.

Рассмотрим другую единицу измерения производительности — FLOPS (или рейтинг FLOPS). Идея такого критерия возникла, когда сравнивали различные типы компьютеров для решения научно-технических задач. В подобных алгоритмах велика доля вычислений с действительными числами, для которых используется тип данных с плавающей точкой.

Если использовать единицы измерения FLOPS (вместо числа инструкций MIPS), то в приведенных выше формулах размерность [инстр] заменяется на [оп] и рассчитывается новое $l_{\text{ср}}^{\text{оп}}$ [такт/оп]

$$R_{\text{пик}} = N \frac{f}{l_{\text{ср}}^{\text{оп}}}$$

Отметим **достоинства** этого метода. Рейтинг FLOPS предназначен для более справедливого сравнения различных машин между собой, так как зависит от числа выполненных операций, а не от числа выполненных машинных команд. Многие программисты полагают, что одна и та же программа, реализованная на различных компьютерах, может выполняться за различное число команд, но всегда за одно и то же число вещественных операций.

Недостатки измерения производительности во FLOPS тоже достаточно очевидны.

Во-первых, эта характеристика предназначена для оценки производительности программ, состоящих из вещественных операций, а значит, плохо подходит для другого класса задач. Например, рассматривая системы с разной производительностью, и оценивая их рейтинг FLOPS для программ компиляторов всегда получим результат, близкий к нулю. Это происходит потому, что компиляторы (трансляторы программ с языка высокого уровня в эквивалентные программы на Ассемблере или в машинном коде) в основном используют команды обработки строк, логические и специальной арифметики и редко содержат вещественные операции.

Во-вторых, наборы вещественных операций, как и системы команд различных компьютеров в целом, не всегда совместимы. Исследователи приводят такой наглядный пример. В суперкомпьютерах фирмы *Cray Research*, чья высокая производительность не подвергается сомнению, рейтинг FLOPS может оказаться низким для программ, содержащих большую долю операций деления. Вместо одной операции деления (которая в этих ВС отсутствует) ис-

пользуются две операции: вычисление обратной величины числа с плавающей точкой и умножение делимого на обратную величину.

В-третьих, рейтинг FLOPS зависит не только от вычислительной системы, но и от программы, на которой он вычисляется. Рейтинг FLOPS изменяется в зависимости как от доли целочисленных операций по сравнению с долей операций с плавающей точкой в программе, так и от доли быстрых и медленных операций с плавающей точкой. Например, программа со 100 % операций сложения (быстрые операции) будет иметь более высокий рейтинг, чем программа со 100 % операций деления (медленные операции).

4. Влияние различных параметров на количественную оценку производительности

Таким образом, производительность ВС (процессора) зависит от трех параметров:

- тактовой частоты (частоты синхронизации);
- среднего времени выполнения команды;
- числа выполняемых команд.

Все указанные параметры взаимосвязаны. Частота синхронизации (тактовая частота) определяется технологией аппаратных средств и архитектурой ВС. Среднее число тактов на команду зависит от функциональной организации процессора и архитектуры системы команд. Число выполняемых в программе команд определяется архитектурой системы команд и технологией компиляторов. Когда сравниваются две машины, необходимо рассматривать все три параметра, чтобы понять относительную производительность этих машин.

Следует также различать пиковую и реальную производительность. Строго говоря, пиковая производительность — это величина теоретическая и недостижимая для реального приложения (в отсутствии конфликтов обращения к памяти при равномерной загрузке всех устройств).

В реальных условиях на выполнение конкретной программы очень сильно влияет время взаимодействия с памятью, что определяется строением, объемом и способом доступа к устройствам памяти. В большинстве современных ВС используется многоуровневая иерархическая память с разграничением по скорости доступа и объему (регистры, КЭШ, оперативная память, виртуальные и жесткие диски, разнообразные запоминающие устройства внешней памяти). Наиболее часто используемые данные хранятся в быстрой и небольшой памяти верхнего уровня иерархии с минимальным временем доступа. Требуемые данные сюда подгружаются заранее и на их поиск не тратится лишнее время.

Из других параметров, влияющих на реальную производительность, можно назвать, например, объем и число уровней КЭШ-памяти (особенно для задач, обрабатывающих большие массивы данных, а значит, часто обращающихся к ОП в вычислениях).

КЭШ — это ассоциативная память, построенная на принципах обеспечения высокой скорости доступа к данным, она хранит копии блоков ОП. Чем больше объем КЭШ-памяти, тем выше вероятность, что нужные задаче данные будут выбраны из более быстрой КЭШ-памяти и не придется обращаться к более медленной ОП. Если все-таки это произошло, то обращение во второй уровень КЭШ, а затем и в третий, и возможно, в четвертый (если таковые имеются) произойдет гораздо быстрее, чем в ОП.

Для учета этих факторов можно использовать такой критерий, как номинальная производительность — скорость процессора с ОП. Он определяется средним числом операций, выполняемых подсистемой "процессор—память" с учетом статистического веса каждого типа операций без учета операций ввода/вывода:

$$V_{\text{ном}} = V_{\text{пик}}\alpha,$$

где $V_{\text{ном}}$ — номинальная производительность; α — некий собирательный коэффициент, учитывающий долю операций доступа к памяти в общем числе операций, и их скорость.

Реальная производительность вычислительной системы, достигаемая на данной программе (приложении), зависит также от взаимодействия программной модели, в которой реализовано приложение, с архитектурными особенностями машины, на которой это приложение запускается. Безусловно, важнейшей характеристикой вычислительной системы, от которой напрямую зависит ее производительность, является тактовая частота. Однако системы с одной и той же тактовой частотой могут иметь различную производительность, так как на выполнение одной операции разным системам может потребоваться различное число тактов (обычно от долей такта до десятков тактов). Кроме того, системы, использующие конвейерную и параллельную обработку, могут на одних и тех же тактах выполнять одновременно несколько операций.

5. Альтернативные методики оценки производительности вычислительной системы

Вывести теоретически формулу для расчета производительности с учетом всех вышеназванных факторов не представляется возможным. В связи с этим многие исследователи [3, 5—7] склоняются к использованию в качестве единицы измерения производительности компьютера **время** (в секундах) выполнения реальных программ: компьютер, выполняющий тот же объем работы за меньшее время, является более быстрым (меньшее время подразумевает большую производительность). В качестве такой программы используется широко распространенный тест LINPACK (LINear algebra PACKage) [8], который реализует алгоритм решения системы линейных алгебраических уравнений с плотно заполненной матрицей коэффициентов большой раз-

мерности (первоначально 100×100 , в последнем варианте 1000×1000) методом исключения Гаусса с числами в формате с плавающей запятой двойной точности. Для многопроцессорных систем также имеются параллельные версии LINPACK. Этот тест достаточно прост и понятен, и используется для формирования списка Top500 — пятисот самых мощных компьютеров мира. Однако LINPACK имеет существенный недостаток: программа распараллеливается, поэтому невозможно оценить эффективность работы коммуникационного компонента суперкомпьютера.

Кроме часто критикуемого канонического теста LINPACK (за недостаток системности, переоценивание роли процессоров и КЭШ-памяти, недооценивание роли системы коммутации) предлагаются другие тестовые программы: Livermore Loops, Dhrystone, NAS Kernel Benchmark Program и более поздние ScaLAPACK, NAS Parallel Benchmarks, Perfect Club, ParkBENCH, SLALOM, HINT [5, 6, 8—11]. Кроме того вместо теста LINPACK Джеком Донгарра* разработан более эффективный тест LAPACK (Linear Algebra PACKage), который пока массово не принят.

Такое обилие тестовых программ привело к тому, что в 1988 г. учредили бесприбыльную корпорацию SPEC (*Standard Performance Evaluation Corporation*) [12]. Основной целью этой организации является разработка, поддержка и публикация стандартизованного набора специально подобранных тестовых программ для оценки производительности новейших поколений высокопроизводительных компьютеров.

Однако окончательного решения проблема оценки производительности параллельных суперкомпьютеров пока не получила. По этому поводу часто цитируется доклад и статья Дэвида Бэйли [13] (Twelve Ways to Fool the Masses When Giving Performance Results on Parallel Computers) от 1991 г. За прошедшее время число тестовых программ, а значит и способов этого своеобразного жульничества только увеличилось.

В заключение хотелось бы отметить, что в будущем во избежание недоразумений, ошибочных оценок, спекуляций будут использоваться оценки мощности вычислительной системы одновременно по нескольким различным показателям. Это могут быть и названные выше показатели, но также возможно появление новых измеряемых или рассчитываемых величин в качестве таких оценок. Например, в публикуемых списках рейтинга самых мощных компьютеров мира "Top500" [5] приведены три показателя — R_{PEAK} , R_{MAX} и эффективность (в %) использования вычислительной мощности системы. R_{PEAK} вычисляется как сумма производительности отдельных процессоров, т. е. это уже

* Джек Донгарра — составитель известного списка "Top 500" — рейтинга мощнейших суперкомпьютеров мира.

известная теоретическая оценка — предельная производительность, вычисленная по приведенным выше формулам во FLOPS. Второй показатель R_{MAX} (также во FLOPS) — измеренная мощность (наивысший результат, полученный при использовании системы тестов LINPACK). Недостатки таких расчетов и измерений мы уже оговорили. А вот третий показатель — эффективность — относительно новый. Он вычисляется как отношение измеренной производительности к расчетной:

$$E = \frac{R_{max}}{R_{peak}} \cdot 100 \%,$$

и представляется наиболее справедливым, так как дает представление о реальной нагрузке разработанной вычислительной системы по сравнению с ее гипотетическими возможностями.

Список литературы

1. **Большой энциклопедический словарь.** 2-е изд.-е., перераб. и доп. М.: Большая Российская энциклопедия, СПб.: Норин, 1997.

2. **Шнитман В.** Современные высокопроизводительные компьютеры: Информационно-аналитические материалы Центра информационных технологий, 1996 г. URL: <http://citforum.ru/hardware/svk/contents.shtml>

3. **Таненбаум Э.** Архитектура компьютера. 5-е изд. СПб.: Питер, 2011.

4. **Top500** — список 500 самых быстрых в мире. URL: <http://www.top500.org/>

5. **Хамахер К., Вранешич З., Заки С.** Организация ЭВМ. СПб.: Питер, Киев: BHV, 2003.

6. **Тесты** производительности процессора. URL: <http://www.parallel.ru/computers/benchmarks/perf.html>

7. **NAS Parallel Benchmarks.** URL: <http://www.nas.nasa.gov/publications/npb.html>

8. **Сайт CSA** (Computational Science Alliance)/раздел Сравнительная производительность. URL: <http://www.csa.ru/CSA/performance1.shtml>

9. **Черняк Л.** Флопсы и лошадиные силы // Открытые системы. 2011. № 07. URL: <http://www.osp.ru/os/2011/07/13010474/>

10. **Сайт корпорации SPEC.** URL: <http://www.spec.org/>

11. **Bailey D. H.** Twelve Ways to Fool the Masses When Giving Performance Results on Parallel Computers, Ref: Supercomputing Review. Aug. 1991. P. 54—55. URL: <http://www.pdc.kth.se/training/twelve-ways.html>. Пер. на русск.: Двенадцать способов обмана, представляя производительность параллельных компьютеров. URL: <http://favorit-studio.com/novostu-vusokix-technology/desyat-sposobov-obmana-na-rezultatax-izmereniya-proizvoditelnosti-gpu.html>

УДК 004.384

С. Г. Бобков, д-р техн. наук, зав. отд.,

П. А. Чибисов, зав. сектором

НИИ системных исследований РАН, г. Москва

e-mail: bobkov@cs.niisi.ras.ru

Повышение качества тестирования высокопроизводительных микропроцессоров методами встречного тестирования с анализом функционального тестового покрытия выделенных приложений

Современные методы тестирования и верификации сложных микропроцессоров не позволяют избежать многочисленных ошибок в проекте. Предлагается методика повышения качества тестирования таких микропроцессоров за счет введения в маршрут тестирования предложенного метода "встречного" тестирования с анализом полного функционального покрытия выделенных задач для ответственных применений.

Ключевые слова: "встречное" тестирование микропроцессора, план верификации, тестирование, управляемое покрытие, метрика полноты тестирования, критерии готовности проекта

Введение

Число транзисторов в современных высокопроизводительных микропроцессорах достигает нескольких миллиардов. Большие объемы СБИС приводят к неизбежности наличия ошибок в микропроцессорах. Главной причиной увеличения числа ошибок является усложнение микропроцессоров и, соответственно, увеличение числа возможных комбинаций состояний и их взаимозависимости у микропроцессоров. То есть, чем проще микропроцессор, тем меньше вероятность наличия ошибок в нем, и чем сложнее — тем больше вероятность ошибок. Непрерывная погоня за повышением производительности и требование сохранения программной совместимости приводит к неизбежному усложнению микропроцессоров и, соответственно, к появлению все новых и новых ошибок. Так, по данным на апрель 2005 года у микропроцессора Pentium 4 [1] ошибок и несоответствий документации — 100, но исправлено — 49, не будут исправляться — 40, неизвестны пути обхода — 22. В микропроцессоре Intel Core i7-600 на декабрь 2011 г. найдено уже 123 ошибки, из которых не планируется исправлять 118, неизвестны пути обхода — 62 [2]. Таким образом, задача повышения эффективности тестирования микропроцессоров на этапе их создания является актуальной задачей, значительно усложняющейся по мере усложнения микропроцессоров.

Для повышения качества тестирования создают многоуровневый и многостадийный процесс тестирования [3].

В данной работе предлагается расширение маршрута тестирования современных сложных микропроцессоров за счет предложенной методики встречного тестирования с анализом полного функционального покрытия выделенных задач для ответственных применений. Понятие "сложный микропроцессор" определяется числом возможных состояний (для любого современного микропроцессора попытка перебора всех возможных комбинаций состояний приведет к комбинаторному взрыву) и выполняемых функций, число транзисторов в таких микропроцессорах составляет более нескольких десятков миллионов для нерегулярной части. Реализация этих предложений в настоящее время возможна только с использованием операционных систем (ОС), т. е. проект микропроцессора должен пройти многочисленные стадии тестирования и позволять загрузить ОС и отдельные приложения.

Использование ОС для повышения качества тестирования

Применение операционных систем в маршруте тестирования проекта позволяет:

- использовать механизмы, недоступные на предыдущих этапах тестирования;
- оптимизировать процесс тестирования под выделенный круг задач;
- найти ошибки с редкими комбинациями состояний и последовательностями выполняемых операций [загрузка ОС (выполнение до 10^9 – 10^{10} команд микропроцессора) сама по себе является хорошим тестом].

Невысокая вероятность появления ошибок определяет иной подход в организации процесса тестирования.

Редкие комбинации состояний — это:

— редко встречаемая последовательность инструкций;

— редко встречаемое состояние микропроцессора или отдельных его узлов и следующей команды (зависимость результата выполнения инструкции от предыстории);

— редко встречаемая диаграмма синхронизации тракта передачи данных: ошибочное направление передачи данных (из/в регистров, кэш-память различного уровня или ОЗУ), появление на одной линии нескольких синхроимпульсов, неправильная длительность синхроимпульсов и др.;

— редкие систематические электрические отклонения в функционировании кристалла в критических местах — редкие комбинации сигналов на соседних проводниках, приводящие к перекрестным наводкам, редкие локальные провалы напряжения питания и пр.

подавляющая часть последних ошибок исправляется средствами САПР на этапе топологического проектирования микросхемы и зачастую может быть устранена изменением режимов функционирования микросхем.

Предлагаемая методика тестирования включает в себя три основных этапа тестирования с использованием ОС.

1. Тестирование тестами высокого уровня из созданной базы знаний, включающей, в свою очередь, несколько уровней тестирования.

2. Тестирование тестами низкого уровня из созданной базы знаний, включая перебор возможных состояний в наиболее консервативных к изменению состояний местах; определение "консервативных" мест проводится по отдельной методике.

3. Определение максимально полного круга задач потенциальных пользователей и тестирование микропроцессора под этими задачами. Сужая область тестирования, мы добиваемся более качественного тестирования в выделенных областях, что особенно важно для ответственных применений, в результате микропроцессор оптимизируется не только по производительности под выделенный круг задач, но и по надежности.

Необходимость загрузки ОС на ранней стадии разработки аппаратуры подтверждена многими авторитетными разработчиками микропроцессоров. Подчеркивается также, что успешная загрузка ОС Linux — первый индустриальный тест кристалла СБИС универсального микропроцессора. Так как время загрузки ОС на логической модели (RTL-модель) может составлять до нескольких недель в зависимости от сложности проекта, процесс распараллеливают, разбивая всю трассу выполнения кода ОС на последовательные кванты времени, каждый из которых выполняется на отдельном вычислительном узле. При этом ошибки в разрабатываемом проекте, как показывает опыт, выявляются при запуске тестов под ОС даже после огромного числа успешно прошедших псевдослучайных, направленных, регрессионных тестов [4].

Функциональное покрытие блоков разрабатываемой модели МП

В работе [5] показано, что код приложений пользователя интегрально воздействует на покрытие HDL-дизайна, увеличивая его на 10...30 % для отдельных блоков по сравнению с направленными и псевдослучайными тестами. За последние несколько лет ведущие специалисты по верификации пришли к выводу, что даже при 100 %-ном покрытии кода такими метриками как *code coverage*, *functional coverage* нельзя утверждать, что в проекте нет ошибок [6].

Если вы не измеряете какие-либо характеристики объекта, это несомненно означает, что вы не сможете что-то осмысленно в нем улучшить. Это утверждение особенно важно для процесса функ-

циональной верификации проекта, который в принципе не может быть полностью завершен в требуемое время ввиду стремящегося к бесконечности возможного числа комбинаций состояний сигналов внутри процессора. Поэтому одна из задач — это тщательный выбор точек функционального покрытия, которые должны быть подвергнуты мониторингу. Конечно, разработчики направленных тестов могут тщательно разрабатывать тесты для проверки отдельных функциональных блоков, при этом места, неохваченные тестами, определяются с помощью метрик структурного покрытия HDL-кода.

В целом, этот подход является необходимым, но недостаточным для современных сложных микропроцессоров. Первая причина заключается в том, что современные модели процессоров содержат множество одновременно (параллельно) взаимодействующих друг с другом независимых машин состояния, и то, как именно они взаимодействуют, не может быть обнаружено метрикой "покрытие кода" (*code coverage*). Вторая причина — это огромная трудоемкость и временные затраты на написание направленных тестов. В итоге, отрасль, связанная с разработкой и верификацией СБИС микропроцессоров, приняла в качестве одного из основных методов генерацию случайных стимулов на основе задаваемых ограничений (*constrained random pattern generation*). Несмотря на имеющиеся недостатки при условии сбора значений различных метрик покрытия, эта методика может быть применена для верификации на уровне блоков.

Функциональное покрытие было создано для того чтобы дать независимую оценку эффективности создаваемых псевдослучайных тестов. С помощью него стало возможным определить, какие аспекты проекта были затронуты поступающими на вход стимулами. И если задать в плане верификации, какие аспекты приоритетны, то инженер-разработчик тестов получает возможность количественно измерить ход работ по верификации. Конечно при этом важно правильно задать точки функционального покрытия, так как если их будет избыточное количество, то группа разработчиков потеряет много времени впустую, а если точек будет недостаточно (или они будут неверно выбраны), то в модели останутся необнаруженные критические ошибки. Эта область функциональной верификации справедливо может быть описана как "искусство" [7].

Повышение качества верификации RTL-модели

Для вычислительных модулей, предназначенных для специальных применений, требуется гарантия повышенного качества как программного, так и аппаратного обеспечения, т. е. требуется формализовать подход к верификации, собрав в единую цепь все имеющиеся методы и средства для минимизации рисков, связанных с описанным выше "искусством". Существует несколько промышленных

стандартов для разработчиков сложной аппаратуры ответственного применения, описывающих необходимую совокупность требований и методик, например, используемый в авиационной технике стандарт RTCA/DO-254 [8]. Совместимость со стандартом DO-254 все чаще становится обычным требованием при реализации коммерческих, промышленных и авиационных проектов. Данный стандарт охватывает весь цикл разработки аппаратуры, при этом в соответствующем разделе описываются создание HDL-модели и подход к ее верификации, базирующийся на понятии элементного анализа и тестов, основанных на функциональных требованиях (*requirement-based tests*). Термин "элементный анализ" специфичен для данного стандарта; под элементом подразумевается некоторая минимальная на данном уровне единица RTL-дизайна (то есть структуры кода, создаваемые разработчиком для реализации одной минимальной функциональной единицы). Предполагается, что разработчик в своем созидательном процессе изначально мыслит не в терминах строк кода, а рассматривает части функциональности, которые ему предстоит создать, и то, как они будут взаимодействовать с другими блоками для выполнения требуемых функций. Также постулируется, что элемент — это то, что подлежит проверке, и то, что должно быть оценено выбранной метрикой покрытия. Заметим, что такое определение схоже с определением, вводимым в верификации на основе встроенных утверждений (*assertion-based verification*).

При этом все такие единицы связаны с требованиями, перечисленными в тестовом плане, и на каждое требование должен быть создан соответствующий тест. Элементный анализ был создан, как попытка ответить на вопрос о полноте тестирования и о критерии успешного завершения верификации. Предположим, что у нас есть абсолютно полный список требований к функциональности проекта, а также идеальный критерий, согласно которому точно определено условие окончания работ по верификации проекта, тогда мы просто можем сделать следующее:

- 1) создать все необходимые тесты (независимо от реализации HDL-дизайна) для проверки всех требований;
- 2) выполнить этот набор тестов, контролируя полноту тестирования с помощью идеальной метрики;
- 3) проанализировать полученные результаты.

В итоге, если какой-либо элемент в коде остался неотмеченным метрикой покрытия, то это может означать, что:

- этот элемент является излишним (и его следует удалить);
- нет функционального требования, связанного с этим элементом, либо нет теста, затрагивающего этот элемент (следует добавить тест либо требование);
- этот код является тестовой логикой либо логикой, помогающей восстановить ошибку.

Таким образом, согласно стандарту DO-254 вводится понятие "функциональный элемент", а также разрабатывается метрика полноты тестирования и критерий успешного завершения верификации, основанный на метрике покрытия кода RTL-модели тестами, созданными по функциональным требованиям. Для более точного анализа покрытия сюда добавляются покрытие ветвей исходного HDL-кода (*branch coverage*), покрытие автоматов с конечным числом состояний (*finite state machine coverage*), а также функциональное покрытие, основанное на встроенных утверждениях. Отметим, что метрики покрытия не всегда могут дать уверенность в работоспособности проекта, поэтому проект тестируется, в том числе пользовательскими приложениями под ОС.

Тестирование ядра микропроцессора под ОС для нахождения как можно большего числа ошибок перед изготовлением СБИС можно ускорить включением в маршрут функциональной верификации прототипа на базе микросхем ПЛИС (ПЛИС-прототипа). Это позволяет проверить работоспособность разрабатываемого ядра в условиях, приближенных к реальным с точки зрения асинхронных процессов, таких как транзакции на внешних интерфейсах, взаимодействия ядра с периферией, возникающих в системе прерываний, изменения областей памяти через DMA-каналы. При этом во время загрузки ОС задействуется большое количество встроенных в HDL-код утверждений (*assertions*).

В результате анализа найденных тестами под ОС ошибок, а также на основании изучения кода реальных задач пользователя под ОС, создаются новые шаблоны для генератора псевдослучайных тестов, направленные на проверку недостаточно охваченных функциональных областей согласно плану верификации.

Критерии готовности проекта к передаче в изготовление

Факт загрузки ОС и работоспособность приложений пользователя под ОС является одним из критериев готовности проекта к передаче в изготовление (*tape-out readiness*). Такими критериями, как правило, являются:

- выполнение плана верификации, включающего в себя отсутствие ошибок на регрессионной базе направленных тестов;
- достижение требуемой производительности на выбранных тестах в ПЛИС-прототипе;
- отсутствие ошибок на 10^5 псевдослучайных тестов для каждого имеющегося шаблона тестов;
- достижение определенных значений покрытия кода и функционального покрытия;
- HDL-код зафиксирован и не изменялся заданное время (3–4 недели);
- новые ошибки перестали поступать за определенное время (2–3 недели);

- нет ошибок во время загрузки ОС и при работе определенного набора приложений пользователя под ОС;
- исчерпано доступное согласно плану время на разработку.

Существуют также варианты оценки числа найденных ошибок в проекте и функции распределения вероятности возникновения ошибок (распределение Рэля — пример показан на рис. 1).

Число ошибок оценивается сверху как число строк HDL-кода/150 для достаточно опытных групп разработчиков (пример функции показан на рис. 2) [9].

В случае выполнения перечисленных критериев, а кроме того, успешной работы ОС и выполнения тестов под ОС, проект передается на следующий по маршруту проектирования этап для изготовления кристалла СБИС. То есть метрики, позволяющие судить о готовности проекта, основываются, в том числе на отсутствии ошибок в тестах под ОС для определенных задач за фиксированное время.

В случае ответственных применений микропроцессор испытывается под определенным кругом задач заданное время в определенных температурных и прочих условиях — для подобных испытаний выбираются приложения пользователя по определенной методике.

При измерении продвижения процесса верификации RTL-модели микропроцессора с использованием различных метрик покрытия необходимо проводить анализ и интерпретацию результатов

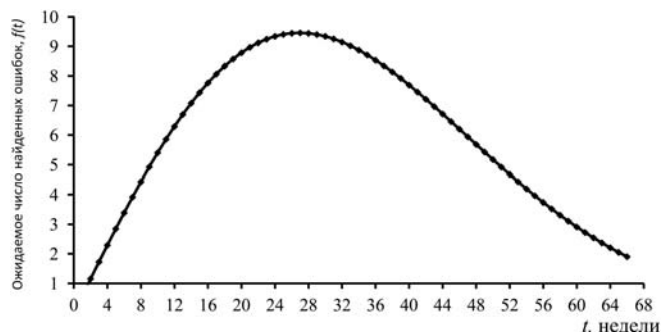


Рис. 1. Функция распределения вероятности возникновения ошибок

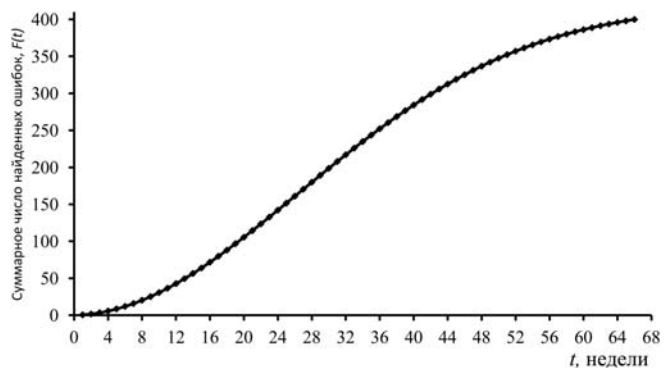


Рис. 2. Оценочный рост числа ошибок как функция от времени разработки проекта

измерений. Анализ динамики степени покрытия позволяет оценить эффективность процесса верификации, которая возрастает при достижении новых областей покрытия, а также при минимизации наблюдаемых дублирований уже достигнутых областей в единицу времени. Анализ получаемых значений позволяет контролировать пробелы покрытия отдельных блоков, уменьшая их расширением пространства входных воздействий, а также путем введения ограничений на запрещенные ситуации, требующие исключения из пространства покрытия.

Таким образом, создается первый контур обратной связи по измеряемым значениям покрытия в рамках одной выбранной модели покрытия (*coverage-directed generation*). В дополнение к нему вводится в рассмотрение второй путь обратной связи, замыкающийся через расширение и уточнение модели покрытия по результатам интерпретации результатов измерений.

Согласно теории, рассмотренной в работе [10], изначально создается так называемая исследовательская модель пространства покрытия, назначение которой — показать, как области всего пространства покрытия отображаются на функциональные требования проекта. С помощью нее проводится отбор релевантных метрик, которые затем уточняются и расширяются. Для микропроцессора, создаваемого для определенного класса задач и имеющего определенный набор инструкций, в качестве одной из метрик покрытия может быть рассмотрено подмножество всего набора инструкций, создаваемое компилятором языка высокого уровня. Далее будет показано применение такой метрики в процессе верификации модели микропроцессора.

Для решения выделенных задач, а также для ответственных применений может потребоваться спе-

циализированный процессор, который оптимизируется под класс задач. Наиболее простой и в то же время эффективный механизм повышения производительности микропроцессора для выделенных задач — включение в его состав вычислительных сопроцессоров. Для разработки и отладки ПО в условиях, приближенных к реальным, а также для оценки производительности такого процессора, может потребоваться работающий прототип на как можно более раннем этапе проектирования. При этом разработка процессора может проводиться итерационным путем с последовательным усложнением ядра, внедрением новых блоков, реализующих новые алгоритмы и механизмы. В предлагаемой методике необходимо также введение возможности перенастраивания либо отключения отдельных функций в случае, если критичная ошибка в соответствующем аппаратном узле не обходится программными средствами.

Методика тестирования микропроцессоров пользовательскими приложениями под ОС

Метод тестирования микропроцессоров пользовательскими приложениями под ОС позволяет в случае успешного многократного прохождения сложных тестов дать уверенность в правильном функционировании всей системы, требующемся для успешного выполнения пользовательских задач, что позволяет этот метод использовать при решении вопросов управления рисками на предприятии.

На рис. 3 приведена в графической форме рассмотренная методика тестирования (сплошными линиями показаны переходы между уровнями тестирования для разных уровней абстракции проекта, пунктирными — переходы в случае нахождения ошибки).



Рис. 3. Методика тестирования микропроцессора на разных стадиях разработки

Как видно из рис. 3, методика охватывает практически все уровни абстракции представления модели; при этом, в случае нахождения ошибки, например в ПЛИС-прототипе, требуется воспроизвести и исправить ее на более низком уровне, после чего повторить цикл тестирования. Под полевыми испытаниями понимается изучение стабильности работы СБИС микропроцессора вне лабораторных (нормальных) условий, т. е. в полном рабочем диапазоне температур, в составе многопроцессорных кластеров под продолжительной вычислительной нагрузкой и т. п. Под указанными на рис. 3 "направленными/регрессионными тестами" следует понимать все имеющиеся в базе тестов направленные тесты (*directed tests*), переборные тесты, а также псевдослучайные тесты с зафиксированным зерном (*seed*), находившие когда-либо ошибку, объединенные в автоматизированную тестовую базу регрессии.

Методика встречного тестирования микропроцессоров, их моделей и прототипов

В случае, когда микропроцессор разрабатывают для работы в системах ответственного применения, в рассмотренный выше маршрут тестирования предлагается включать дополнительные этапы, связанные с получением тестовых (контрольных) задач от потенциальных пользователей, компиляцией и запуском этих программ на прототипе процессора с последующей оптимизацией аппаратной и программной частей будущего комплекса.

Такой подход (рис. 4) позволяет также заранее выявить архитектурные особенности создаваемого микропроцессора, приводящие к снижению производительности на выделенных задачах и стандартных тестах универсального микропроцессора, и дать рекомендации разработчикам аппаратной и программной частей микропроцессоров по оптимизации требуемых алгоритмов функционирования микропроцессора, а также получить информацию для создания шаблонов стохастических тестов. Назовем рассмотренную выше методику "встречным тестированием", используя аналогию с понятием "встречной оптимизации" [11], под которой понимают достижение высокой эффективности проектируемой ЭВМ на определенном классе задач при оптимизации конструкции ЭВМ и параметров реализации вычислительных алгоритмов.

Рассмотрим подробнее предлагаемую методику "встречного тестирования" (рис. 5). Пусть A — множество всех инструкций микропроцессора, определяемое его спецификацией (порядка 100 инструкций). Предположим, что буфер инструкций исследуемого МП считывает из памяти или кэш-памяти инструкций по четыре инструкции (окно выполнения), выдаваемые на декодирование и дальнейшее суперскалярное исполнение в конвейере. Протестировать все возможные комбинации инструкций из всего набора A , очевидно, не представляется возмож-

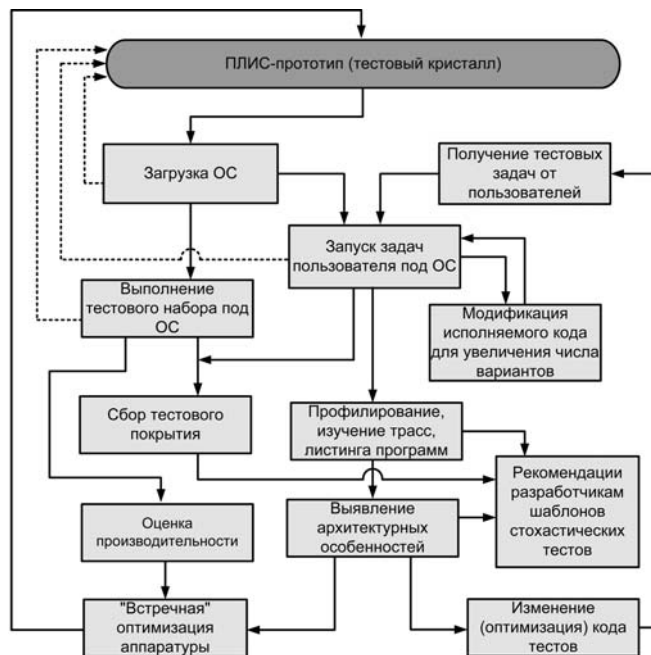


Рис. 4. Методика тестирования микропроцессора с привлечением контрольных (тестовых) задач от потенциальных пользователей ("встречное тестирование")



Рис. 5. Получение метрики функционального покрытия для "встречного тестирования" модели микропроцессора

ным (да и не требуется) за обозримое время, поэтому проведем редуцирование множества A к множеству B , опустив из рассмотрения инструкции, не создаваемые (либо создаваемые крайне редко в строго определенных ситуациях, поддающихся переборному тестированию) компиляторами программ, написанных на языках программирования высокого уровня (речь идет в основном о C, C++ и Fortran). Удаленные из рассмотрения инструкции и их комбинации тестируются обычными методами.

Проведем факторизацию групп инструкций множества B по классам эквивалентности, основываясь на спецификации микропроцессора, а также используя рекомендации группы разработчиков RTL-модели. Результатом факторизации будет множество C классов инструкций, в которое, например, будут входить 6–7 классов (такие как класс арифметических команд или класс команд ветвления), каждый из которых состоит из 1–4 групп инструкций, обозначенных как эквивалентные (в таком масштабе рассмотрения инструкция сложения ADD будет эквивалентна инструкции вычитания SUB). Объединим получившиеся группы в множество D .

В итоге, наша исследовательская (отправная) модель пространства покрытия будет состоять из всех возможных комбинаций по четыре инструкции, каждая из которых берется из множества D . Такие комбинации являются, по сути, элементарными тестовыми воздействиями для модели микропроцессора, причем все такие комбинации регистрируются, используя тестовое покрытие, задаваемое на уровне инструкций. Так как у каждой инструкции может быть множество явных и неявных аргументов и па-

раметров (так, для инструкции LOAD — это сегмент виртуальной памяти, адрес, смещение, разрядность загружаемых данных, попадание либо промах в кэш-память $L1$, $L2$, политика обращения в кэш-память и другие), то с их учетом суммарное число подлежащих тестированию комбинаций в одном окне выполнения инструкций микропроцессора будет слишком велико. Для ограничения сверху числа тестовых ситуаций и уменьшения избыточности в рамках встречного тестирования множество комбинаций параметризуют (уточняют) с учетом результатов, полученных при анализе листингов и трасс задач пользователя под ОС. Затем исходя из заданных ограничений, создается метрика функционального покрытия, основанная на группах инструкций из рассмотренного выше множества D с учетом выбранных аргументов и параметров.

Метрика задается следующим образом. Обозначим Ω — пространство всех допустимых элементарных тестовых воздействий в рамках нашей исследовательской модели. Пусть при прохождении набора тестов A на тестируемую модель подается совокупность тестовых воздействий Ω_A . Тогда значение заданной этим пространством метрики будет составлять $q(t) = |\Omega_A|/|\Omega|$. То есть мы получили функцию, ставящую в соответствие набору тестов число от 0 до 1. Тогда за один из критериев завершения тестирования можно взять достижение этой метрикой некоторого порогового значения, например 0,95.

С учетом созданной метрики функционального покрытия создаются шаблоны для генерации стохастических (псевдослучайных) тестов, в результате запуска которых могут быть найдены новые ошибки. Также накапливаются данные по покрытию, используемые для замыкания первого контура (обозначен I на рис. 5) обратной связи (по измеряемым значениям покрытия) в рамках одной выбранной модели покрытия. Данные по покрытию, получаемые непосредственно в процессе выполнения задач пользователя, также учитываются.

Как отмечалось выше, вводится и второй путь обратной связи, замыкающийся через расширение и уточнение модели покрытия по результатам интерпретации результатов измерений (обозначен II на рис. 5).

Результаты применения методики "встречного" тестирования на практике приведены на рис. 6 (▲ — тестирование с учетом рассматриваемого подхода, ◆ — без учета подхода "встречного" тестирования).

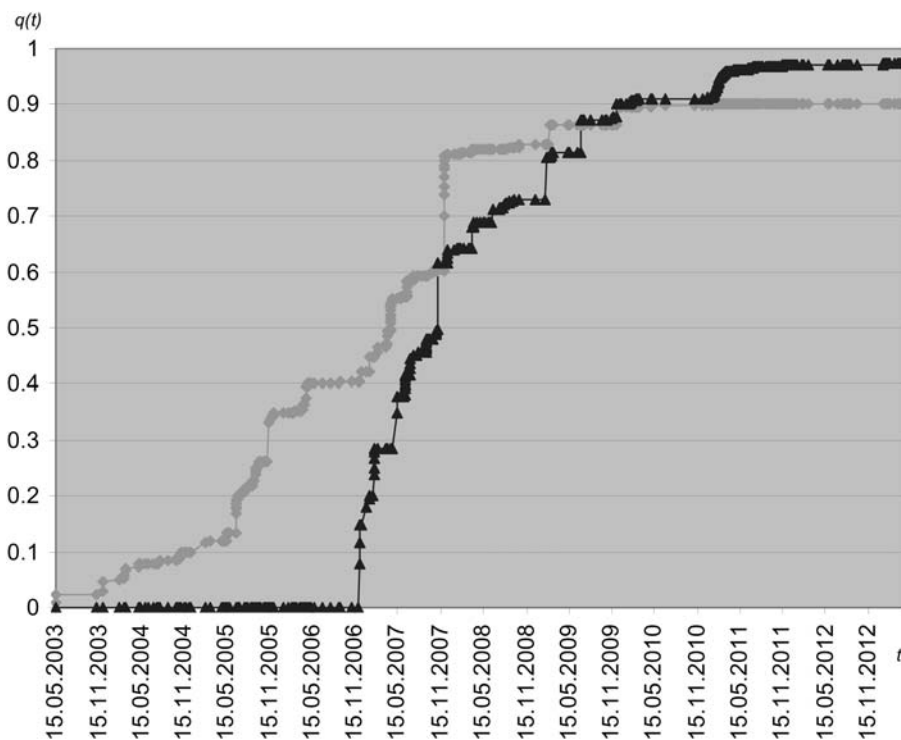


Рис. 6. Значение одной из метрик функционального покрытия в течение времени

Заключение

Применение предложенного маршрута тестирования микропроцессора с привлечением контрольных задач от потенциальных пользователей позволило поднять производительность микропроцессора 1890ВМ6Я на задачах обработки сигналов почти в 2 раза по сравнению с производительностью микропроцессора 1890ВМ5Ф. Благодаря использованию предложенной методики тестирования микропроцессора 1890ВМ6Я в реальных системах ответственного применения не выявлено ни одной ошибки данного микропроцессора на выборке 200 микропроцессоров при исследовании в течение 2 лет. В то же время на менее сложных микропроцессорах предыдущего поколения примерно 1—2 ошибки выявлялись в течение 1 года на меньшей выборке. Таким образом, предложенная методика позволяет существенно улучшить качество тестирования сложных проектов и оптимизировать архитектуру микропроцессоров под выделенный круг задач.

Список литературы

1. **Pentium 4 Processor Specification.** Материалы компании Intel. Update № 249199 — 059, 2005.

2. **Intel® Core i7-600, i5-500, i5-400 and i3-300 Mobile Processor Series.** Материалы компании Intel. Specification Update № 322814, December 2011.

3. **Бобков С. Г.** Методика тестирования микросхем для компьютеров серии "Багет". // Программные продукты и системы. 2007. № 3. С. 2—5.

4. **Чибисов П. А.** Тестирование микропроцессоров и их RTL-моделей приложениями пользователя под ОС Linux // Программные продукты и системы. 2012. № 3. С. 112—116.

5. **Balston K., Karimibiuki M., Hu A. J., Ivanov A., Wilton S. J. E.** Post-Silicon Code Coverage for Multiprocessor System-on-Chip Designs, IEEE Transactions on Computers, 25 June 2012. IEEE computer Society Digital Library. URL: <http://www.computer.org/csdl/trans/tc/2013/02/ttc2013020242-abs.html> (дата обращения: 30.05.2013).

6. **Wilson R.** Verification metrics: When is enough enough? URL: <http://www.edn.com/electronics-news/4326023/Verification-metrics-When-is-enough-enough-> (дата обращения: 30.05.2013).

7. **Bailey B.** Coverage for functional verification. URL: <http://www.eetimes.com/electronics-blogs/practical-chip-design/4401929/Coverage-for-functional-verification> (дата обращения: 30.05.2013).

8. **RTCA DO-254.** Design Assurance Guidance for Airborne Electronic Hardware, RTCA Inc., Washington, DC, 2000.

9. **Smith G.** Verification Bug Metrics. URL: http://www.dvclub.org/images/Presentations/Greg_Smith.pdf (дата обращения: 30.05.2013).

10. **Piziali A.** Functional Verification Coverage Measurement and Analysis. Boston: Springer, 2004. 232 p.

11. **Бетелин В. Б., Бобков С. Г., Зендрикова С. А., Кравченко А. А., Кушниренко А. Г., Николаев В. К.** Теоретические оценки эффективности суперЭВМ с распределенной памятью. М.: НИИСИ РАН, 2003. 180 с.

ИНФОРМАЦИЯ

8—10 ноября 2013 в г. Львов пройдет

Четырнадцатая Международная конференция
в области обеспечения качества ПО



"SOFTWARE QUALITY ASSURANCE DAYS"

Конференция охватит широкий спектр профессиональных вопросов в области обеспечения качества, ключевыми из которых являются:

- методики и инструменты тестирования ПО;
- автоматизация тестирования ПО;
- подготовка, обучение и управление командами тестировщиков;
- процессы обеспечения качества в компании;
- управление тестированием и аутсорсинг;
- совершенствование процессов тестирования и инновации.

Предыдущая 13-я конференция проходила в Санкт-Петербурге, ее участниками стали более 600 профессионалов. Организатором традиционно выступает компания "Лаборатория тестирования" (<http://www.sqalab.ru/>).

Обращаем внимание, что 10 ноября пройдет дополнительный день SQA Days English Day, в рамках которого пройдут доклады на английском языке. Это отдельное событие в рамках конференции. Число мест на этот день будет ограничено.

Сайт конференции: http://sqadays.com/index-news.sdf/sqadays/sqa_days14

УДК 004.934.2

Б. С. Агаев, Ph. d,

Институт информационных технологий,
НАНА, г. Баку,
e-mail: agayev_sayil@mail.ru

Ф. Г. Пашаев, Ph. d,

Институт кибернетики, НАНА, г. Баку,
e-mail: pasha.farhad@gmail.com

Метод оценки качества речи в корпоративных VoIP-сетях

Рассматриваются вопросы оценки качества передаваемой речи в сетях пакетной коммутации. Предложен гибридный метод, который упрощает процесс оценки качества и уменьшает погрешность результатов по сравнению с существующими методами.

Ключевые слова: речевые сигналы, IP-телефония, методы оценки качества, задержка доставки, потери пакетов, гибридный метод

Введение

В последние годы пакетные технологии передачи голоса развиваются быстрыми темпами. Если в 2006 г. объем хранимой и передаваемой в цифровом виде информации состоял из 161 Эбайт, то в 2010 г. эта цифра составила уже 986 Эбайт. Больше половины этой информации является аудиоинформацией [1].

Наряду с этим растут требования к качеству передаваемых по телекоммуникационным сетям аудио-сообщений. Поэтому разработка эффективных методов оценки качества речевых сигналов в таких сетях является важной и актуальной задачей. Эти вопросы являются одними из важнейших не только для пользователей, но и для разработчиков и создателей аудиоинформационных сетей и систем.

Вопросами качества речи, передаваемой с использованием телекоммуникационных сетей и систем, а также его оценкой, занимаются международные организации ITU-T (International Telecommunication Union), ETSI (European Telecommunications and Standard Institute) и комиссия по вопросам качества речи STQ (Speech Transmission Quality). Эти организации разработали несколько субъективных (оценка проводится экспертами) и объективных (технических — оценка проводится с использова-

нием технических средств) методов. В настоящее время субъективный экспертно-артикуляционный метод по ITU-T Rec. P.800 [2] и объективный метод E-модель по ITU-T Rec. G.107 [3] приняты как стандартизированные методологии оценки качества для классических телефонных сетей канальной коммутации (PSTN, ISDN) и пакетной телефонии (IP-телефонии). По первому методу эксперты (дикторы) специально подготовленной артикуляционной бригады озвучивают выражения из стандартной слогово-артикуляционной таблицы на входе системы. Другие эксперты (аудиторы) на выходе системы принимают их. Качество принятой речи оценивается по психофизиологическим состояниям auditors, т. е. по степени комфортности принятия аудиосообщений. В другом варианте качество определяется как отношение правильно принятых слогов к общему числу переданных по сети. Мера оценки качества речи по этой методике называется MOS (Mean Opinion Score). E-модель, как технический (инструментальный) метод, позволяет оценить качество речи вычислением функции так называемого R-фактора по некоторым сетевым параметрам:

$$R = R_0 - I_s - I_d - I_e,$$

где $R_0 = 93,4$ — базовое качество неискаженной речи, озвученной на входе сети; I_s — фактор, учитывающий степень искажений речи, вызванных эффектом "эхо" и процессом квантования оцифрованного сигнала; I_d — фактор, учитывающий суммарные сетевые задержки; I_e — фактор, зависящий от используемого оборудования.

С некоторыми гибридными методами, используемыми в VoIP-сетях, можно ознакомиться в ITU-T Rec. P.861 [4].

Постановка задачи

Наряду с определенными положительными свойствами вышеуказанные стандартизированные методы оценки качества имеют существенные недостатки [5–8], в том числе:

субъективные методы:

- наличие или привлечение экспертно-артикуляционной бригады, что требует значительных финансовых затрат;
- трудоемкость процедуры оценки и субъективизм в полученных результатах;
- невозможность автоматизировать процесс оценки качества и т. д.;

объективные методы:

- сложность учета и расчета слагаемых, входящих в E-модель;
- большая погрешность результатов и т. д.

В статье предложен гибридный метод, позволяющий упростить процесс и уменьшить погрешность процесса оценки качества передаваемой речи в корпоративных VoIP-сетях.

Решение задачи

Как известно, в классической телефонии аналоговые звуковые сигналы (в частности, речевые сигналы абонентов в сеансе обмена), передаются по сети связи по одному и тому же каналу непрерывно и синхронно, т. е. в темпе ведения речевого обмена, и принимаются таким же образом. Качество речи в таких сетях в основном зависит от амплитудно-частотной характеристики канала, уровня шума и нелинейных искажений. Эти параметры в определенном длительном интервале времени достаточно стабильны. Кроме того, для кодирования оцифрованных речевых сигналов используется один единственный метод кодирования — импульсно-кодовая модуляция (PCM — Pulse Code Modulation). Эти обстоятельства позволяют осуществить оценку качества с помощью субъективных методов, т. е. с участием экспертов по артикуляции, и согласно ITU-T эти методы — самые достоверные.

Принципы пакетной передачи информации требуют применения дополнительного сетевого оборудования (кодеков разных типов, речевых шлюзов, маршрутизаторов и т. д.) и определенных методов обработки (компрессии, пакетирования, маршрутизации и т. д.). Следовательно, возникают другие искажающие факторы, воздействующие на речевые сигналы, которые передаются по IP-сетям. Эти факторы не стабильны, т. е. их влияние может изменяться в достаточно сильной степени за короткий интервал времени. Поэтому для оценки качества передачи пакетной речи в IP-сетях, в частности VoIP-сетях, требуются критерии и методы, отличные от тех, которые используются для сетей телефонии канальной коммутации. В таблице показаны результаты оценки качества в разных системах оценки, полученные субъективно-экспертным методом согласно ITU-T Rec. P.800.

Согласно Rec. G.107, значения R-качества, определенные по любому методу, должны быть трансформированы в систему оценки MOS.

Поэтому в последний столбец таблицы добавлены значения R-качества, вычисленные по формуле $R = f(\text{MOS})$ по модифицированным в дальнейшем рекомендациям Rec.107 (последние изменения — декабрь 2011 г).

Эти преобразования удобно осуществить в соответствии с таблицей $R = f(\text{MOS})$ Rec. G.107 или по [9], так как вычисления R по формулам очень трудоемки и сложны. Однако табличные результаты, наряду с простотой использования, имеют достаточно большие погрешности. Более точное вычисление R-качества VoIP-сетей можно достичь использованием методов расчета теории систем массового обслуживания. Эти вопросы достаточно подробно рассмотрены в работе [10]. Однако полученные здесь формулы сложны и основаны на серьезных упрощениях, поэтому целесообразность использования их в практических целях вызывает некоторые сомнения.

По мнению авторов, только экспериментальным путем, объединив определенные процедурные правила объективных и субъективных методов, можно более просто и точно оценить качество передачи речи в пакетной IP-телефонии.

Как видно из таблицы, каждому психофизиологическому состоянию (степень комфортности приема) экспертов (аудиторов), принимающих тестовые речевые сообщения, соответствует определенное качество (в значениях MOS), которое в свою очередь характеризует состояние сети.

Значит, если в цепи передачи эксплуатируемой сети искусственным путем эмулировать совокупность значений воздействующих факторов (параметров) на качество речи, то соответствующее состояние экспертов можно оценить как значение качества (в MOS) сети в данный период времени. Таким образом, используя разовую услугу экспертной бригады, можно экспериментальным путем определить взаимосвязь между некоторыми сетевыми параметрами и качеством передаваемой речи, т. е. создать "метрику качества" сети. В дальнейшем периодическую работу экспертов можно заменить измерительно-вычислительным процессом по определению соответствующих параметров и нахождением значений качества по этой метрике. Целесо-

Результаты оценки качества

Психофизиологическое состояние экспертов при приеме	Система оценки		
	Словесная	MOS	R
Полное отсутствие напряженности при разговоре	Наилучшая (best)	4,34...4,50	$90 < R < 100$
При разговоре следует быть внимательным, но не требуется никаких заметных усилий	Высокая (high)	4,03...4,34	$80 < R < 90$
При разговоре требуется постоянное, но не исключительное напряжение внимания	Средняя (medium)	3,60...4,03	$70 < R < 80$
Разговор возможен, но испытываются большие затруднения	Низкая (low)	3,10...3,60	$60 < R < 70$
Разговор невозможен или возможен, но с очень большими трудностями	Плохо (poor)	2,58...3,10	$50 < R < 60$

образность такой замены с технологической, финансовой и других точек зрения не вызывает сомнения, так как измерение (вычисление) параметров сети намного проще, чем обеспечение работы экспертной бригады или же применение сложной Е-модели.

Согласно ITU-T Rec. Y.1541 [11] (дополнительно [10, 12]), основными факторами, воздействующими на качество речи в IP-сетях, являются:

- суммарная задержка сети (IP TD);
- вариация задержки пакетов (IP DV);
- коэффициент потери пакетов (IP LR);
- коэффициент пакетов с ошибками (IP ER).

Здесь под термином "вариация задержки пакетов" понимается джиттер пакетов, т. е. максимальная разница во времени поступления пакетов на оконечный абонентский терминал (например, в речевой шлюз в варианте абонентского подключения "телефон—телефон"), которая возникает из-за очереди и маршрутизации пакетов одного фрагмента разговора по разным путям. Пакеты, задержки которых больше установленного размера джиттера (30...50 мс), отбрасываются, т. е. не участвуют в депакетизации. Поэтому джиттер можно воспринимать одним из составляющих суммарной задержки. Так как современные сетевые технологии обеспечивают надежную защиту от ошибочных битовых преобразований в пакетах ($IPER < 10^{-10}$), то можно не учитывать его искажающего воздействия на качество речи. Вследствие этого в тестовых испытаниях в качестве воздействующих факторов можно рассматривать только суммарную сетевую задержку (t) и потери пакетов (n).

В экспериментах использован стандартный кодек G.711 (64 кбит/с), так как корпоративная сеть обладает достаточной пропускной способностью и нет необходимости в экономии полосы пропускания. Также не использованы устройства подавления пауз (VAD — Voice Activity Detection) и устройство нейтрализации потери пакетов (PLC — Packet Loss Concealing). Причина в том, что во многих случаях требуется качественная запись переговоров в корпоративных аудиоинформационных сетях, поэтому не допускаются потери пауз в речевом обмене, поскольку VAD подавляет слабые и короткие по времени шумовые и звуковые сигналы, которые могут оказаться очень информативными. Кроме того, механизм обработки информации VAD понижает качество речи приблизительно на 5...10 единиц R . Также PLC целесообразно использовать при применении низкоскоростных кодеков, так как на обработку пакетов в нем затрачивается много времени (около 40 мс), что увеличивает длительность суммарной задержки. Кроме того, как показывает практика, даже 3 %-ные одиночные потери меньше влияют на качество восприятия речи, чем возникающие задержки при обработке сигналов в этом вспомогательном устройстве.

При текущем состоянии сети (при существующих значениях t и n) оценка качества выполняется обычным образом, т. е. согласно Rec. P.800. Генерацию других значений t , n удобно осуществить с помощью соответствующих программных продуктов. Для этого, например, можно использовать программы PsyVoIP или NetIQ Chariot фирм Psytechnics и NetIQ соответственно. Эти программы могут быть установлены в компьютере администратора сети или на речевом шлюзе в зависимости от архитектуры абонентского подключения VoIP. С помощью этих программ можно генерировать речевой трафик, регулировать задержки, размер джиттера, создавать потери пакетов и т. д. Имеется возможность вычислить R по известным значениям MOS на основе функции $R = f(\text{MOS})$. Однако эти средства достаточно дорогие и с низкой достоверностью результатов. Поэтому эксперименты лучше проводить в автономном режиме.

Допустим, в ходе экспериментов получены некоторые значения зависимостей $\text{MOS} = g(t)$ и $\text{MOS} = f(n)$.

Параметры достаточно регулировать в пределах от максимально допустимого до минимально возможного для данной сети, значения которых установлены в ITU-T Rec. G.114 [13] (для аудиоинформационного трафика номинальное и максимальное значения: $t_{\text{ном}} = 150$ мс, $t_{\text{max}} = 400$ мс, $n_{\text{max}} = 3$ %, в исключительных случаях $n_{\text{max}} = 5$ %, если t достаточно низкий [14]). Максимальные уровни t и n соответствуют качеству $R = 50$ в стобалльной системе (в пятибалльной системе — $\text{MOS} = 2,28$ по Rec. G. 107). ITU-T Rec. G.109 [15] рекомендует не проектировать VoIP-задержки, потери которых хуже этих показателей.

Обычно для построения гладких кривых графиков $\text{MOS} = g(t)$ и $\text{MOS} = f(n)$ технически трудно генерировать все необходимые значения этих сетевых параметров, особенно t_{min} , n_{min} и их близлежащие значения. В этом случае необходимо с помощью известных математических методов, например, простейшего метода наименьших квадратов, восстановить (предсказать) недостающие точки графиков.

Если требуется более точно вычислить предсказанные значения данных, можно использовать другие методы, например нейронные сети. Однако эти методы имеют более сложный вычислительный алгоритм. Рассмотрим этот вариант.

Для расчетов применим трехслойную нейронную сеть с алгоритмом обратного распространения [16, 17 с. 49, 52, 58] для каждого отрезка (число отрезков k) первичного графика. В качестве функции активизации выберем сигмоиду, которая определяется по формуле

$$f(x) = \frac{1}{1 + e^{-x}}. \quad (1)$$

Структурные элементы сети определим следующим образом:

- $W1[i, j], W2[i, j], W3[i, j]$ — весовые коэффициенты 1-, 2- и 3-го слоев;
- $M1$ — число входных данных;
- $N1, N2, N3$ — число нейронов 1-, 2- и 3-го слоев; пусть $M1 = 1, N2 = 1, N3 = 1$;
- $S1[j], S2[j], S3[j]$ — параметры состояния нейронов 1-, 2- и 3-го слоев;
- $Y1[i], Y2[i], Y3[i]$ — выходные функции нейронов 1-, 2- и 3-го слоев;
- входные данные обозначим как x_i .

Входные данные t_{ki}, n_{ki} , где $k \in [1; k], i \in [1; M1]$, и выходные данные $MOS[k, i]$ берутся из полученных экспериментальных данных. Выходная функция и состояние нейронов 1-, 2- и 3-го слоев определяются по нижеуказанным формулам соответственно:

$$\begin{aligned}
 S1[j] &= \sum_{i=1}^{M1} x_i W1[i, j], j = 1, \dots, N1; \\
 Y1[i] &= \frac{1}{1 + e^{-S1[i]}}; \\
 S2[j] &= \sum_{i=1}^{N1} Y1 W2[i, j], j = 1, \dots, N2, \\
 Y2[i] &= \frac{1}{1 + e^{-S2[i]}}; \\
 S3[j] &= \sum_{i=1}^{N2} Y2[i] W3[i, j], j = 1, \dots, N3, \\
 Y3[i] &= \frac{1}{1 + e^{-S3[i]}}.
 \end{aligned} \tag{2}$$

Нейронная подсеть для каждого отрезка обучается, как самостоятельная сеть одинаковой структуры с различными коэффициентами. Эти сети образуют множество сетей. Предварительные входные и выходные данные образуют множество данных.

Если $MOS[k, i]$ — желаемый результат (предсказанные значения), тогда параметры коррекции для 1-, 2- и 3-го слоев вычисляются следующим образом:

$$\begin{aligned}
 \delta 3[i] &= (MOS[k, i] - Y3[i])(f3)'(S3[i]), i = 1, \dots, N3, \\
 \delta 2[i] &= (f2)'(S2[i]) \sum_{j=1}^{N3} w3[i, j] \delta 3[j], i = 1, \dots, N2, \\
 \delta 1[i] &= (f1)'(S1[i]) \sum_{j=1}^{N2} w2[i, j] \delta 2[j], i = 1, \dots, N1.
 \end{aligned} \tag{3}$$

Весовые коэффициенты для каждого слоя, соответственно, корректируются по следующему алгоритму.

$$\begin{aligned}
 WS1[i, j] &= W1[i, j] + \eta \delta 1[j] x[i], \\
 i &= 1, \dots, M1; j = 1, \dots, N1; \\
 W1[i, j] &= WS1[i, j]; \\
 WS2[i, j] &= W2[i, j] + \eta \delta 2[j] Y1[i], \\
 i &= 1, \dots, N1; j = 1, \dots, N2; \\
 W2[i, j] &= WS2[i, j]; \\
 WS3[i, j] &= W3[i, j] + \eta \delta 3[j] Y2[i], \\
 i &= 1, \dots, N2; j = 1, \dots, N3; \\
 W3[i, j] &= WS3[i, j],
 \end{aligned} \tag{4}$$

где η — скорость процесса обучения.

Для всех слоев ошибка обучения E вычисляется по формуле

$$E = \frac{1}{2} \sum_{i=1}^{N3} (Y3[i] - MOS[k, i])^2. \tag{5}$$

Обучение продолжается до достижения приемлемой погрешности. Вычисленные коэффициенты каждого отрезка сети после завершения обучения сохраняются отдельно в базе данных для дальнейшего использования. Далее строятся графики $MOS = g(t), MOS = f(n)$ по табличным и вычисленным промежуточным данным (рис. 1 и 2).

Если требуется трансформировать значения пятибалльной системы в 100-балльную систему R ,

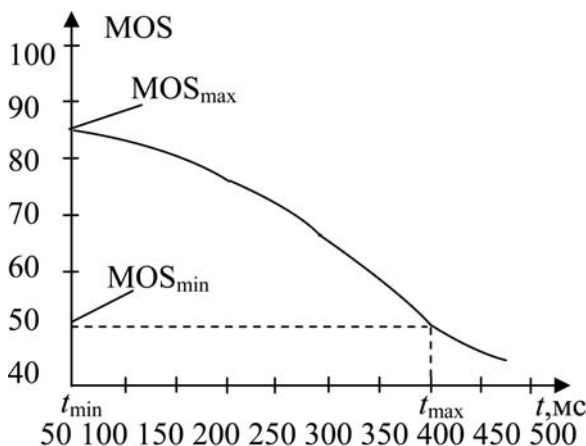


Рис. 1. Зависимость MOS оценки от суммарной задержки

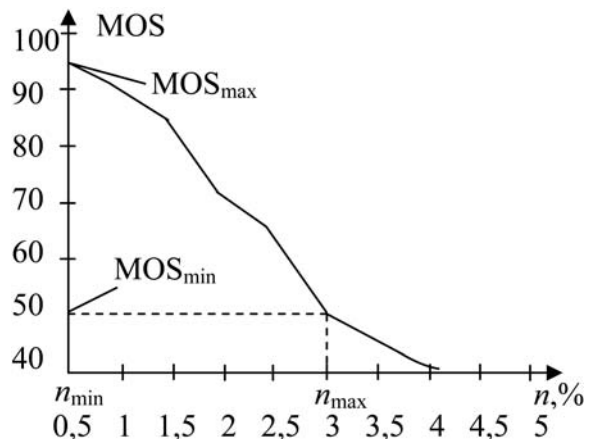


Рис. 2. Зависимость MOS оценки от потери пакетов

то лучше воспользоваться простым графическим вариантом, чем сложными математическими формулами Res.G.107. Иногда в практических целях достаточно воспользоваться грубой формулой $R = 20\text{MOS}$ (погрешность — 11 %).

Если нет в наличии программных средств для генерации промежуточных сетевых параметров, то их можно создать искусственным путем (в определенном интервале), регулируя размер буфера джиттера, длительность очереди пакетов в коммутаторах и т. д. Максимальные и минимальные значения t , n обычно возникают во время соответствующих сетевых нагрузок. Если экспертные испытания провести в эти временные периоды, то можно получить крайние значения MOS-оценок.

С точки зрения точности целесообразно учитывать совместное воздействие t и n на качество оценки. Однако осуществить это технически гораздо труднее. Наверное, по этой причине в технической литературе встречаются варианты, где используется зависимость MOS от одного из параметров при некоторых постоянных значениях другого. Поэтому здесь рассматривается решение вопроса аналитическим путем.

Допустим, даны графики зависимостей $\text{MOS} = g(t)$ и $\text{MOS} = f(n)$, полученные в процессе оценки качества речевых сигналов в VoIP-сетях. Примем, что эти функции определены в областях $t \in [t_{\min}, t_{\max}]$ и $n \in [n_{\min}, n_{\max}]$ соответственно. Множества значений обеих функций определены в общем отрезке $\text{MOS} \in [\text{MOS}_{\min}, \text{MOS}_{\max}]$, т. е.

$$[t_{\min}, t_{\max}] \xrightarrow{g} [\text{MOS}_{\min}, \text{MOS}_{\max}];$$

$$[n_{\min}, n_{\max}] \xrightarrow{f} [\text{MOS}_{\min}, \text{MOS}_{\max}].$$

Априори известно, что функции являются нелинейными и убывающими. Требуется построение такой функции $\text{MOS} = F(z)$, в которой учтена совместная взаимосвязь t и n с MOS-функцией. Для этого, в первую очередь, множество данных должно быть приведено в общий масштаб (процесс нормирования). Разделим область обеих функций на равные части в количестве M :

$$\Delta t = \frac{t_{\max} - t_{\min}}{M}; \Delta n = \frac{n_{\max} - n_{\min}}{M}. \quad (6)$$

Для крайних значений:

$$\begin{aligned} t_{\min} &= t_{\min} + 0 \cdot \Delta t; \\ t_{\max} &= t_{\min} + M\Delta t. \end{aligned} \quad (7)$$

Соответственно для n :

$$\begin{aligned} n_{\min} &= n_{\min} + 0 \cdot \Delta n; \\ n_{\max} &= n_{\min} + M\Delta n. \end{aligned} \quad (8)$$

Значения функций в крайних (граничных) точках областей известны.

Однако для некоторых z промежуточные значения $g(t_{\min} + z\Delta t)$ и $f(n_{\min} + z\Delta n)$ могут быть неизвестными. Для определения этих предсказанных (восстановленных по некоторым известным данным) значений можно использовать один из известных методов. Применим в качестве примера метод интерполяции для $g(t)$.

Допустим, для некоторого $z \in [0, M]$ существуют значения $g(t_r)$ и $g(t_{r+1})$ для которых справедливо следующее:

$$\begin{aligned} g(t_r) &> g(t_{\min} + z\Delta t) > g(t_{r+1}), \\ t_r &< t_{\min} + z\Delta t < t_{r+1}. \end{aligned} \quad (9)$$

Ясно, что это неравенство следует от условия убываемости $g(t)$. Поэтому

$$g(t_{\min} + z\Delta t) = \frac{g(t_{r+1}) - g(t_r)}{t_{r+1} - t_r} (t_{\min} + z\Delta t - t_r) + g(t_r). \quad (10)$$

Тогда функцию $F(z)$ можно определить следующим образом:

$$F(z) = \frac{g(t_{\min} + z\Delta t) + f(n_{\min} + z\Delta n)}{2}. \quad (11)$$

Область определения этой функции будет $z \in [0, M]$, множество значений $\text{MOS} \in [\text{MOS}_{\max}, \text{MOS}_{\min}]$. Например, значения функции в конечных точках области определения:

$$\begin{aligned} F(0) &= \frac{g(t_{\min} + (0) \cdot \Delta t) + f(n_{\min} + (0) \cdot \Delta n)}{2} = \\ &= \frac{g(t_{\min}) + f(n_{\min})}{2} = \frac{\text{MOS}_{\max} + \text{MOS}_{\max}}{2} = \text{MOS-} \\ &\text{max}; \end{aligned} \quad (12)$$

$$\begin{aligned} F(M) &= \frac{g(t_{\min} + M\Delta t) + f(n_{\min} + M\Delta n)}{2} = \\ &= \frac{g(t_{\max}) + f(n_{\max})}{2} = \frac{\text{MOS}_{\min} + \text{MOS}_{\min}}{2} \end{aligned}$$

В итоге $F(z)$ определяется следующим образом:

$$F(z) = \begin{cases} \text{MOS}_{\max}; & z = 0 \\ \frac{g(t_{\min} + z\Delta t) + f(n_{\min} + z\Delta n)}{2}; & z \in (0, M); \\ \text{MOS}_{\min}; & z = M. \end{cases} \quad (13)$$

Используя вычисленные значения и первичные экспериментальные данные, можно построить график $\text{MOS} = F(z)$.

Заключение

Проанализированы стандартизированные объективные и субъективные методы ITU-T и ETSI для оценки качества передаваемой речи в сетях IP-телефонии, выявлены их основные недостатки (сложность процесса оценки, большая погрешность результатов, невозможность автоматизации

процесса оценки и т. д.). Предложен более простой и точный гибридный метод, объединяющий процедурные правила обоих методов. По этому методу эксперты оценивают качество по числовым значениям основных сетевых факторов (суммарная сетевая задержка, потери пакетов), искажающих речевые сигналы (пакеты), которые инструментальным путем генерируют в сети.

Созданная разовым привлечением экспертно-артикуляционной бригады "метрика качества" $MOS = g(t)$ и $MOS = f(n)$ в дальнейшем позволит заменить процесс экспертной оценки (субъективный метод) на измерительно-вычислительный процесс (объективный метод) текущих числовых значений сети. Обоснована методика проведения оценки, показан пример использования нейронной сети для расчета некоторых предсказанных значений параметров в целях построения метрики. Предложенный технологически более простой и точный метод может быть использован для оценки качества передачи речи в корпоративных VoIP-сетях.

Список литературы

1. Харламов А. А. Редакционная статья // Речевые технологии. 2010. № 1. С. 111.
2. ITU-T Rec. P.800. Methods for subjective determination of transmission quality. URL: <http://www.itu.int/rec/T-REC-P.800/en>.
3. ITU-T. Rec. G.107. The E-model a computational model for use in transmission planning. URL: <http://www.itu.int/rec/T-REC-G.107/en>.
4. ITU-T. Rec. P.861. Objective quality measurement of telephone-band (300-3400 Hz) speech codecs. URL: <http://www.itu.int/rec/T-REC-P.861/en>.
5. Нгуен Фи Туен. Разработка объективных методов оценки качества речевой информации по экзозащищенным каналам связи. Автореферат дис. ... канд. тех. наук. М.: 1999. 170 с.
6. Курилов О.С. Объективный анализ качества речи в IP-телефонии // Технологии и средства связи. 2002. № 4. С. 76–78.
7. Муравьев В. Е. Оценка качества передачи речи в современных сетях связи // Мобильные связи. 2003. № 8. С. 43–46.
8. Терехов А. Н. Относительная методика экспресс-оценки качества передачи речи в аудиоинформационных системах // Материалы научно-технической школы-конференции. МИРЭА, М.: 2006. Часть 1. С. 175–179.
9. Байорек Крис. Качество VoIP: корреляция оценки MOS и R-фактора // Сети и системы связи. 2003. № 6. С. 98.
10. Фомин А. Ф. Оценка параметров сетей IP-телефонии // Информационные технологии. 2003. № 2. С. 22–32.
11. ITU-T Rec. Y.1541. Network performance objectives for IP-based services. URL: <http://www.itu.int/rec/T-REC-Y.1541/en>.
12. Тарасов Б. Ю., Дроздов К. И. Качество речи в IP-сетях // Сети связи. 2007. № 12. С. 39–45.
13. ITU-T Rec. G.114. One-way transmission time. URL: <http://www.itu.int/rec/T-REC-G.114/en>.
14. Коган А. Контроль качества услуг IP-телефонии: теория и практика // Технологии и средства связи. 2004. № 3. С. 42.
15. ITU-T. Rec. G.109. Definition of categories of Speech transmission quality. URL: <http://www.itu.int/rec/T-REC-G.109/en>.
16. Алиев Т. А., Гулуев Г. А., Рзаев А. Г., Пашаев Ф. Г., Юсифов И. Б., Саттаров И. Р. Помехотехнология и системы определения дебита нефтяных скважин // Доклады Национальной Академии Наук Азербайджана 2010. Т. LXVI. № 3. С. 20–28.
17. Круг П. Г. Нейронные сети и нейрокомпьютеры. М.: Изд-во МЭИ, 2002. 177 с.

ИНФОРМАЦИЯ



С 13 по 17 октября 2013 г. в Ялте
на базе комплекса "Россия" Верховной Рады
Автономной республики Крым состоится

Международная научно-практическая конференция "КИБЕРБЕЗОПАСНОСТЬ—2013"

Конференция выступает платформой диалога представителей органов государственной власти и управления, бюджетного и частного секторов производства, сферы хозяйствования и предоставления услуг, науки, образования и гражданского общества в целях обсуждения общих проблем кибербезопасности, выявления общих решений, предотвращения угроз, обеспечения эффективной защиты и безопасности активов в киберпространстве.

Формат конференции предполагает проведение пленарных заседаний, секций, круглых столов, обмен положительным опытом и укрепление сотрудничества, а также проведение семинаров повышения квалификации по международным стандартам этой сфере.

Программа работы конференции ориентирована на рассмотрение следующих вопросов:

- Общие вопросы кибербезопасности и киберзащищенности личности, организации, общества и государства.
- Развитие и гармонизация нормативно-правовой базы кибербезопасности.
- Обращение с рисками кибербезопасности в бизнесе, секторе государственных услуг и в приватной жизни.
- Кибербезопасность ключевых информационных систем объектов критической инфраструктуры и промышленных систем автоматизированного управления.
- Кибербезопасность инфраструктуры открытых ключей и юридически значимого электронного документооборота.
- Практические решения, технологии и исследования в сфере кибербезопасности.
- Кибербезопасность Smart Grid.
- Развитие осведомленности и подготовка кадров в сфере кибербезопасности

Развернутая информация по конференции представлена на сайте
www.confere.ru

УДК 004.4'22

Е. А. Паршикова, аспирант,
e-mail: inara@korolev-net.ru
Московский институт
электроники и математики

Национального исследовательского университета
"Высшая школа экономики"

Возможные подходы к автоматизации проектирования систем поддержки принятия управленческих решений¹

Рассмотрены теоретические вопросы, касающиеся систем поддержки принятия решений, проведен обзор современных тенденций в проектировании систем поддержки принятия решений и рассмотрены возможные подходы к автоматизации проектирования систем поддержки принятия управленческих решений.

Ключевые слова: системы поддержки принятия решений, автоматизация проектирования систем поддержки принятия решений

Введение

Современный этап развития общества характеризуется его глубокой и всесторонней информатизацией. Устойчивой тенденцией развития современных процессов, протекающих в любой сфере человеческой деятельности, становится усложнение задач, которые приходится решать на различных уровнях управления. Это связано с ростом объемов разнородной информации, необходимой для обоснованного принятия решений, и с сокращением времени, имеющегося для принятия решений.

Многопараметричность и многокритериальность управленческой деятельности предполагает подключение на помощь человеку инструмента, способного помочь ему в принятии решений. Одним из таких инструментов являются системы поддержки принятия решений.

Современные *системы поддержки принятия решений (СППР)* представляют собой системы, ориентированные на решение задач повседневной уп-

равленческой деятельности, являются инструментом, призванным оказать помощь лицу, принимающему решение (ЛПР). С помощью СППР может проводиться выбор решений неструктурированных и слабоструктурированных задач, в первую очередь многокритериальных [1].

Под СППР понимают человеко-машинные системы, использующие методы поддержки принятия решений, которые позволяют ЛПР использовать данные и знания объективного и субъективного характера для решения слабоструктурированных (плохо формализованных) проблем [2].

СППР обладают следующими основными характеристиками [3]:

- СППР используют и данные, и модели;
- СППР предназначены для помощи менеджерам в принятии решений для слабоструктурированных и неструктурированных задач;
- они поддерживают, а не заменяют процесс принятия решений менеджерами;
- цель СППР — повышение эффективности решений.

1. Современные тенденции в проектировании СППР

Сегодня при разработке многих классов автоматизированных систем начинает доминировать типовое проектирование за счет развития методологии типового проектирования и ее практической реализации. Инструментальные средства разработки автоматизированных систем все больше ориентируются на архитектуру готовых программных продуктов. Это объясняется требованиями по сокращению времени на создание и внедрение систем, сокращению затрат на их проектирование и трудозатрат на сопровождение эксплуатируемых приложений при их адаптации к изменениям в предметной области (ПрО).

При этом в области создания СППР можно видеть доминирование индивидуального проектирования над другими подходами. Вопросы же типового проектирования и создания инструментальных средств, направленных на сокращение времени и стоимости проектирования за счет унификации модулей СППР, не освещены в должной мере [4]. Это объясняется тем, что устойчивой тенденцией является превалирование узкоспециализированных СППР, создаваемых для конкретной предметной области под конкретную задачу [5–8]. Доминирующее развитие узкоспециализированных СППР

¹ Данное исследование осуществлено в рамках программы фундаментальных исследований НИУ ВШЭ в 2012 году.

связано с тем, что в области принятия решения не существует универсальных средств решения, поэтому для каждой задачи фактически приходится создавать свою СППР, являющуюся реализацией конкретного метода или группы методов [9]. Таким образом, перед разработчиками СППР в рамках различных подходов и направлений встает одинаковая задача создания с нуля модулей, которые являются общими для разных СППР и могли бы быть созданы единожды с последующим их использованием во множестве проектов.

Все предпосылки для разработки средств, направленных на типовое проектирование СППР, существуют [4]: проработаны архитектурные особенности СППР различных классов, определены методы принятия решений для широкого класса решаемых задач, созданы средства, успешно эксплуатируемые разработчиками автоматизированных систем.

Методология автоматизированного проектирования СППР позволит:

- получить новый подход к проектированию программного обеспечения СППР;
- создать открытые для модернизации и развития СППР;
- получить унифицированные подходы для проектирования элементов СППР;
- качественно повысить уровень и скорость проектирования, сократить расходы на проектирование;
- развить технологии автоматизации программирования приложений по типовым проектам;
- ускорить процессы разработки СППР;
- сделать использование этих систем более массовым.

2. Краткий обзор современных подходов к автоматизации проектирования систем поддержки принятия управленческих решений

Сегодня область СППР крайне коммерциализована, технологии разработки систем являются коммерческой тайной, поэтому в зарубежных источниках, как правило, описываются интерфейс и возможности предлагаемых авторами разработок, однако не раскрываются методологии их проектирования [10].

Представленные на сегодняшний день в литературе возможные подходы к автоматизации проектирования систем поддержки принятия управленческих решений классифицированы автором следующим образом:

1. Использование специальных языков для описания и последующего автоматического создания СППР.

2. Использование системы-конструктора для автоматизированного создания СППР силами опытного специалиста.

3. Использование специализированной экспертной системы, результатом работы которой должна являться готовая СППР.

4. Онтологический подход к разработке СППР.

5. Автоматизация создания СППР на основе лингвистических моделей.

2.1. Использование специального языка для описания и последующего автоматического создания СППР

Примером вышеназванного подхода является технология создания СППР на основе свободно расширяемого набора типовых элементов.

Данная технология предназначена, в первую очередь, для задач, где основой для принятия решения является информация числового характера. В основу предлагаемой в работе [9] технологии положен тот факт, что в принятии решений можно выделить ряд действий, которые семантически и с формальной точки зрения совпадают. Соответственно, для реализации такого рода действий могут быть использованы типовые элементы как некоторые структуры, формируемые специальным образом.

Использование типовых элементов предполагает наличие двух этапов при решении прикладной задачи (в данном случае это создание программного обеспечения):

1) анализ прикладной области в целях формирования набора типовых элементов (типового набора), что подразумевает выделение и описание типовых элементов, используемых при решении широкого круга задач данной прикладной области;

2) описание способа решения стоящей задачи на основе созданного типового набора.

В качестве типового элемента в данной технологии предлагается использовать *формальное описание задачи*. Полагается, что переход от формальной задачи к конкретной прикладной требует задания модели ПрО, а также метода решения задачи — правил обработки информации, содержащейся в модели. При определении типового элемента такая конкретизация не требуется. Она необходима лишь на этапе непосредственного формирования СППР или ее составных частей. В итоге предлагаемые типовые элементы могут быть представлены как оболочки с предопределенными семантикой формальных задач входами-выходами и структурой организации взаимосвязей между моделями и методами, присущей той или иной формальной задаче.

Набор типовых элементов может обеспечить решение задач не только в соответствующих формальных постановках, но и задач, которые могут быть декомпозированы и сведены к решению логически увязанной совокупности этих задач.

В качестве средства описания логики решения задач, как совокупности взаимосвязанных подзадач, в работе [9] предлагается *объектно-ориентированный язык*. Основной его особенностью является ис-

пользование в качестве особого вида оператора задач из типовых наборов, идентифицируемых своими именами. Каждой указанной задаче может ставиться в соответствие уточненная математическая постановка. Эта постановка в немалой степени является также формальной, поскольку конкретные данные для этой задачи могут формироваться лишь непосредственно при решении исходной задачи выбранным методом.

Данный подход имеет следующие *недостатки*:

- основой для принятия решения является информация числового характера, поэтому нечисловую задачу придется сводить к числовой, что может повлечь за собой ошибки и принятие неправильного решения;
- необходимость разработки набора типовых элементов и задания модели конкретной предметной области требует тесной кооперации между специалистом в ПрО и специалистом по принятию решений, так как работа выполняется в терминах теории принятия решений.

2.2. Использование системы-конструктора для автоматизированного создания СППР силами опытного специалиста

Примером такой реализации служит Web-конструктор для построения СППР [11]. Его структуру логически можно разделить на две основные зоны: *экспертную и рабочую*.

Экспертная зона включает в себя экспертный модуль и базу данных. Основные задачи, решаемые в этой зоне:

- настройка модуля для построения классификаций объектов выбора и клиентов;
- определение состава ограничений, накладываемых на объект в процессе построения конфигураций;
- разработка системы критериев для описания объекта в целом;
- формулировка общих свойств клиента и объектов выбора (если они есть);
- формулировка полного списка задач клиента;
- адекватная связь свойств объектов выбора с задачами клиента;
- заполнение и сопровождение базы данных объектов выбора.

Рабочая зона системы содержит клиентский модуль и логический модуль (алгоритмическую часть СППР), решающий следующие задачи:

- классификация клиента (классы предварительно определены в экспертной зоне);
- формирование списка задач, которые должен решать объект выбора;
- выделение ЛПР значений накладываемых ограничений (состав ограничений определен в экспертной зоне).

Алгоритм работы логического модуля СППР следующий:

- 1) строятся все возможные конфигурации объекта;
- 2) для каждой конфигурации решается "задача о многокритериальной упаковке" на основе сформированного списка ключевых задач;

3) осуществляется модификация клиентом построенных конфигураций (если это необходимо) и определяется лучшая конфигурация или несколько потенциально лучших конфигураций.

К *недостаткам* данного подхода можно отнести:

- необходимость привлечения к разработке СППР эксперта в ПрО, который хорошо ориентируется в методах принятия решений;
- ограниченность решаемых задач: фактически все множество альтернатив разбивается на два класса — заслуживающие и не заслуживающие внимания ЛПР.

2.3. Использование специализированной экспертной системы, результатом работы которой должна являться готовая СППР

Примером этого подхода может служить *алгоритм адаптивного конфигурирования подсистем СППР на основе экспертной системы и мнений группы экспертов*. В работе [12] предложен алгоритм синтеза адаптивной системы поддержки принятия решений (АСППР) с использованием функции полезности (ценности) и мнений экспертов для оптимизации ее функционирования. Предложенная архитектура многозадачной АСППР обеспечивает оперативное конфигурирование модулей системы при решении задач, относящихся к различным ПрО.

Ядро СППР составляет *экспертная система*, обеспечивающая построение модулей СППР на основе экспертных процедур и предпочтений ЛПР в целях эффективного анализа слабоструктурированных проблем. Адаптивность системы обеспечивается за счет экспертных процедур, позволяющих построение систем и анализ проблем в условиях неопределенности и нечеткости исходной информации.

Схема действия алгоритма заключается в следующем. Устанавливается тип задачи на основе разработанной классификации задач. Используя мнения экспертов, определяется объем и вид информации о задаче в целях дальнейшего выбора необходимых методов принятия решений и модулей АСППР. Находится решение по данной задаче, и система переходит к решению новой проблемы.

Каждая из задач N характеризуется определенным набором информации, которую описывает система критериев K . На основе полученной системы критериев определяются необходимые этапы поддержки принятия решений с помощью экспертов или экспертной системы. Каждый из этапов соответствует конкретному программному модулю, который обеспечен необходимым математическим аппаратом. Набор программных модулей определяет последовательность функционирования СППР. Адаптивность СППР дает возможность получения

адаптивного решения любой задачи на основе предложенного алгоритма.

Математическая модель конфигурирования структуры АСППР, сформированная на основе предпочтений ЛПР и мнений экспертов, позволяет определять необходимый минимальный набор подсистем при решении задач различной структуры в условиях неопределенности и слабой структурированности исходной информации.

Функционально многозадачная адаптивная СППР построена таким образом, что при изменении типа задач, относящихся к различным ПрО, система перенастраивается на конкретную проблемную ситуацию. Возможность организации базы данных и базы знаний больших объемов позволяет использовать накопленную информацию при решении различных задач, с учетом методов генерации решений, формирования системы критериев и выбора решений, в зависимости от ПрО.

К *недостаткам* данного метода можно отнести большую трудоемкость создания системы и необходимость участия в конфигурировании СППР эксперта в ПрО, хорошо ориентирующегося в методах принятия решений.

2.4. Онтологический подход к разработке СППР

Онтологии — модели, являющиеся компьютерной формой представления знаний о предметных областях в виде семантических информационно-логических сетей взаимосвязанных объектов, где в качестве главных элементов выступают *понятия* (или классы объектов, наблюдаемых в ПрО) с их *свойствами и отношениями* между объектами. Онтологии выполняют интегрирующую функцию, обеспечивая общий понятийный базис в процессах принятия решений и единую платформу для объединения разнообразных информационных систем в ПрО [13].

В качестве интегрирующей основы для множества информационных систем в [13] предлагается использовать онтологии, которые содержат концептуальную информацию и в явном и неявном виде уже являются основой создания и функционирования частных СППР. Идентификацию систем в онтологиях предлагается осуществлять с учетом персональных знаний ЛПР и конкретизировать в объектных денотативных моделях.

Идентификацию систем предлагается осуществлять с помощью их анализа в целях формализации в виде семантических информационно-логических моделей. Такие модели отображают данные в виде совокупности объектов и связей между ними. Для их построения предлагается использовать онтологический анализ ПрО, в соответствии с которым строятся два типа моделей: онтологии, описывающие понятийную структуру исследуемых ПрО (классы объектов и характеризующие классы атри-

буты), и денотативные объектные модели ПрО, описывающие конкретные системы.

Схематично в построении онтологии можно выделить следующие этапы:

- 1) изучение системы;
- 2) выявление классов объектов и построение их списка;
- 3) фиксация у классов объектов их атрибутов-свойств;
- 4) обнаружение родовидовых связей;
- 5) выявление у классов объектов способностей образовывать различные связи (атрибуты-валентности).

При составлении модели системы эти этапы могут повторяться.

На языке разработанных онтологий описываются объектные денотативные модели конкретных систем.

Подход на основе онтологий обладает существенным *недостатком*: создание онтологии для конкретной ПрО — сложная комплексная задача. Для ее решения необходимо привлекать высококвалифицированных специалистов. Создание и верификация такой онтологии может занять много времени, поэтому в условиях требования быстрого создания СППР такой подход неприемлем.

2.5. Автоматизация создания СППР на основе лингвистических моделей

Для решения трудноформализуемых задач используются появившиеся сравнительно недавно интегрированные системы поддержки принятия решений (ИСППР). Их отличает наличие в составе ИСППР модуля преобразования данных, накопленных в процессе функционирования исследуемой системы, в основные, наиболее устойчивые закономерности ее поведения, представляемые в виде приближенных математических моделей. Такие модели составляют основу базы знаний ИСППР.

В условиях разнотипности информации, в среде которой функционируют исследуемая система или процесс, *лингвистические модели* являются наиболее рациональной формой представления приближенных моделей, составляющих базу знаний ИСППР [14]. Использование лингвистических моделей упрощает согласование типов данных и позволяет настраивать алгоритм извлечения знаний из имеющихся данных для достижения наилучшего результата.

В работе [14] предложен *метод эволюционного синтеза лингвистических моделей (ЭСМ)* в условиях ограниченной длины предыстории функционирования системы, относительно которой принимаются решения, отвечающей условиям рациональности поддерживаемых решений. Он заключается в итерационном формировании модели путем анализа генерируемых гипотез, оценки их истинности с помощью информационной меры и принятия решения о завершении процедуры с помощью информационного тестового критерия. ЭСМ позво-

ляет автоматизировать процесс создания базы знаний для конкретной ПрО.

Метод априорной оценки рациональности решений [14], получаемых с помощью ИСППР на основе лингвистических моделей, заключается в использовании показателей размытости оценок основных аргументов функции полезности, используемой для оценки рациональности на начальной стадии работы ИСППР, а при "накоплении опыта" работы ИСППР — в использовании композиционного правила вывода значения функции принадлежности вектора критериев, отражающего цель решения, по информации "задача — решение — критерий", отражающей "накопленный опыт". Данный метод позволяет дать предварительную оценку целесообразности практической реализации решения, выработанного решающей системой. При получении низкой степени рациональности ЛПР может уточнить некоторые исходные данные, ограничения или по-иному поставить задачу.

Методы ЭСМ и оценки рациональности решений на основе лингвистической модели могут использоваться для решения задач прогнозирования, управления, диагностики в различных областях деятельности. *Недостатком* описанного метода является слабая устойчивость результатов относительно исходных данных, задаваемых экспертом.

Заключение

Приведенный обзор подтверждает рост интереса к автоматизации процесса создания СППР. Это диктуется современными требованиями по сокращению времени на разработку и внедрение СППР, сокращению затрат на их проектирование и на сопровождение эксплуатируемых приложений при их адаптации к изменениям в предметной области.

Каждый из существующих на сегодняшний день подходов к автоматизации проектирования систем поддержки принятия управленческих решений имеет как достоинства, так и недостатки, наиболее распространенными из которых являются большая трудоемкость создания системы и необходимость привлечения эксперта в ПрО, хорошо разбирающегося в методах поддержки принятия решений.

Для дальнейшего развития методов автоматизации проектирования СППР наиболее перспективным представляется подход, связанный с использованием конструктора для создания СППР, предназначенных для решения задач выявления лучшего решения или построения ранжировки решений.

Следует отметить, что одной из современных тенденций в области компьютерных технологий является интеллектуализация информационных сис-

тем (применение генетических алгоритмов, нейронных сетей, агентных технологий и т. д.). Агентные технологии, используемые при создании прикладных систем [15], находят применение и при создании систем поддержки принятия решений. [16, 17], что подтверждает актуальность и активное развитие СППР.

Список литературы

1. **Крупина Е. Е.** Информационные системы поддержки принятия управленческих решений // Конспект лекций. Самара: ПГУТИ, 2011. 100 с.
2. **Тигоренко Г. А.** Информационные технологии управления: учеб. пособие для вузов, 2-е изд., доп. М.: ЮНИТИ-ДАНА, 2003.
3. **Turban E.** Decision support and expert systems: management support systems. N. J.: Prentice Hall, 1995. 887 p.
4. **Семенова И. И.** К вопросу разработки инструментальных средств для типового проектирования систем поддержки принятия решений // Современные наукоемкие технологии. 2010. № 4. С. 94, 95. URL: http://www.rae.ru/snt/?section = content&op = show_article&article_id = 6136
5. **Contono G., Virgilio G.** A spatial decision support system for situational crime prevention in urban environment // 5th AGILE Conference on Geographic Information Science, Palma (Balearic Islands, Spain), 2002.
6. **Cheng S.** Development of a fuzzy multi-criteria decision support system for municipal solid waste management. Regina: Saskatchewan, 2000.
7. **Зайцев И. А.** Информационная система поддержки принятия решений по управлению эргатическими структурами: Автореф. дис. ... д-ра техн. наук. М., 2011.
8. **Долгов С. И.** Разработка системы поддержки принятия решений для выбора методов ликвидации чрезвычайных ситуаций на потенциально опасных объектах на примере аварийных газовых фонтанов: Автореф. дис. ... д-ра техн. наук. Иркутск, 2004.
9. **Чернов А. Г.** Технология создания систем поддержки принятия решений в интегрированных САПР. URL: <http://www.mediplanirovanie.ru/atchernov/rus/docs/rus-dic.htm>
10. **Genova K., Vassilev V., Andonov F., Vassileva M., Konstantinova S.** A Multicriteria Analysis Decision Support System // International Conference on Computer Systems and Technologies — CompSysTech'2004, Rousse, Bulgaria.
11. **Ларичев О.И., Афанасьев В.М., Чугунов Н.В.** Web-конструктор для построения систем поддержки принятия решений. URL: www.raai.org/library/ainews/2003/4/chugunov.doc
12. **Черкасов А. Н.** Разработка математического и алгоритмического обеспечения адаптивных систем поддержки принятия решений в ситуационных центрах: Автореф. дис. ... канд. техн. наук. Краснодар, 2011.
13. **Ситников П. В.** Построение систем поддержки принятия решений на основе онтологий: Автореф. дис. ... канд. техн. наук. Самара, 2009.
14. **Мерцалов А. А.** Программное инструментальное средство для разработки систем поддержки принятия решений на основе лингвистических моделей: Автореф. дис. ... канд. техн. наук. Тула, 2003.
15. **Евгеньев Г. Б.** Многоагентная методология — новая информационная технология создания прикладных систем // Информационные технологии. 2008. № 11. С. 33—41.
16. **Кононов И. В.** Агентно-ориентированный подход к созданию системы поддержки принятия решений, предназначенной для прогнозирования развития производственных структур: Автореф. дис. ... д-ра техн. наук. Тверь, 2006.
17. **Швецов А. Н.** Модели и методы построения корпоративных интеллектуальных систем поддержки принятия решений: Автореф. дис. ... д-ра техн. наук. Санкт-Петербург, 2004.

М. А. Бульонков, канд. физ.-мат. наук, зав. лаб.,
e-mail: mike@iis.nsk.eu,

Н. Н. Филаткина, научн. сотр.,
Институт систем информатики СО
РАН, г. Новосибирск

Ситуационный анализ в системе транспортного прогнозирования МИКС-ПРОСТОР¹

Рассматривается система автоматизации научных исследований в применении к задаче прогнозирования развития опорной транспортной сети России. Описан процесс формирования, расчета и анализа прогнозируемых вариантов. Введенные в работу методы комбинирования вариантов позволяют реализовать концепцию ситуационной комнаты, предоставляющую пользовательский интерфейс для лиц, принимающих решения.

Ключевые слова: автоматизированная система научных исследований, ситуационная комната, развитие транспортной системы

Введение

Современный уровень глобализации и рост международной торговли приводят к существенному увеличению нагрузки на транспорт. Географическое положение России дает ей потенциальную возможность извлекать прибыль из участия в международных перевозках в качестве поставщика транзитных услуг. Для того чтобы реализовать этот потенциал, формирование опорной транспортной сети России должно осуществляться с учетом возможных вариантов включения ее в систему тех развивающихся международных транспортных коридоров, которые отвечают как внешним, так и внутренним экономическим интересам Российской Федерации.

Прогнозирование развития единого транспортного пространства страны включает в себя рассмотрение проблемы на многих уровнях, по сути это система иерархически связанных задач. Для решения этих задач требуется объединение методов и подходов, наработанных в разных областях научных исследований. Конечно, ключевую роль здесь играет построение экономико-математических моделей и методов их расчета, но их практическое использование невозможно без существенной информационной поддержки. В нашем случае необходимо создание автоматизированной информационно-аналитической системы с поддержкой ГИС-технологий и

технологий автоматизации процессов научных исследований.

В настоящее время развивается достаточно много программных систем, ориентированных на проблемы развития транспорта и инфраструктуры на уровне города или даже региона. Наиболее крупной такой разработкой, на наш взгляд, является программный комплекс PTV Vision® VISUM, который позволяет заниматься как оперативным, так и стратегическим планированием процесса функционирования транспортной системы в различных условиях [1]. Чтобы подчеркнуть теоретическую базу и потенциал системы, разработчики ссылаются на обширные связи компании с фундаментальными исследованиями в области методологии транспортного моделирования.

В статье [2] обсуждаются способы решения задачи моделирования развития транспорта для городской агломерации на основе методов математического моделирования транспортных потоков. Вместе с тем, есть работы, которые теоретически обосновывают понятия потенциала различных систем и транспортных систем в частности, а также решают задачи исследования использования этого потенциала, т. е. прогнозирования развития системы. Задачи исследования потенциала систем относятся к классу новых задач, решение которых должно улучшить функционирование ряда систем, в состав которых входят взаимодействующие коллективы людей и комплексы технических устройств [3].

Существуют также и работы, посвященные описанию модели развития транспортной сети на уровне страны. Так, основной особенностью модели в статье [4] является то, что у транспортной системы отсутствует единый управляющий центр и ее динамика является результатом самоорганизации системы, происходящей в результате процессов удовлетворения потребностей в перевозках товаров и развития инфраструктуры узлов сети.

Стоит отметить, что когда мы говорим о построении прогнозов на уровне всей страны, то во многих случаях удобно оперировать объектами на уровне транспортных коридоров. О важности формирования и развития системы транспортных коридоров России можно прочитать в статье [5].

В нашем подходе для задания схемы формирования опорной транспортной сети России используется группа экономико-математических моделей составления транспортно-экономических балансов (ТЭБ) [6], которые позволяют описывать и анализировать межотраслевые и межрегиональные связи в условиях конкуренции разных видов транспорта.

В такой постановке получается, что мы задаем прогноз развития параметров, описывающих услуги, которые предоставляют конкурирующие виды транспорта, и нам необходимо построить вариант

¹ Работа поддержана грантом РФФИ № 13-02-00413.

развития опорной транспортной сети, обеспечивающий рациональное взаимодействие входящих в нее видов транспорта, при условии выполнения этого прогноза. Иными словами, мы решаем задачу прогнозирования путем имитационного моделирования типа "что-если".

Для решения поставленной задачи на основе общей системы МИКС [7], разработанной совместно ИЭиОПП СО РАН, ИСИ СО РАН и Новосибирским госуниверситетом, была создана подсистема МИКС — ПРОСТОР (Прогнозирование Развития Опорной Сети Транспортной Отрасли России). Система дает возможность оперативно строить и проверять различные гипотезы развития опорной транспортной сети, согласованно корректировать исходные данные, работать в режиме интервального задания прогнозов по потенциальным объемам транспортировки грузов и разным уровням тарифов.

Постановка задачи

Используемая нами модель данных для решения задач прогнозирования развития транспортной сети [8] включает в себя следующие **сущности**:

- *продукт* — условная группа товаров для перевозок, к которой приводятся все рассматриваемые производимые и потребляемые товары;
- *тип работы* — используются пять типов работы, связанных с функционированием транспортной сети: погрузка, выгрузка, перевозка, транзит по видам транспорта, а также перевалка с одного вида транспорта на другой для допустимых продуктов и пар видов транспорта;
- *вид транспорта* — используемые в транспортной сети виды перевозки с указанием тарифов для выполнения каждого типа работ для данного вида транспорта;
- *узел (транспортный узел)* — географический пункт, для которого определены возможные операции по переработке продуктов определенного вида. Поскольку узел является оператором, то для него задаются показатели максимальной производительности и коэффициенты тарифной стоимости по выполнению каждого типа работ. Во множестве узлов выделяются узлы-производители (пункты начальной погрузки) и узлы-потребители (пункты конечной выгрузки). Кроме того, к узлу добавлена географическая привязка для возможности последующей визуализации;
- *плечо* — элементарная единица транспортной сети, условный вектор, соединяющий смежные узлы. Плечо параметризуется длиной, пропускной способностью и тарифами на перевозку по видам транспорта и продуктам. Аналогично узлам плечи дополняются последовательностью географических координат, позволяющих изображать соответствующий путь на карте.

Переменные модели:

X_{ir}^j — объем продукта вида i , принятый в узле r к погрузке на транспорт вида j ; $i \in I$, где I — множество видов груза; $r \in R$, где R — множество всех транспортных узлов; $j \in J$, где J — множество всех видов транспорта;

Y_{ir}^j — объем груза вида i , принятый в узле r к выгрузке с транспорта вида j ;

$V_{ir}^{(j\tau)}$ — объем груза вида i , перегружаемого в узле r с транспорта вида j на транспорт вида τ ;

$\bar{V}_{ir}^{(\tau j)}$ — объем груза вида i , перегружаемого в узле r с транспорта вида τ на транспорт вида j ;

Z_{ir}^j — объем груза вида i , проходящего транзитом через узел r транспортом вида j ;

W_{irs}^j — объем грузопотока продукта вида i , отправляемого из узла r в смежный узел s транспортом вида j ;

\bar{W}_{irs}^j — объем грузопотока продукта вида i , принятый из узла s в смежный узел r транспортом вида j ;

B_{ir} — верхняя граница на объем продукции вида i , предназначенный для вывоза из узла r ;

A_{ir} — нижняя граница на объем продукции вида i , ввозимой в узел r всеми видами транспорта;

$l_{(rs)}^j$ — расстояние между смежными узлами для сети транспорта вида j ;

ϕ_{irs}^j и $\bar{\phi}_{irs}^j$ — удельные (в пересчете на единицу грузооборота) затраты на перевозку продукции между парами смежных транспортных узлов в прямом и обратном сообщении соответственно.

Условия, задаваемые моделью:

1. Ограничения на объем отдельного вида произведенной в узле продукции и принятой разными видами транспорта к погрузке для вывоза за пределы узла:

$$\sum_j X_{ir}^j \leq B_{ir}, \quad i \in I, r \in R, j \in J,$$

2. Ограничения на объем отдельного вида продукции, принятой в узле от разных видов транспорта к выгрузке для потребления внутри узла:

$$\sum_j Y_{ir}^j \geq A_{ir}, \quad i \in I, r \in R, j \in J,$$

3. Определение объема отдельного вида продукции, проходящей транзитным грузопотоком через узел отдельным видом транспорта:

$$Z_{ir}^j = \sum_s \bar{W}_{irs}^j - Y_{ir}^j - \sum_\tau V_{ir}^{(j\tau)},$$

$$s, r \in R, r \neq s; i \in I; j, \tau \in J, \tau \neq j.$$

4. Определение объемов продукции, перегружаемой в узле с одного вида транспорта на другие:

$$\sum_\tau V_{ir}^{(j\tau)} = X_{ir}^j + \sum_s \bar{W}_{irs}^j + \sum_\tau \bar{V}_{ir}^{(j\tau)} - Y_{ir}^j - \sum_s W_{irs}^j,$$

$$r, s \in R_1, r \neq s; j, \tau \in J, \tau \neq j; i \in I.$$

Целевая функция модели:

представляет собой минимум суммарных приведенных затрат на развитие опорной транспортной системы, который обеспечивает фиксированные по узлам зарождения и погашения объемы перевозок. Структура минимизируемого функционала является суммой следующих групп слагаемых, каждая из которых отражает формирование издержек, связанных с определенными операциями в сфере транспортного процесса:

- с переработкой грузов непосредственно в транспортных узлах;
- с движущими операциями, т. е. непосредственно с перевозкой грузов.

Таким образом,

$$\begin{aligned} & \sum_{i,r,j} c_i^j X_{ir}^j + \sum_{i,r,j} \bar{c}_i^j Y_{ir}^j + \sum_{i,r,j} c^j Z_{ir}^j + \sum_{i,r,(j\tau)} c_{ir}^{(j\tau)} V_{ir}^{(j\tau)} + \\ & + \sum_{i,r,(j\tau)} \bar{c}_{ir}^{(j\tau)} \bar{V}_{ir}^{(j\tau)} + \sum_{j,i,(rs)} \phi_{(rs)}^j l_{(rs)}^j W_{irs}^j + \\ & + \sum_{j,i,(sr)} \bar{\phi}_{(sr)}^j l_{(rs)}^j \bar{W}_{irs}^j \rightarrow \min, \end{aligned}$$

где коэффициенты при переменных — удельные (в пересчете на единицу груза) затраты на погрузку, выгрузку, обслуживание транзитных грузопотоков и перегрузку с одного вида транспорта на другой соответственно.

Таким образом, мы имеем задачу оптимальной перевозки от узлов-производителей к узлам-потребителям. Решатель для этой оптимизационной задачи был доступен и использовался в течение многих лет [9]. Необходимо было обеспечить интерактивную визуальную и геоинформационную поддержку всего исследовательского процесса. В данном случае в работе исследователя мы выделили следующие задачи:

1. *Редактирование транспортной сети.* Для этого необходимо задать список перевозимых продуктов, используемые виды транспорта, обозначить актуальные транспортные узлы и их параметры, определяющие производство продуктов и их обработку внутри узла, а также плечи, определяющие направление, пропускную способность и тарифы перевозки.

2. *Моделирование.* Когда текущее состояние сети уже задано, можно обратиться собственно к задачам прогнозирования, отвечающим на вопросы "А что будет, если тарифы на данном участке пути уменьшатся вдвое?" или "Как повлияет на распределение перевозок увеличение производства этого товара в данном узле?" или "Насколько должны измениться тарифы на данном отрезке пути, чтобы он стал участвовать в перевозке выбранного продукта?" Для этого необходимо предоставить исследователю возможность формулировать возможные изменения параметров сети, сохранять наборы этих параметров, а также иметь возможность решать оптимизационную задачу с разными наборами изменений и

проводить сравнительный анализ результатов. Чтобы оперировать с наборами изменений, мы ввели понятие серии вариантов, которое позволяет задавать не только новые единичные значения для выбранного параметра, но и указывать диапазон, в котором должно изменяться значение и число шагов изменения. Это дает возможность исследователю одновременно просчитать целую последовательность задач и получить соответствующую им последовательность решений, просмотреть их, провести сравнительный анализ, выбрать нужный из вариантов, а также на его основе создать новые серии вариантов, включающие в себя дополнительные изменения.

3. *Ситуационная комната.* Специфика задачи не позволяет однозначным образом интерпретировать результаты решения, предоставляемые решателем. Решением является описание сети перевозок между узлами-производителями и узлами-получателями, а интерпретация результатов, выводы об их свойствах и перспективности остаются за исследователем. Более того, результатом исследования в данном случае является правильная постановка вопроса прогнозирования, которую исследователь может сформулировать в процессе моделирования различных вариантов решения. Чтобы иметь возможность поделиться интересными вариантами постановки задачи прогнозирования с другими специалистами, не являющимися экспертами в предметной области, было предложено создать ситуационную комнату.

Далее мы рассмотрим эти задачи более подробно.

Архитектура системы

Система МИКС-ПРОСТОР является одним из приложений системы МИКС, которая была разработана в целях обеспечения визуальной и геоинформационной поддержки экономико-математического моделирования и анализа различных систем, таких как система взаимодействия политико-административных субъектов, система анализа статистических данных регионов, система прогнозирования развития транспортной сети, система визуализации работы лесного хозяйства.

В данный момент активно развиваются два приложения системы МИКС: МИКС-РЕГИОН и МИКС-ПРОСТОР. В общем случае каждое приложение имеет свою специфическую терминологию, систему связей между информационными объектами, вводит дополнительные объекты, использует собственный программный аппарат, может создавать дополнительные хранилища промежуточной информации. Но при этом визуальные методы поддержки работы с разными экономико-математическими моделями имеют значительную общую реализационную часть и пользовательскую логику. Таким образом, разные приложения имеют в своем составе одинаковые визуальные компоненты, такие

как компонента с картографической информацией, табличные отображения данных или деловая графика.

В задаче анализа транспортных потоков во многих случаях наиболее естественно выбирать узлы и плечи для задания прогноза, используя картографическое представление. Вместе с тем, в ряде случаев удобно вносить коррективы к прогнозу из компоненты, отвечающей за общее редактирование транспортной сети. Таким образом, поскольку система предоставляет несколько разных способов редактирования информации о прогнозе, более того, за это редактирование отвечают различные компоненты визуализации, то возникает необходимость синхронизации информации, поступающей от различных компонент. В целях реализации взаимодействия различных компонент визуализации мы изменили логическую структуру системы МИКС, выделив в ней общее ядро (*MIX Engine*), отвечающее за доступ к данным, а также генерацию и распространение событий. Это ядро имеет предметно-независимую часть, а также включает в себя ряд специализированных функциональных элементов, ориентированных на работу с конкретными предметными областями (*Transport Gear*, *Region Gear*). Таким образом, все знания о специализированных хранилищах данных распределены между соответствующими функциональными элементами.

Масштабируемость входила в изначальную постановку задачи при планировании архитектуры системы МИКС, причем масштабируемость подразумевается в двух различных измерениях. С одной стороны, необходимо иметь возможность добавлять к рассмотрению новые модели, связанные с новыми предметными областями, с другой стороны — наращивать визуальные возможности каждой специфической подсистемы.

Первое измерение обеспечивается строением ядра системы и возможностью добавления новых

функциональных элементов, на рис. 1 эта часть изображена в центре в виде взаимодействующих шестеренок. Визуальные компоненты изображены на этом рисунке прямоугольниками. Для работы с компонентами используется технология плагин. Полный набор компонент, используемых системой, задается в конфигурационном файле. При работе системы компоненты загружаются динамически, и их список можно менять в зависимости от типа рассматриваемых задач.

Поскольку каждая визуальная компонента имеет доступ ко всем подсистемам ядра, то это дает возможность в перспективе строить компоненты, оперирующие одновременно информацией из разных предметных областей и работающие на стыке двух разных задач. Например, решать транспортную задачу для лесного хозяйства, опираясь на конкретные данные, доступные в результате анализа структуры лесного фонда.

Различные комбинации визуальных компонент составляют приложения системы МИКС. В данной статье мы рассматриваем конфигурацию МИКС-ПРОСТОР, работающую с транспортными моделями. Обмен событиями и обращение к данным в приложении МИКС-ПРОСТОР происходит через часть ядра, называемую *Transport Gear* и отвечающую за организацию транспортной модели. В частности, *Transport Gear* обеспечивает построение промежуточных файлов, поступающих на вход оптимизационного пакета (с помощью которого решается задача нахождения оптимальной транспортной сети), отвечает за запуск этого пакета и преобразование результатов в вид, необходимый компонентам для визуализации. Кроме того, она обеспечивает доступ к содержимому транспортной сети, а также интерфейс для задания вариантов с соответствующим набором событий.

Сценарии работы

В начальной реализации системы МИКС-ПРОСТОР мы ориентировались на три группы пользователей:

- администратор транспортной сети — человек задающий основополагающие части транспортной сети, в первую очередь ее геометрию, а также общий список используемых продуктов и рассматриваемых видов транспорта;
- эксперт в задачах моделирования и прогнозирования — человек, который на основе заданной экономико-математической модели транспортной задачи строит некоторые варианты развития за счет изменения количественных параметров сети;
- лицо, принимающее решения (ЛПР) — человек, который, пользу-

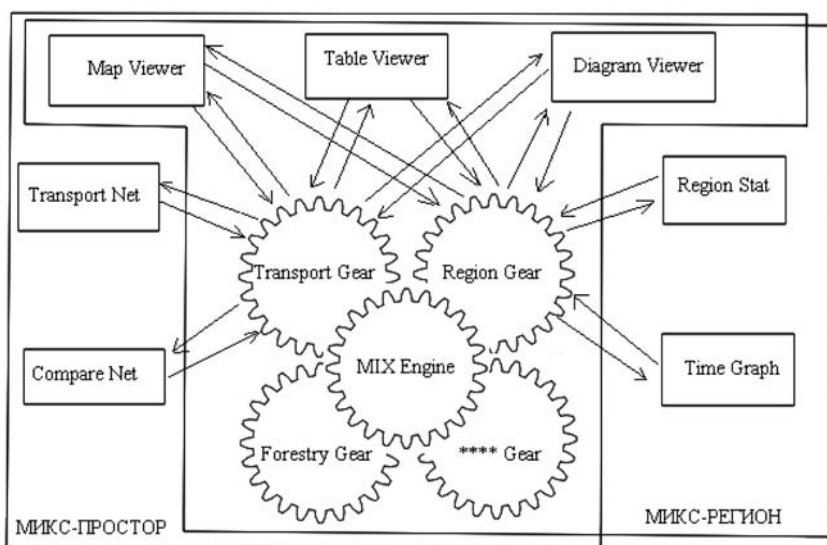


Рис. 1. Структура МИКС

ьясь некоторым набором глобальных ситуаций, сможет сделать общие выводы о перспективах развития, а также провести эксперименты с разными сценариями развития, не вникая в количественные параметры задачи.

Однако на практике оказалось, что при рассмотрении задач прогнозирования не имеет смысла разделять первые две роли, поскольку в процессе моделирования возникает потребность менять все части сети, включая геометрию и набор перевозимых грузов, а также добавлять к рассмотрению новые виды транспорта. Следовательно, мы разработали архитектуру, обеспечивающую специализированные интерфейсы для двух видов пользователей. Первым типом является эксперт в области экономико-математических моделей для транспортной задачи, в функции которого входит создание и редактирование транспортной сети, а также создание вариантов развития этой сети и набора интересных ситуаций развития. Этот набор ситуаций является результатом работы эксперта и используется в качестве входных данных для второго типа пользователей — экономистов, не являющихся экспертами в решении транспортных задач, а также лиц, ответственных за принятие решений в области развития транспортной системы. Для них интерфейс должен максимально скрывать математическую составляющую задачи и позволять на макроуровне задавать сценарии развития событий, используя различные сочетания ситуаций. И мы предоставляем такой интерфейс в виде информационно-аналитического комплекса поддержки решений — ситуационной комнаты [10].

Редактирование сети

Чтобы обеспечить информационную поддержку для работы эксперта с транспортными задачами, надо в первую очередь создать среду, которая бы предоставляла удобный интерфейс для создания и редактирования параметров экономико-математической модели транспортной системы.

Пример типичной панели редактирования приведен на рис. 2. Это интерфейс для редактирования значений производства и потребления для каждого продукта, положительное число означает произ-

Настройка производства и потребления ресурса	
<input type="checkbox"/> Показать все города	
Узел	Значение
Ляньюньган	-20
Новосибирск	-5
Франкфурт	+30
Ченджу	-5
Итого	0

Рис. 2. Редактирование производства и потребления ресурсов

водство, отрицательное — потребление продукции в данном узле. Последняя строчка позволяет при редактировании следить за соблюдением баланса производства-потребления, проверять таким образом корректность введенных значений. Аналогичным образом можно редактировать остальные динамические параметры транспортной сети.

На данный момент система позволяет создавать и редактировать только одну транспортную сеть, но структура хранения данных дает нам возможность легко перейти к рассмотрению нескольких сетей одновременно.

Моделирование сети

Наиболее гибкой частью транспортной сети мы считаем коэффициенты, связанные с тарифами и пропускной способностью, соответственно в первую очередь была реализована вариативность в отношении этих коэффициентов. Мы ввели понятие *вариатора*, символизирующего единицу изменения. Вариатор состоит из трех значений: *начальное значение, конечное значение, число шагов*. Первые два определяют диапазон изменения параметра, т. е. задают начальное и конечное значения для варьирования, а третий параметр специфицирует число промежуточных значений, которое необходимо вычислить. Значением вариатора является последовательность значений параметра, имеющая длину, равную числу шагов и представляющую изменение параметра от начального значения до конечного. В приложениях может быть актуально рассмотрение как аддитивного, так и мультипликативного процесса изменения параметра.

В качестве начального варианта мы остановились на первом процессе, т. е. параметр принимает значения, равномерно расположенные по диапазону, заданному вариатором. Рассматриваются вариаторы трех типов: вариатор для тарифов на выполнения работ внутри узла, параметризующийся видом транспорта; вариатор для тарифов на выполнение перевозки вдоль выбранного плеча транспортной сети; вариатор для значения пропускной способности для плеча. Вариатор задается диалогом, представленным на рис. 3. В данном примере задан вариатор, который говорит, что необходимо изменить пропускную способность на обобщенном железнодорожном транспортном плече между Урумчи и Ташкентом. Значениями данного вариатора будет последовательность 50, 75, 100.

Для описания задачи прогнозирования исследователь создает так называемую серию вариантов

Урумчи - Ташкент (Жд)	
Пропускная способность	50 - 100
	3

Рис. 3. Редактирование вариатора

сети. Серия вариантов — это набор вариаторов, в котором каждый параметр сети может использоваться не более одного раза.

Если значением одного вариатора является последовательность значений параметра, то значением серии вариантов являются многомерная матрица из всевозможных сочетаний значений вариаторов. Таким образом, серия вариантов определяет последовательность оптимизационных задач, каждая из которых использует текущие значения вариаторов для выбранного набора параметров вместо имеющихся умолчательных значений. Исследователь имеет возможность просчитать выбранную серию задач, в результате чего система генерирует пакет оптимизационных задач, просчитывает их с помощью решателя и получает соответствующую последовательность решений. Эта последовательность решений и предоставляется пользователю в качестве решения для варианта.

В реализованной модели вычисления все вариаторы независимы, т. е. множество решений получается перебором всех комбинаций значений. Однако в некоторых случаях возникает желание изменять параметры некоторых элементов транспортной сети синхронно. Например, если мы хотим провести эксперимент, связанный с изменением тарифа перевозки по Транссибирской магистрали, рассматривая ее как единое целое, то мы должны обеспечить согласованное изменение параметров плеч, представляющих ее в нашей модели. Для задания такого изменения мы должны иметь возможность создать виртуальный объект, агрегирующий однотипные объекты модели, и варьировать параметры для всего этого объекта.

Поскольку на данный момент основным приложением системы являются задачи глобального прогнозирования, то следует учитывать, что в описании прогнозируемого варианта развития сети, как правило, участвует не одна точка этой сети, а некоторая цепочка связанных участков. Так например, для описания ситуации со снижением тарифов на Северном морском пути нам необходимо выделить все участки, задающие этот путь, и указать согласованные изменения коэффициентов. Наиболее удобным и наглядным в этом случае является использование картографического представления. Пользователь может непосредственно на карте выбрать интересующие его участки дороги (плеча) и узлы транспортной сети и задать параметры для выбранных узлов, при этом на карте подсвечиваются элементы сети, участвующие в текущем варианте.

Для нахождения интересных вариантов прогнозирования исследователю нужен широкий спектр операций над сериями вариантов. Система предоставляет возможность группировки серий и создания иерархии, копирования серий вариантов и их содержимого, добавления содержимого одной серии

к другой. Формат хранения позволяет обмениваться сериями вариантов в случае групповой работы.

Визуальные возможности для изображения результатов решения

Для отображения решения мы используем два визуальных представления: картографическое и табличное. Картографическое представление является наиболее естественным, поскольку мы имеем дело с реальными географическими объектами [11]. Но вместе с тем, для сравнительного анализа числовой информации оказываются более уместными средства бизнес-графики.

В начальный момент на карте и в таблице отображается решение, объединяющее продукты всех типов. Это позволяет оценить общую картину: посмотреть загруженность сети, выявить "узкие" места. Отфильтровав результаты по типу продукта, можно увидеть конкретные числовые значения для перевозок этого продукта в рамках общей постановки задачи. Эти значения доступны пользователю в виде таблицы, в которой собраны все плечи, которые по результатам решения целесообразно использовать для перевозок данного продукта: указаны начальный и конечный пункт плеча, вид транспорта, объем перевозки и процент загруженности участка. Таблица предоставляет возможность отсортировать содержимое по любому из параметров, а также выделить цветом те участки, на которых загруженность превышает заданный уровень. Кроме того, объем перевозки можно посмотреть, выделив соответствующий транспортный узел на карте.

Для оценки перспективности полученного решения или необходимости более радикального изменения начальных параметров часто используется изучение степени загруженности дорожных участков. Система предоставляет специальный режим отображения, позволяющий выделять загруженность цветом. При изображении транспортной сети на карте цвет участка меняется градиентно от синего до красного, где синий соответствует нулевой загруженности дороги, а красный — 100 %. Это позволяет быстро и наглядно отобразить проблемные места. Так например, на рис. 4 (см. четвертую сторону обложки) видно, что в сложившейся сети перевозок "узким" местом для дальнейшего увеличения мощностей является цепочка связанных участков от Екатеринбурга до Новосибирска.

Второй режим отображения картографической компоненты меняет ширину участка дорог пропорционально объемам перевозок выбранного продукта. Таким образом, тот же вариант решения будет отображен, как показано на рис. 5 (см. четвертую сторону обложки), наглядно демонстрирующим, что основной объем перевозок приходится на морские пути через Суэцкий канал, а Трансказахстанская магистраль используется совсем незначительно.

Информативность отображения может быть значительно улучшена, если отобразить оба этих показателя одновременно, что для тех же входных данных дает изображение, приведенное на рис. 6 (см. четвертую сторону обложки). Этот рисунок показывает, что при дальнейшем увеличении объемов перевозки или включении в рассмотрение дополнительных продуктов перевозки Транссибирская магистраль перестанет справляться с нагрузкой (на ней есть участки пути со 100 %-ной загруженностью, изображаемые красным цветом), и Трансказахстанская магистраль является реальным конкурентом в данной постановке задачи, причем видно, что проблемным является не вся Транссибирская магистраль, а несколько ее отрезков.

Ситуационная комната

При работе над глобальными прогнозами развития у эксперта возникает желание понизить уровень детализации. Первый раз мы сделали это на уровне логической модели транспортной сети, выделив обобщенные транспортные узлы и плечи, которые интересны с точки зрения международных транспортных потоков. Далее, после проведения экспериментов с этой моделью эксперт создает варианты изменений, которые определяют некоторую "элементарную ситуацию" с точки зрения глобальных изменений. Например, эксперт выделяет все участки Транссибирской магистрали в один вариант и ставит согласованные изменения на всех этих участках, задавая таким образом ситуацию, которая характеризуется общим снижением тарифов на Транссибирской магистрали и позволяет использовать эту ситуацию как единое целое, сочетая его в прогнозах с изменениями в других частях мировой транспортной системы. Таким образом, эксперт может создать вариант сети и сказать, что этот вариант является элементарной ситуацией. Задав набор элементарных ситуаций, мы получим ситуационную комнату. Такая комната позволит оперировать изменениями на более глобальном логическом уровне, например, рассматривать транспортные коридоры как единое целое. При этом различные сочетания ситуаций позволяют оценить

структуру перевозок при синхронных согласованных изменениях во всех составляющих транспортного коридора.

Для создания ситуационной комнаты эксперт подбирает интересные элементарные ситуации и задает к ним подробное словесное описание, характеризующее экономический смысл количественных изменений скрытых внутри соответствующего варианта транспортной сети. Ситуационная комната позволяет эксперту задавать всевозможные комбинации ситуаций, задавая таким образом некоторый глобальный прогноз развития, и получать решение транспортной задачи при условии исполнения всех элементарных прогнозов, описывающих изменения в различных частях общей международной транспортной сети.

В задачу эксперта входит проработать и подобрать согласованный набор ситуаций, чтобы соответствующие прогнозы были актуальны и интересны. Под "согласованностью" здесь понимается то, что множество параметров, на которых основываются ситуации, не пересекаются и, следовательно, каждый параметр задан однозначно. Если же такой конфликт все же возникает, то по умолчанию он разрешается в пользу ситуации, которая появляется в списке позже.

Некоторые ситуации могут содержательно отличаться тем, что задают разные значения одного и того же параметра. В этом случае мы позволяем оформить эти ситуации как альтернативные так, чтобы пользователь мог выбрать не более одной из них. Например, мы можем рассмотреть три варианта изменения тарифов провоза по Северному морскому пути: сильное снижение — 50 %, среднее снижение — 30 % и небольшое снижение — 10 %. Неявно мы имеем и четвертый вариант, оставляющий уровень тарифов неизменным. В этом случае разумно будет рассматривать эти ситуации не как независимые "элементарные" ситуации, а как альтернативы одного прогноза. Для отображения таких альтернатив мы расширили язык описания ситуационной комнаты и способы визуального представления. Пример таких альтернативных ситуаций можно увидеть на рис. 7.

2. Ситуация с Северным морским путем: льготный тарифный режим без ледовых сборов.

В настоящее время здесь практически нет сквозного плавания. В 1989 г. перевозки составили около 6 млн. т. в основном на западном плече Мурманск - Дудинка (Норильск). Поставлена задача возродить этот вариант международных транзитных перевозок, воссоздать базы береговой поддержки. По сравнению с вариантом через Суэцкий канал, этот маршрут на 7 тыс. км короче, но, естественно труднее (льды даже в летнюю навигацию). П_Р_И_М_Е_Ч_А_Н_И_Е: СМП потенциально конкурентоспособен, но требует серьезной модернизации. Может быть доходы от транзита - не главная цель?

Понижение провозных тарифов на 50%

Понижение провозных тарифов на 30%

Понижение провозных тарифов на 10%

Рис. 7. Альтернативные варианты ситуаций

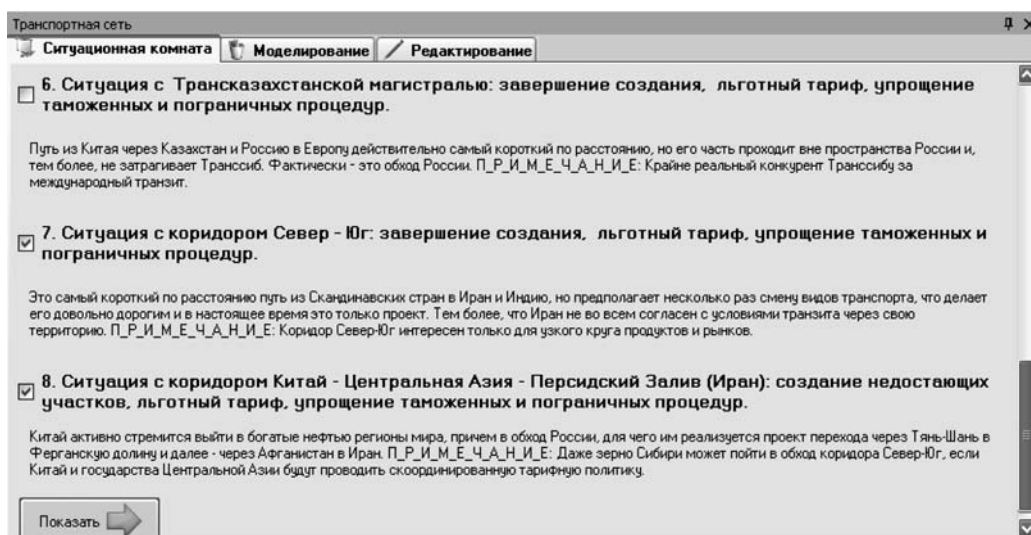


Рис. 8. Ситуационная комната

Диалог ситуационной комнаты (рис. 8) строится динамически по файлу описания, который создается экспертом в процессе моделирования. Динамическая генерация дает возможность легко расширять ситуационную комнату и контролировать внесенные изменения. Таким образом, можно говорить о гибкости и настраиваемости разработанных средств: система может быть достаточно просто перенастроена на любую другую (в том числе и прогнозируемую, а не реально существующую) транспортную сеть и другой набор базовых ситуаций.

Заключение

Мы рассмотрели систему автоматизации научных исследований в области прогнозирования развития опорной транспортной сети России. Существенным для успешного использования этой системы является использование адаптивного интерфейса, ориентированного на конкретную категорию пользователей. В частности, концепция ситуационной комнаты повышает понятийный уровень и, как следствие, позволяет проводить содержательные эксперименты лицам, не обладающим экспертными знаниями в области экономико-математического моделирования.

Дальнейшее развитие системы нам видится, во-первых, в реализации методов сравнения различных вариантов прогнозов на основе использования общего подхода, принятого в системе МИКС-ПРОСТОР, при котором существенную роль в процессе анализа играют директивы эксперта. Во-вторых, в расширении контекста, учитывающего не только развитие транспортной системы, но и тенденции соответствующего ему экономического развития регионов России и их взаимодействия.

Список литературы

1. Швецов В. Л., Прохоров А. В., Ильин И. В. Транспортные модели в системе государственного управления // Научно-технические ведомости СПбГПУ. 2009. № 5 (85). С. 20.
2. Швецов В. И. Алгоритмы распределения транспортных потоков // Автоматика и Телемеханика. 2009. № 10. С. 148—157.
3. Гейда А. С., Лысенко И. В. Автоматизация решения задач исследования потенциала систем и эффективности их функционирования // Труды СПИИРАН. 2012. Вып. 22. С. 260—281.
4. Агапова Г. И., Гавдаева А. В., Степанцов М. Е. Моделирование динамики развития железнодорожных сетей // Препринты ИПМ им. М. В. Келдыша. 2011. № 73. 12 с. URL: <http://library.keldysh.ru/preprint.asp?id = 2011-73>
5. Сергеев В. И., Гончаренко С. С., Прокофьева Т. А. Формирование международных транспортных коридоров в регионах Европейского Севера, Сибири и Дальнего Востока // Логистика сегодня. 2011. № 5. С. 5—21.
6. Воробьева В. В., Малов В. Ю., Марусин В. В., Радченко В. В. Прогнозирование формирования опорной транспортной сети: инструментарий вариантных расчетов // Сибирь и Дальний Восток в долгосрочной стратегии развития интегрированной транспортной инфраструктуры Евразии / Под ред. С. Н. Васильева и др. Иркутск—Москва—Новосибирск: ИрГУПС, 2011. С. 516—522.
7. Бульонков М. А., Карпан В. В., Малов В. Ю., Марусин В. В., Радченко В. В. Концептуальные вопросы построения Модельно-Информационно-Картографической Системы (МИКС) // Моделирование производственных и региональных систем на основе ГИС и информационных технологий: Сб. науч. тр. / Под ред. Ю. Ш. Блама, В. В. Радченко. Новосибирск: ИЭОПП СО РАН, 2011. С. 5-28.
8. Воробьева В. В., Малов В. Ю., Радченко В. В., Поттер М. В., Серебрянников И. Е. Модель прогнозирования развития опорной транспортной сети // Моделирование производственных и региональных систем на основе ГИС и информационных технологий: сб. науч. тр. / под ред. Ю. Ш. Блама, В. В. Радченко. — Новосибирск: ИЭОПП СО РАН, 2011. С. 68—96.
9. Забияко Г. И. Пакет программ целочисленного линейного программирования // Дискретный анализ и исследование операций. 1999. Сер. 2. Т. 6, № 2. С. 32—41.
10. Райков А. Н. Ситуационная комната для поддержки корпоративных решений // Открытые системы. 1999. № 7—8. С. 56—66.
11. Турлапов В. Е. Геоинформационные системы в экономике: учебно-метод. пособие // Нижний Новгород: НФ ГУ-ВШЭ, 2007.

УДК 004.932

С. С. Садиков, д-р техн. наук, проф.,

e-mail: sadykovss@yandex.ru

Ю. А. Буланова, аспирант,

e-mail: yuliyabulanova@yandex.ru

Муромский институт (филиал)

ГОУ ВПО "Владимирский государственный университет имени Александра Григорьевича и Николая Григорьевича Столетовых", г. Муром

Е. А. Захарова, врач-рентгенолог 1 кат.,

НУЗ Отделенческая больница

на станции Муром, ОАО "РЖД"

Алгоритм выделения области кисты на малококонтрастных маммограммах

Дан анализ известных алгоритмов предварительной обработки применительно к малококонтрастным маммографическим снимкам. Разработан алгоритм гистограммного преобразования маммограмм, учитывающий средний контраст снимка. Предложен алгоритм выделения области новообразования типа "Киста молочной железы" на маммограмме.

Ключевые слова: маммограмма, киста, предварительная обработка, гистограммные преобразования, средний контраст, текстурная сегментация, выделение области новообразования, выделение контуров

Введение

Заболевания молочной железы находятся на первом месте в списке заболеваний женщины [1, 2]. Одним из самых распространенных доброкачественных заболеваний является киста молочной железы [3]. Киста молочной железы — это полостное образование, заполненное жидкостью, которое появляется в результате уничтожения просвета в молочных ходах вследствие избыточного разрастания соединительной ткани. При пальпации они напоминают мягкий виноград или очень маленькие воздушные шарики, наполненные водой. Чаще всего кисты формируются у женщин в возрасте 30—55 лет [4].

Эффективным методом диагностики новообразований является рентгенологическая маммография. На маммограмме киста дает однородную тень, по плотности сходную с железистой тканью молочной железы, округлой или овальной формы. На маммографическом снимке вокруг кисты в большинстве

случаев образуется частичный или полный ободок просветления [4, 5]. Несмотря на то, что кисты не представляют опасности для здоровья женщины, они могут доставлять дискомфорт и боль в молочной железе, а также являются признаком возникновения злокачественных заболеваний, поэтому важным и актуальным является своевременная диагностика кисты и ее лечение. В последнее время увеличилось количество случаев диагностики рака молочной железы в кисте, что также свидетельствует о важности выявления кисты, особенно потому, что на маммографическом снимке она трудно различима. Для решения этой задачи следует использовать современные методы обработки и анализа изображений, а также средства информационных технологий.

Цель настоящей статьи — исследование алгоритмов выделения области кисты на маммографическом снимке.

Алгоритм выделения области кисты на маммограмме

Основная сложность, возникающая при автоматической обработке и анализе маммографических снимков с целью выделения области кисты, связана с большой зашумленностью снимка и некоторыми особенностями их регистрации с помощью конкретных маммографов. Поэтому использовать напрямую широко известные алгоритмы выделения областей на изображениях (сегментация) не дает желаемых результатов.

Общая блок-схема предлагаемого алгоритма дана на рис. 1.

Рассмотрим особенности каждого блока алгоритма.

1. *Исходная маммограмма $f(x, y)$* имеет размер $m \times n$ пикселей. Градации яркостей точек распределены в диапазоне [0, 255].

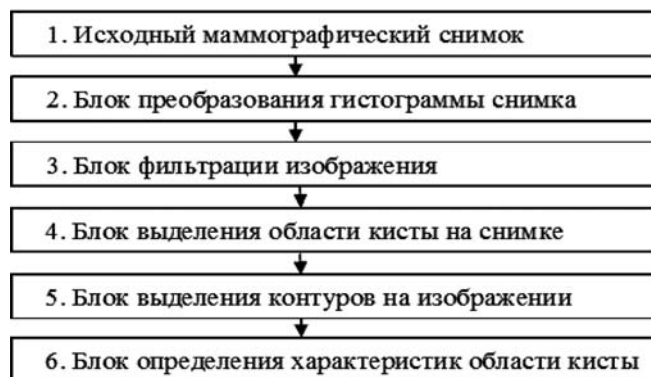


Рис. 1. Общая блок-схема алгоритма выделения области кисты на маммограмме

2. Блок преобразования гистограммы снимка включает несколько известных алгоритмов изменения значений яркости точек маммограммы, позволяющих получить выходные снимки с линейным, экспоненциальным, гиперболическим и другими распределениями гистограмм [6, 7]. Такие преобразования обычно дают возможность осуществить перераспределение числа точек в различных диапазонах яркости и тем самым визуализировать часто явно невидимые объекты (области) на снимке.

Исследования показали, что известные алгоритмы гистограммного преобразования плохо работают на малоконтрастных снимках, каковыми являются маммограммы, поскольку они в основном оперируют максимальными и минимальными значениями яркости.

Мы предлагаем использовать в гистограммном преобразовании так называемый средний контраст снимка. Средний контраст определяется следующим образом:

$$k = \frac{f_{\max}(x, y)}{f_{\text{cp}}}, \quad (1)$$

где k — средний коэффициент контрастности изображения; $f_{\max}(x, y)$, f_{cp} — максимальное и среднее значения яркости исходного снимка, соответственно.

Гистограммное преобразование, учитывающее средний контраст, имеет следующий вид:

$$g(x, y) = kt[f(x, y)] - \frac{t[f(x, y)]}{mn - t[0]}, \quad (2)$$

где $g(x, y)$ — выходное изображение; k — средний коэффициент контрастности исходного снимка; $f(x, y)$ — текущее значение яркости исходного снимка;

$$t[i] = \sum_{j=0}^i \frac{v[j]}{mn}, \quad (3)$$

где i — значение градации яркости точки обработанного снимка ($i = 0..255$); j — значение градации яркости точки исходного снимка ($j = 0..255, j \leq i$); $v[j]$ — значение элемента гистограммы исходного снимка при яркости j ; $t[i]$ — значение элемента гистограммы обработанного изображения при яркости i ($t[i] \in T$); $t[0] = t[f(x, y) = 0]$ — значение элемента гистограммы обработанного изображения при яркости точки, равной 0; T — область изменения элементов гистограммы выходного изображения $g(x, y)$; $m \times n$ — размер изображения.

3. Блок фильтрации изображения. Обычно преобразование гистограммы вызывает некоторое дополнительное зашумление изображения, поэтому изображение $g(x, y)$ подвергается медианной фильтрации маской 3×3 [8]:

$$d(x, y) = M[g(x, y)], \quad (4)$$

где $d(x, y)$ — выходное изображение; M — оператор медианной фильтрации.

4. Блок выделения области кисты на снимке осуществляет текстурную сегментацию.

Для сегментации снимка использован текстурный алгоритм из работы [9]:

$$e(x, y) = S[d(x, y)], \quad (5)$$

где $e(x, y)$ — выходное изображение; S — оператор текстурной сегментации.

Следующим этапом является бинаризация полученного изображения — перевод полутонового в монохромное изображение, содержащее только пиксели черного и белого цветов, путем обработки по максимальному порогу яркости сегментированного изображения:

$$b(x, y) = \begin{cases} 1 & \text{если } e(x, y) \geq Q, \\ 0, & \end{cases} \quad (6)$$

где $b(x, y)$ — яркость точки с координатами (x, y) выходного изображения; Q — пороговое значение, которое равно максимальному значению яркости точек на изображении $e(x, y)$,

$$Q = \max_{i \in [0, 255], j \in [0, 255]} e(i, j). \quad (7)$$

5. Блок выделения контуров на изображении. Выделение контуров осуществляется с помощью алгоритма из работы [9]:

$$r(x, y) = L[b(x, y)], \quad (8)$$

где $r(x, y)$ — выходное изображение, L — оператор выделения контуров.

6. Блок определения характеристик области кисты выполняет вычисление:

— площади области кисты S_k (число точек области);

— периметра области P — приближенно равен числу точек контура области;

— коэффициента формы $K_\Phi = \frac{P^2}{S_k}$.

и другие геометрические характеристики кисты.

Экспериментальное исследование алгоритма

В ходе исследования решались следующие задачи:

1. Сравнительный анализ предлагаемого и известных алгоритмов гистограммного преобразования изображений.

2. Определение возможностей формирования алгоритма выделения области кисты.

На рис. 2, а, б представлены два исходных мало-контрастных изображения кисты на маммограмме. Традиционно врач-рентгенолог диагностирует новообразования визуально. При анализе снимка, представленного на рис. 2, а, врачом-рентгенологом очаговое образование зрительно не визуализируется. На втором снимке (рис. 2, б) выявлено округлое образование с четкими неровными контурами, однородное, средней плотности. Данное образование можно интерпретировать как доброкачественное объемное образование, а также как кистозное образование.

Для уточнения диагноза следует осуществить более контрастную визуализацию снимков, поэтому используют несколько известных гистограммных преобразований и предлагаемое в данной работе гистограммное преобразование.

Гистограммы снимков указывают, что распределения их яркостей находятся в малом диапазоне [$v_{\min} = 140, v_{\max} = 170$] и [$v_{\min} = 153, v_{max} = 230$] с небольшим разбросом от средних значений $v_{\text{cp}} = 152$ и $v_{\text{cp}} = 140$, соответственно.

В экспериментах по гистограммным преобразованиям были использованы и определены такие характеристики как:

v_{\min} — минимальное значение гистограммы исходного изображения;

v_{\max} — максимальное значение гистограммы исходного изображения;

v_{cp} — среднее значение гистограммы исходного изображения;

w_{\min} — минимальное значение гистограммы обработанного изображения;

w_{\max} — максимальное значение гистограммы обработанного изображения;

w_{cp} — среднее значение гистограммы обработанного изображения;

σ — СКО изображения;

H — энтропия изображения.

После преобразования гистограмм этих снимков с помощью линейного метода [7] получим изображения, показанные на рис. 3, *a, б*, с соответствующими характеристиками. Видно, что существенных изменений снимков и их характеристик не произошло. Аналогичные результаты дали использование экспоненциального распределения и распределения Рэля.

Небольшое растяжение гистограмм на рис. 3, *a, б*, происходит в области серого, что незначительно улучшает изображение. Контрастность снимков практически не изменилась. Распределение яркостей от среднего значения на рис. 3, *a* равно $\sigma_a = 15,11$ и на рис. 3, *б*, $\sigma_b = 22,49$, и изменилось незначительно. Увеличилось значение энтропии обработанных снимков H_a и H_b , что показывает увеличение неравномерности распределения яркостных характеристик изображения. Следовательно, можно сказать, что применение распространенных алгоритмов преобразования гистограмм изображений не подходит для данного вида снимков.

На рис. 4 показаны результаты обработки исходных снимков (см. рис. 2, *a, б*) предложенным алгоритмом гистограммного преобразования на основе среднего коэффициента контрастности. Видно, что произошло растяжение гистограмм как в темной, так и в светлой области, хотя максимальные значения изменились незначительно. Увеличилось значение СКО σ для снимков, что характеризует увеличение разброса значений яркости от среднего как в сторону темных, так и в сторону светлых областей. Существенно изменилось значение энтро-

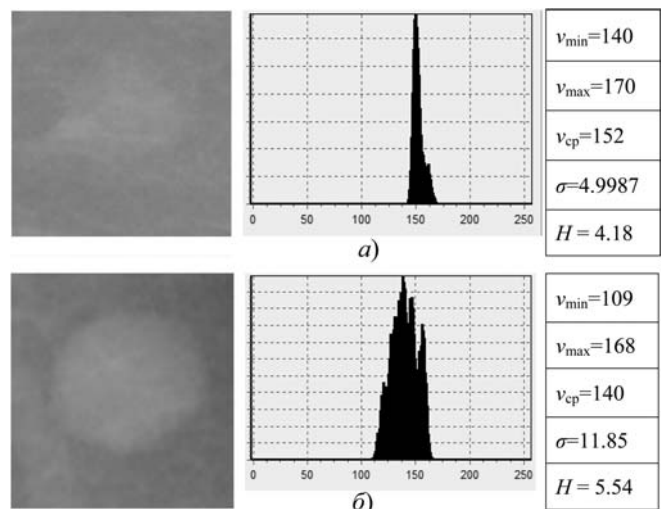


Рис. 2. Исходные изображения кисты на маммограмме

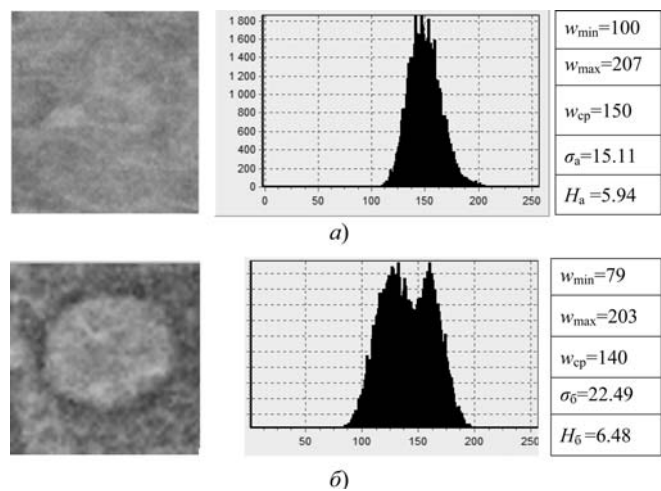


Рис. 3. Результаты линейного преобразования гистограмм: *a* — линейный метод преобразования гистограмм рис. 2, *a*; *б* — линейный метод преобразования гистограмм рис. 2, *б*

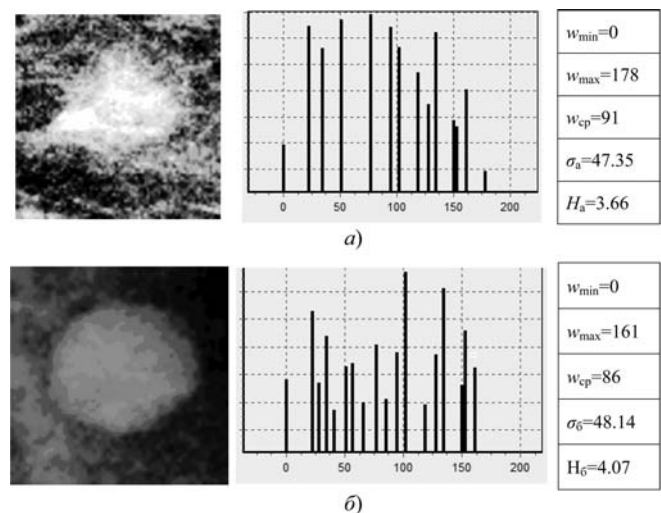


Рис. 4. Результаты гистограммного преобразования с использованием среднего коэффициента контраста: *a* — предложенный алгоритм преобразования гистограмм для рис. 2, *a*; *б* — предложенный алгоритм преобразования гистограмм для рис. 2, *б*

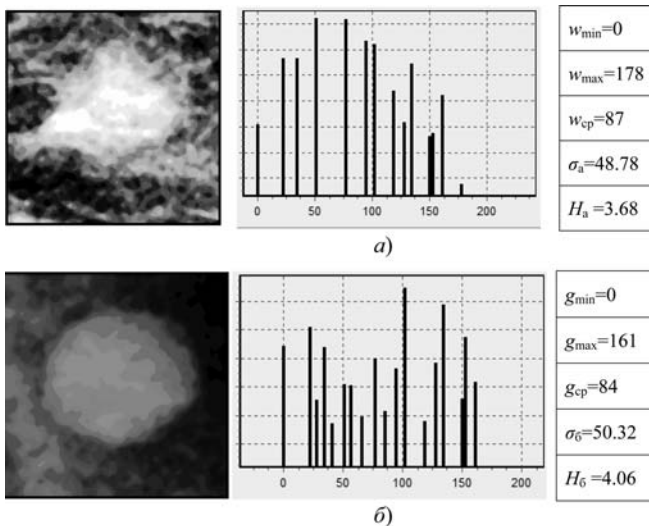


Рис. 5. Результаты медианной фильтрации: *a* — медианная фильтрация изображения для рис. 4, *a*; *б* — медианная фильтрация изображения для рис. 4, *б*

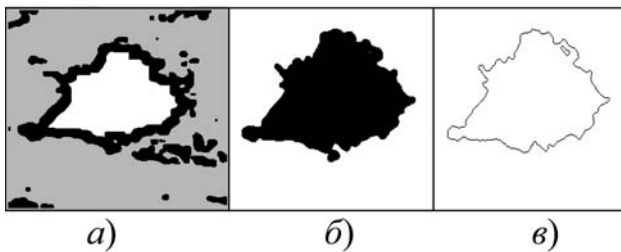


Рис. 6. Этапы выделения области новообразования на рис. 5, *а*: *a* — текстурная сегментация изображения; *б* — бинаризация; *в* — определение контура области кисты

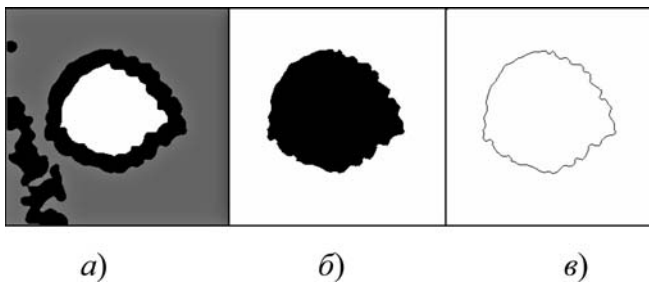


Рис. 7. Этапы выделения области новообразования на рис. 5, *б*: *a* — текстурная сегментация изображения; *б* — бинаризация; *в* — определение контура области кисты

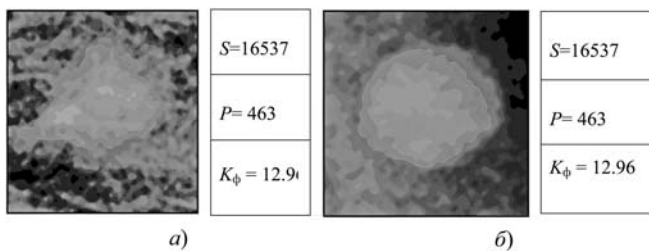


Рис. 8. Наложение контура на изображение изображения рис. 5, *а*, и рис. 5, *б* соответственно

пии H , что характеризует уменьшение неравномерности распределения яркостей изображения.

При преобразовании снимка, приведенного на рис. 2, *a* (см. рис. 4, *a*) визуализируется овальной формы образование, однородной структуры, высокой плотности с четкими контурами, интимно связанное с протоком молочной железы. Наиболее вероятен диагноз внутрипротоковой кисты молочной железы.

При анализе снимка на рис. 4, *б* выявлено округлое образование высокой плотности, однородное, частично визуализируются стенки новообразования, контур четкий. Результаты такого преобразования позволили установить четкий диагноз.

С помощью предложенного и известных гистограммных преобразований обработаны более 50 маммографических снимков с областью кисты. Известные гистограммные преобразования (линейный метод, экспоненциальное распределение и распределение Рэлея) не дали удовлетворительных результатов, в то время как предложенные гистограммные преобразования на основе среднего коэффициента контрастности дали на всех снимках удовлетворяющие врачей-рентгенологов результаты.

Дальнейшие исследования были связаны с локализацией области кисты и определением требуемых геометрических характеристик.

На рис. 5, *a*, *б* приведены результаты медианной фильтрации изображений, которая была применена к преобразованным снимкам для уменьшения зашумленности изображения.

На рис. 6 и 7 отображены операции текстурной сегментации и выделения контуров. На этих рисунках видно, что новообразование имеет четкие контуры, а также однородную структуру внутри контура, это характеризует новообразование как "кисту молочной железы".

Наложение контура на снимок позволяет более точно оценить форму и размер образования, оценить связь новообразования с сосудистым компонентом, протоками молочной железы (если таковая имеется).

При анализе снимков на рис. 8, *a* и *б* выражены четкие контуры, повторяющие контурное образование, размер образования несколько меньше видимых глазом, к тому же на рис. 8, *a* имеется интимная связь с протоком молочной железы образования рисунка.

Расчет геометрических характеристик необходим для определения точных размеров образования.

Результаты обработки достаточно сложного для визуального анализа снимка молочной железы (рис. 9) представлены на рис. 10.

При оценке рентгенограммы на рис. 9 и рис. 10, *a* консилиумом врачей-рентгенологов возникли определенные сложности, так как при отсутствии жалоб со стороны пациента на фоне жировой инволюции молочной железы выявлено очаговое образование неправильной округлой формы, визуальное с нечеткими контурами. Принимая во внимание возраст пациента (65 лет) и рентгенологическую кар-

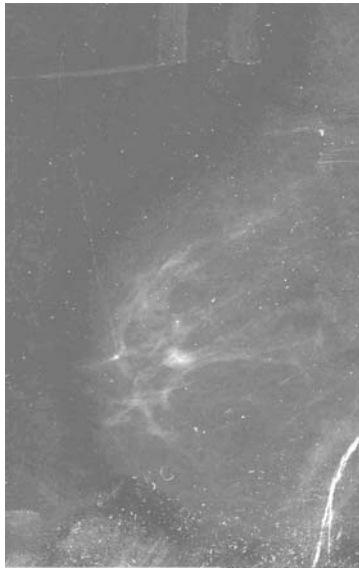


Рис. 9. Исходная маммограмма

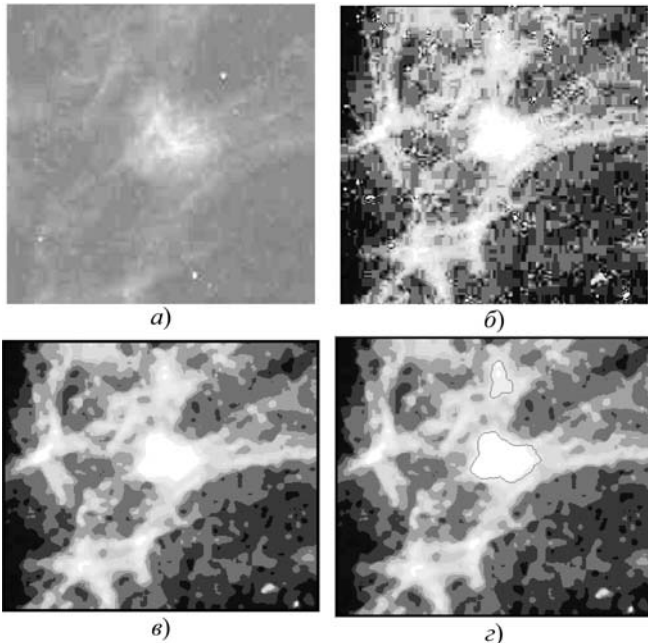


Рис. 10. Результаты обработки сложного снимка: *a* — исходная область интереса; *b* — предложенное гистограммное преобразование; *v* — медианная фильтрация области интереса; *z* — выделение контуров

тину, врачи-рентгенологи подозревали наличие злокачественного новообразования молочной железы.

При цифровой обработке рентгенограммы с использованием предложенного метода с наложением и без наложения контура определяются два идентичных округлых однородных образования с четкими ровными контурами, диаметром 7 и 4 мм соответственно, одно из которых связано с протоком молочной железы (проток в ортогональном срезе). Дано заключение "Кисты молочной железы". Женщина взята под амбулаторное наблюдение. Таким образом, удалось избежать дополнительного обследования пациента (УЗИ) и биопсии молочной железы.

Заключение

Результаты исследований 50 маммографических снимков с областью кисты показали:

- качество большинства маммографических снимков затрудняет визуальное выявление новообразований;
- широко распространенные гистограммные преобразования не дают удовлетворительных результатов обработки малоcontrastных маммографических снимков;
- предложенное преобразование гистограммы маммографического снимка на основе среднего коэффициента контраста обеспечивает хорошее улучшение характеристик малоcontrastных изображений, так как распределяет градации яркости по всему диапазону яркостей, уменьшая тем самым энтропию изображения;
- совокупность гистограммного преобразования на основе среднего контраста, медианной фильтрации, текстурной сегментации представляет собой алгоритм выделения области кисты на малоcontrastных маммографических снимках, который отвечает достаточно жестким требованиям врачей по локализации участков кисты;
- использование предложенного алгоритма обработки маммограммы дает возможность выделения области кисты для последующего распознавания новообразований в молочной железе.

Список литературы

1. Садыков С. С., Буланова Ю. А., Захарова Е. А. Диалоговая система анализа маммографических снимков // Алгоритмы, методы и системы обработки данных. 2012. Вып. 1 (19). С. 155—187.
2. Социально значимые заболевания населения России в 2010 году (Статистические материалы) / Департамент организации медицинской профилактики, медицинской помощи и развития здравоохранения ФГУ "Центральный научно-исследовательский институт организации и информатизации здравоохранения и Росздрова". Москва, 2011. 66 с.
3. Захарова Е. А., Буланова Ю. А. Обзор медицинской системы КМИС и формирование статистики заболеваемости молочных желез // Алгоритмы, методы и системы обработки данных. 2012. Вып. 1 (19). С. 54—61.
4. Заболевания молочных желез: кисты молочной железы, выделения из соска. URL: <http://www.protectmed.ru/zabolevaniya-molochnyix-zhelez>
5. Корженкова Г. П. Комплексная рентгено-сонографическая диагностика заболеваний молочной железы. М.: Фирма СТРОМ, 2004. 128 с.
6. Sankar P. R., Srinivasa Rao B. K. N. Parallel Architecture for Implementation of Contrast Limited Adaptive Histogram Equalization // International Journal of Advanced Engineering sciences and Technologies (IJAEST). 2011. Vol. 10, Is. 1. P. 047—051.
7. Rai R., Gour P., Singh B. Underwater Image Segmentation using CLANE Enhancement and Thresholding // International Journal of Emerging Technology and Advanced Engineering. January 2012. Vol. 2, Is. 1. P. 118—123.
8. Прэтт У. Цифровая обработка изображений: Пер. с англ. М.: Мир, 1982. Кн. 1 — 312 с.; кн. 2 — 493 с.
9. Nguyen H. T., Worring M., Boomgaard R. Watersnakes: Energy-Driven Watershed Segmentation // IEEE Transactions on Pattern Analysis And Machine Intelligence. March 2003. Vol. 25, No. 3. P. 330—342.

ГЕОИНФОРМАЦИОННЫЕ СИСТЕМЫ И СИСТЕМЫ ПРИРОДОПОЛЬЗОВАНИЯ

УДК 004.021:519.612

М. А. Солодовниченко, вед. разработчик iOS,
ООО "Сибирские информационные системы",

А. В. Замятин, канд. техн. наук,

проректор-директор,

Институт кибернетики

Национального исследовательского

Томского политехнического университета,

e-mail: zamyatin@tpu.ru

Алгоритм распространения лесного пожара с использованием модели Ротермела и клеточных автоматов

Рассматривается процесс моделирования распространения лесного пожара согласно модели Ротермела на основе клеточных автоматов, позволяющий предсказывать территории леса, охваченные огнем и выгоревшие в определенный момент времени, по известному ландшафту местности, скорости и направлению ветра, расположению очага возгорания. Приводится алгоритм для программной реализации, а также рассматривается один из вариантов подготовки тестовых данных. Приведены результаты сравнительных тестов разработанного программного обеспечения с известной реализацией модели Ротермела — FARSITE, сравнивающие зависимости модели от изменения направления и силы ветра, а также сопоставляющие результаты моделирования на территории с одинаковым ландшафтом.

Ключевые слова: моделирование лесных пожаров, клеточный автомат, модель Ротермела

Введение

Бурное развитие информационных технологий обеспечивает возможность применения компьютерного моделирования в качестве дополнительного инструмента в борьбе с лесными пожарами. Различные программные средства для прогнозирования распространения или расчета определенных параметров пожара успешно применяют во всем мире.

Существующие на рынке средства прогнозирования распространения лесных пожаров, такие как FARSITE, FlamMap или BehavePlus требуют зна-

чительного количества разного рода данных, вычислительных ресурсов и достаточно высокой квалификации пользователя [1–3]. Система BehavePlus позволяет с приемлемой точностью оценить различные параметры пожара, такие как поверхность распространения огня, форма, площадь и периметр пожара, продолжительность и другие. Достоинством этой системы является то, что она была включена в национальную систему определения пожарной опасности США, где прошла проверку в течение многих лет, она имеет высокую точность и универсальность. Система BehavePlus не использует пространственную информацию, характерную для географических информационных систем (ГИС), а результат представляется в виде таблиц, графиков и простых диаграмм. Это значительно снижает ее ценность ввиду невозможности визуальной оценки динамики пожара и игнорирования изменения некоторых параметров в пространстве и времени.

Система FlamMap позволяет оценить эффективность горения топлива на ландшафте, используя, в отличие от BehavePlus, данные, изменяющиеся в пространстве. Эта система вычисляет эффективность горения топлива на элементарном участке ландшафта независимо от остальных, не учитывая смену условий среды во времени, что не позволяет рассматривать изменения погоды и поведения пожара во времени. Однако полученные данные хорошо подходят для сравнения эффективности топлива, выделения и устранения топливных угроз.

Система FARSITE позволяет прогнозировать поведение пожара с учетом данных о топографии, горючем материале и состоянии среды, изменяющемся в пространстве и во времени. Достоинства этой системы — высокая точность и универсальность. В то же время для ее практического использования необходимо иметь доступ к обширному набору данных, в том числе иметь поддержку ГИС, так как требуются актуальные данные о различных параметрах каждого участка ландшафта.

Все три перечисленные системы требуют от пользователя высокой квалификации, пользовательский интерфейс выполнен на английском языке. Российских аналогов, позволяющих проводить моделирование распространения лесных пожаров и имеющих сходный набор функций, доступных для широкого практического использования, пока не существует.

В то же время проблема лесных пожаров в России и мире не перестает быть актуальной [4].

Учитывая перечисленные факты, задача разработки алгоритма распространения лесного пожара и реализации программной системы с его применением, в которой сделана попытка минимизировать недостатки существующих решений на рынке, представляется очень актуальной. В связи с этим необходима простая в использовании система, не требующая большого объема входных данных, посредством которой можно получить результаты, аналогичные результатам, полученным с использованием существующих коммерческих дорогостоящих аналогов.

Модель распространения лесных пожаров

Различают три основных класса моделей распространения лесных пожаров: эмпирические, полуэмпирические и физические [5]. Модели эмпирического класса позволяют предсказывать наиболее вероятные сценарии поведения пожара, основываясь на средних условиях и знаниях, полученных после лабораторных или наружных испытаний поведения экспериментального пожара или записях о произошедших пожарах. Результаты, полученные на основании моделей этого класса, являются наименее точными из трех. Полуэмпирические модели основаны на теории глобального энергетического баланса и, как правило, включают использование данных, полученных в результате лабораторных экспериментов. Полуэмпирические модели позволяют получать адекватные результаты и их наиболее часто используют для построения моделей, лежащих в основе программных систем, предназначенных для решения задачи прогнозирования лесных пожаров. Класс физических моделей содержит модели, которые основаны на физических принципах, позволяют описать моделируемую систему в более широком, чем в эмпирических моделях диапазоне входных переменных. Модели этого класса основаны на сложных закономерностях, учитывающих физические и химические процессы, происходящие во время лесного пожара, с максимальной из трех видов моделей детализацией. Физические модели распространения лесных пожаров позволяют получить наиболее точные результаты. Существенным недостатком класса физических моделей является сложность реализации программных систем с их использованием [6]. В связи с изложенным, для решения поставленной задачи наиболее подходит класс полуэмпирических моделей.

В качестве основы для алгоритма распространения пожара выбрана полуэмпирическая модель низового распространения пожара Ричарда Ротермела (1972), использованная в системах BehavePlus, FlamMap и FARSITE [7]. Модель низового пожара Ротермела ведет к вектору распространения максимума пожара. Она опирается на закон баланса

энергии при горении элементарного участка леса и учитывает основные факторы, влияющие на пожар, — состав и характеристики горючего материала, направление и скорость ветра, влажность воздуха и другие. В качестве результата модель выдает такие параметры как скорость распространения пожара, его интенсивность, количество выделяемой теплоты. Данная модель лежит в основе многих систем прогнозирования лесных пожаров в США и в других странах (например, Северной Америке, Южной Африке, Южной Франции и Австралии). Отмечают, что точность предсказания по модели Ротермела в простых системах лугов и саванн адекватна в широком диапазоне параметров. Однако значительные проблемы возникли в более сложных по структуре экосистемах [8].

Для удобства реализации на ЭВМ решено использовать модель низового пожара Ротермела совместно с моделью клеточных автоматов (КА). КА — это дискретная динамическая система, представляющая собой совокупность одинаковых клеток, идентичным образом соединенных между собой. Все клетки образуют так называемую решетку клеточного автомата [9]. Решетки могут быть разных типов, отличаясь как по размерности, так и по форме клеток. В данной работе используются квадратные клетки. Каждая клетка является конечным автоматом, состояния которого определяются состояниями соседних клеток и, возможно, ее собственным состоянием [10]. Модель клеточных автоматов удобна для реализации на ЭВМ, поскольку не требует хранения данных в сложном формате, проста в описании и ее реализация возможна без дополнительных затрат с применением метода параллельных вычислений, что существенно повысит производительность обработки.

Алгоритм

В работе предлагаемого алгоритма для моделирования поведения распространения лесного пожара с применением вероятностного правила КА и основных формул модели низового распространения пожара Ротермела учитывается степень горючести лесного горючего материала (ЛГМ) I_b , скорость ветра v , угол φ_w , соответствующий направлению ветра, угол φ , соответствующий уклону поверхности, и влажность воздуха h .

Область ландшафта, на которой необходимо выполнить моделирование, разбивают прямоугольной сеткой на $R \times C$ (R по ширине и C по высоте) равных квадратных блоков и создают матрицы $R \times C$, содержащие значения, соответствующие этим блокам. Каждый блок характеризуется парой чисел x и y : x — номер столбца, $x \in [0; R - 1]$; y — номер строки, $y \in [0; C - 1]$, самая верхняя строка — 0.

Для каждого такого блока необходимо предоставить среднее значение высоты поверхности. Определим матрицу $\mathbf{M}^{ah} = \{m_{ij}^{ah}, i = \overline{0, C}, j = \overline{0, R}\}$,

содержащую эти значения. Значения указаны в метрах, точка отсчета высоты не важна. Необходимо учитывать длину ребра блока a_{block} , выраженную в тех же единицах, что и высота поверхности. Также необходимы следующие входные данные:

- h — влажность воздуха, %;
- φ_w — плоский горизонтальный угол направления ветра, ° (на Север — 0°);
- v — скорость ветра, м/с;
- v_{max} — максимальная скорость ветра, м/с;
- $\mathbf{M}^{I_b} = \{m_{ij}^{I_b}, i = \overline{0, C}, j = \overline{0, R}\}$ — матрица горючести ЛГМ. Предполагается, что данные о горючести ЛГМ предварительно рассчитаны или известны и выражены в кВт/м;
- I_b^0 — порог возгораемости; ячейки, для которых выполняется неравенство $I_b \geq I_b^0$, способны перейти в состояние "горит", остальные ячейки не способны перейти в это состояние;
- I_b^{max} — максимально возможное значение горючести ЛГМ, значение выбирается экспериментально, важно чтобы $I_b^{max} \geq m_{yx}^{I_b}$;
- t_0 — минимальное время, при котором ячейка находится в состоянии "горит" (при этом ячейка имеет минимальное значение горючести $I_b - I_{b_0}$), выражается в условных логических единицах и подбирается экспериментально;
- t_{max} — максимальное время моделирования (время, по истечении которого процесс моделирования прекращается), выражено в условных логических единицах.

Основу процесса моделирования составляет расчет правил функционирования КА для определения вероятности распространения пожара. Представим область моделирования в виде матрицы $\mathbf{I} = \{i_{yx}, y = \overline{0, C}, x = \overline{0, R}\}$. Каждый КА можно определить как $\mathbf{M}^{KA} = c_{yx}$ порядка d . Каждый КА формируется из матрицы \mathbf{I} наложением на нее скользящего окна такой же размерности, как и у \mathbf{M}^{KA} . В случае, если скользящее окно выходит за границы матрицы \mathbf{I} , то доопределяем недостающие элементы значениями, равными крайним элементам \mathbf{I} .

Определим три основных возможных состояния ячейки:

$$i_{yx} = \begin{cases} 0, & \text{"не горит"} \\ 1, & \text{"горит"} \\ 2, & \text{"выгорела"} \end{cases}$$

В процессе моделирования ячейки, для которых выполняется неравенство $I_b \geq I_b^0$, переходят из состояния $i_{yx} = 0$ в состояние $i_{yx} = 1$ и затем в состояние $i_{yx} = 2$. Вероятность перехода ячейки из состояния $i_{yx} = 0$ в состояние $i_{yx} = 1$ рассчитывается с использованием вероятности перехода КА, обозначаемого

$P_{рез} \in [0, 1]$. Если ячейка перешла из состояния $i_{yx} = 0$ в состояние $i_{yx} = 1$, то для нее вычисляется

$$\text{расчетное время горения ЛГМ } t_{yx}^{burn} = \frac{I_b^0}{m_{yx}^{I_b}} t_0,$$

после истечения которого ячейка переходит в состояние $i_{yx} = 2$.

Пусть:

- K_{cil} — коэффициент силы ветра, $K_{cil} = v/v_{max}$;
- коэффициент $a_{yx}, a_{yx} = \begin{cases} 1, & i_{yx} = 1; \\ 0, & i_{yx} \neq 1; \end{cases}$
- $\mathbf{K}^{dir} = \{k_{ij}^{dir}, i = \overline{1, d}, j = \overline{1, d}\}$ — матрица влияния направления ветра. Для простоты предложены следующие правила формирования матрицы \mathbf{K}^{dir} :
 - центральная ячейка матрицы не имеет значащего элемента;
 - ближайшая ячейка от центра, находящаяся с подветренной стороны, имеет максимальное значение 1, так как $k_{yx}^{dir} = (\cos 0^\circ) = 1$;
 - каждая следующая соседняя ячейка (при обходе по часовой или против часовой стрелки) принимает значение предыдущей ячейки, умноженное на некоторый коэффициент (в данном случае коэффициент предлагается определить исходя из допущения, согласно которому ячейки равномерно распределены вокруг центрального элемента с шагом в 45° , $\cos 45^\circ = \sqrt{2}/2$);
- $k_s \in [0, 1]$ — коэффициент влияния уклона;
- угол $\varphi_{i,j}$, соответствующий уклону поверхности между двумя ячейками, можно вычислить по формуле

$$\varphi_{i,j} = \arctg\left(\frac{M^{ah}(x,y) - M^{ah}(i,j)}{a_{block}}\right).$$

Введем коэффициент Q , вычисляемый для каждой ячейки. Коэффициент Q зависит от скорости ветра, направления ветра и особенности рельефа; Q предлагается вычислять так:

$$Q = \sum_{\substack{i=0 \\ i \neq x}}^d \sum_{\substack{j=0 \\ j \neq y}}^d a_{ij} (1 + K_{cil} K_{ij}^{dir} + k_s \sin \varphi_{i,j}).$$

Введем следующие обозначения:

- $P(I_b, Q)$ — вероятность распространения пожара в зависимости от коэффициента Q , который зависит, в свою очередь, от значения уклона, соответствующего центральному элементу КА, числа ячеек, находящихся в состоянии "горит", и ветра в рассматриваемом скользящем окне, а также от горючести ЛГМ I_b ;
- $P_h(h)$ — вероятность распространения пожара с учетом влажности h воздуха, вычисляется по формуле (1).

Вероятность перехода центрального элемента КА, учитывающую степень горючести материала, ско-

рость и направление ветра, уклон поверхности и влажность воздуха, на основе модели низового распространения пожара Ротермела предлагается вычислять по формуле

$$P_{\text{рез}} \approx P(I_b, Q) + P_h(h).$$

Вероятность распространения может быть найдена как

$$P(I_b, Q) = Q(1 + I_b/I_b^{\max})/d^2.$$

Согласно модели Ротермела, скорость распространения пожара выше в сторону увеличения уклона (подъем) и ниже в сторону его уменьшения (спуск) [6]. Поэтому, чтобы внести в уравнение коэффициента Q , от которого зависят вероятности $P(I_b, Q)$, и $P_{\text{рез}}$, меру влияния угла уклона $\varphi_{i,j}$, выбрана тригонометрическая функция синус, для которой при $x \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$, $f(x) \in (-1, 1)$.

Таким образом, уклон поверхности влияет на горючесть ЛГМ I_b и на поведение распространения пожара, влияя на коэффициент Q , от которого зависит вероятность $P(I_b, Q)$ и $P_{\text{рез}}$. Большой угол уклона при других равных параметрах соответствует большему значению вероятности $P_{\text{рез}}$ и наоборот.

Вероятность возникновения пожара в зависимости от значения влажности предлагается учитывать для простоты по формуле:

$$P_h(h) = k_{P_h}/h, \quad (1)$$

где h — влажность воздуха, %, $h \in [10; 100]$; k_{P_h} — коэффициент, подбираемый экспериментально.

Признаком завершения моделирования служит факт достижения максимального времени моделирования t_{\max} или нахождения всех ячеек в состояниях $i_{yx} = 2$ или $i_{yx} = 0$.

Реализация алгоритма в ПО

Предложенный алгоритм распространения пожара реализован в программной системе, названной АЮ. Пространственные данные система АЮ получает в виде растровых изображений. Реализована возможность сохранения и загрузки параметров модели в формате INI. Программная система работает под управлением операционной системы Microsoft Windows 7. Для работы системы требуются следующие входные данные: карта горючести горючего материала, рассчитанная априори на основании модели Ротермела [7], значения представлены в киловаттах на метр в квадрате {кВт/м²}; карта высоты поверхности, значения представлены в метрах; влажность воздуха — в процентах; угол направления ветра — в градусах; скорость ветра — в метрах в секунду; минимальное значение горючести горючего материала для возгорания, макси-

мально возможное значение горючести горючего материала и некоторые другие параметры.

Результат работы система АЮ отображает в графическом виде в реальном времени, пример работы приложения представлен на рис. 1 (см. третью сторону обложки). Очаги возгорания предлагается вручную располагать на карте или же загрузить из файла.

Результаты экспериментальных исследований

Необходимо проверить адекватность результатов, получаемых в результате моделирования с использованием разработанной системы АЮ.

Поиск входных данных для модели в открытых источниках показал, что одновременно все необходимые данные получить затруднительно. Данные о начальных источниках возгорания, конечной границе выгоревшего горючего материала и других менее всего доступны. Поэтому целесообразно провести моделирование на основании одного и того же набора данных в системах АЮ и одной из лидирующих на рынке, после чего сравнить полученные результаты.

В качестве аналога для сравнительного анализа была выбрана система FARSITE. Главным образом это связано с тем, что систему FARSITE можно сконфигурировать таким образом, чтобы в качестве модели распространения пожара была использована модель на основании модели Ротермела [11]. Кроме того, FARSITE активно использовали в правительственных организациях США. Система FARSITE, как и собственная разработка АЮ, отображает результаты моделирования в графическом виде (существует возможность экспорта результатов моделирования в различные форматы данных), но требует больше различных входных данных.

Для экспериментов были заданы следующие значения констант:

$$I_b^{\max} = 200 \text{ кВт/м}; v_{\max} = 15 \text{ м/с}; t_0 = 20; k_{P_h} = 10.$$

Эксперименты со стандартными моделями топлива [12], анализ документации к FARSITE и собственной реализации показали невозможность использования данных о горючести ЛГМ из FARSITE в разработанной модели. Связано это с разным подходом в реализации модели Ротермела. В FARSITE горючесть ЛГМ полностью генерируется во время моделирования с учетом ветра и уклона [12], в то время как в АЮ горючесть ЛГМ является входным параметром. Причем считается, что уклон поверхности и параметры ветра не учтены в горючести горючего материала и должны быть учтены во время моделирования. Такое разделение связано с применением КА, оптимизации вычислений и упрощением представления имеющихся данных. Кроме того, такие параметры модели АЮ как время сгорания самой медленно горящей ячейки t_0 , коэффициент влияния ветра K_{cil} , порог возгораемости I_b ,

максимум горючести I_b^{\max} и коэффициент влияния уклона k_s должны быть отличными для различных моделей топлива, что приводит к ограничению в одну модель топлива для всех областей ландшафта [1].

Далее представлены результаты экспериментов, подтверждающие соответствие системы АЮ и системы FARSITE положениям модели Ротермела.

Эксперимент, проведенный в целях выявления схожего поведения модели АЮ и модели FARSITE при изменении скорости ветра и неизменных остальных параметрах, выявил схожесть поведения рассматриваемых систем (рис. 2, а и б, см. третью сторону обложки). При остальных неизменных параметрах скорость распространения пожара в направлении вектора направления ветра больше при большей скорости ветра. На рис. 2 изображены области распространения пожара с разным значением скорости ветра, полученные в АЮ (рис. 2, а) и в FARSITE (рис. 2, б), стрелкой показано направление ветра. Для FARSITE как и для АЮ меньшие значения параметров, относящиеся к влиянию скорости ветра, соответствуют меньшей дальности распространения пожара и наоборот.

Эксперимент в целях оценки изменения параметра угла направления ветра выявил схожесть поведения рассматриваемых систем (рис. 3, а и б, см. третью сторону обложки). При остальных неизменных параметрах скорость распространения пожара выше в сторону направления ветра и ниже против направления ветра. На рис. 3 (см. третью сторону обложки) изображены области распространения пожара с разным значением угла направления ветра, полученные в АЮ (рис. 3, а) и в FARSITE (рис. 3, б), стрелкой показано направление ветра.

Эксперимент в целях выявления схожего поведения модели АЮ и модели FARSITE при изменении уклона поверхности и неизменных остальных параметрах выявил схожесть поведения рассматриваемых систем (рис. 4, а, б, см. третью сторону обложки). При остальных неизменных параметрах скорость распространения пожара выше в сторону увеличения высоты поверхности и ниже в сторону уменьшения высоты поверхности.

На рис. 4 (см. третью сторону обложки) изображены области распространения пожара с разным направлением увеличения высоты поверхности, полученные в АЮ (рис. 4, а) и в FARSITE (рис. 4, б), стрелкой показано направление увеличения высоты поверхности. Угол увеличения высоты поверхности составляет 45° .

Цель последнего эксперимента — выявление схожего поведения модели АЮ и модели FARSITE при использовании сопоставимых входных данных, полученных на основании реальных сведений о ландшафте.

Данные о высоте поверхности участка ландшафта получены из сетевой ГИС "Космоснимки" [13]. На основании карты рельефа с использованием

программной системы *CrazyBump* сгенерирована карта высоты [14]. Полученная карта высоты в исходном виде была использована в качестве входного параметра "карта высоты" для системы АЮ. Для системы FARSITE на основании изображения карты высоты с использованием программного средства BMP2ASC сгенерированы в формате ASC карта высоты поверхности и карта уклона поверхности, карта направленности поверхности.

Поскольку система АЮ одновременно оперирует только одной моделью топлива, карта горючести содержала константную величину ($M^{I_b} = \{m_{ij}^{I_b} = 100, i = \overline{0, C}, j = \overline{0, R}\}$), а для системы FARSITE была указана константная модель топлива [13].

В экспериментах с системой FARSITE, кроме описанных выше, были созданы следующие условия:

- значения покрытия (*Canopy Cover*) в ячейках карты для всех областей ландшафта были одинаковыми и составляли 100 %;
- значения подстройки скорости распространения пожара для всех моделей топлива составляли 1, что соответствует отсутствию подстройки;
- ветер отсутствовал;
- температура воздуха составляла 20°C ;
- время симуляции составляло 4 сут.

На рис. 5 изображены результаты моделирования, полученные с использованием системы АЮ за 60 итераций с шагом 10 итераций и системы FARSITE за 1 сутки с шагом 1 ч. Контурные линии кромки пламени пространственно совмещены так, чтобы совпали их точки воспламенения.

Видно, что форма кромки пламени, полученная на основании моделирования с использованием системы АЮ, напоминает форму очага распространения пожара, полученную на основании моделирования с использованием системы FARSITE, направления распространения пламени схожи (рис. 5).

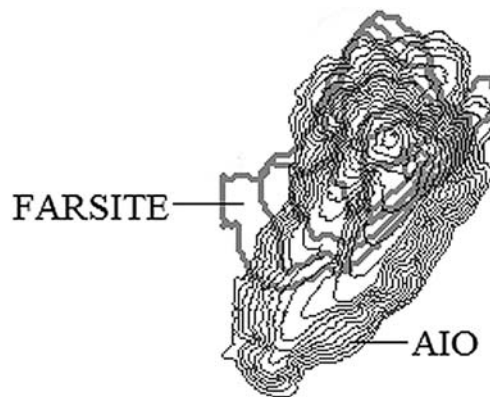


Рис. 5. Результаты моделирования распространения лесного пожара при сопоставимых параметрах, полученные для АЮ и FARSITE

Эксперимент выявил схожесть поведения модели AIO и модели FARSITE при использовании сопоставимых входных данных, полученных на основании реальных сведений о ландшафте. Полученные результаты экспериментов показали соответствие поведения модели собственной разработки AIO и модели FARSITE в рамках одной модели данных для всех рассмотренных положений. Учитывая изложенное выше, есть основания полагать, что откалибровав необходимым образом модель AIO и определив по логическим единицам физические, в результате использования полученной системы моделирования лесных пожаров возможно получение сведений о поведении пожара согласно модели поверхностного распространения пожара Ротермела с использованием одной модели топлива с достаточно высокой адекватностью. Причем получение прогнозных данных о поведении пожара другими средствами с аналогичной адекватностью затруднительно.

Заключение

В результате экспериментальных исследований выявлена возможность использования разработанной системы в задачах моделирования распространения лесных пожаров в рамках одной модели топлива. При этом существует возможность расширить область решаемых задач, усложнив разработанную систему путем добавления дополнительных данных для некоторых параметров. Такие параметры, как время сгорания самой медленно горящей ячейки, коэффициент влияния ветра, порог возгораемости, максимум горючести и коэффициент влияния ветра должны варьироваться для разных моделей топлива, поэтому добавление слоев пространственных данных с данными параметрами модели могло бы снять ограничение в одну модель топлива.

Получен инструмент, который в сравнении с одним из лидеров аналогичного ПО — FARSITE, требует меньшего числа входных данных, более простой в использовании, но одновременно демонстрирующий в высокой степени аналогичные результаты

в рамках одной модели топлива. Очевидно, для более широкого практического применения система AIO требует доработки в части поддержки различных форматов входных данных и визуализации выходных данных, а модель, лежащая в ее основе — внимательного калибрования.

Работа выполнена при поддержке РФФИ (грант № 11-07-00027а) и Госзадания (№ 8.8113.2013).

Список литературы

1. **Finney M. A.** FARSITE: Fire Area Simulator — Model Development and Evaluation. 1998. Research Paper RMRS-RP-4 Revised, USDA Forest Service.
2. **Finney M. A.** An Overview of FlamMap Fire Modeling Capabilities // USDA Forest Service Proceedings RMRS-P-41. 2006. P. 212—220.
3. **Andrews P. L.** BehavePlus fire modeling system: past, present, and future // Proceedings of 7th Symposium on Fire and Forest Meteorology Society, 2007.
4. **Brushlinsky N. N., Hall J. R., Sokolov S. V., Wagner P.** Fire statistics: Report № 11. 2006.
5. **Баровик Д. В., Таранчук В. Б.** Состояние проблемы и результаты компьютерного прогнозирования распространения лесных пожаров // Вестник БГУ. 2011. № 3. С. 78—84.
6. **Гришин А. М.** Математическое моделирование лесных пожаров и новые способы борьбы с ними. Новосибирск: Наука, 1992. 408 с.
7. **Rothermel R. C.** A mathematical model for predicting fire spread in wildland fuels. // Res. Pap. INT-115. Ogden, UT: U.S. Dept. of Agriculture, Forest Service, Intermountain Forest and Range Experiment Station. 1972. 40 p.
8. **Бурасов Д. М.** Математическое моделирование низовых лесных и степных пожаров и их экологических последствий. Томск, 2006.
9. **Packard N. H., Wolfram S.** Two-dimensional cellular automata // Journal of statistical physics. 1985. Vol. 38. P. 901—946.
10. **Bodrozic Lj., Stipanicev D., Seric M.** Forest fires spread modeling using cellular automata approach // Modern trends in control. 2006.
11. **Finney M. A., Andrews P. L.** Application and status of the far-site fire area simulator // Proc. of III International Conference on Forest Fire Research 14th Conference on Fire and Forest Meteorology. 1998. Luso, 16/20 November 1998. Vol. 1. P. 755—760.
12. **Anderson H. E.** Aids to determining fuel models for estimating fire behavior: General Technical Report INT-122, USDA Forest Service, 1982.
13. **Космоснимки.** URL: <http://kosmosnimki.ru/> (дата обновления: 19.05.2012).
14. **Crazy Bump.** URL: <http://crazybump.com/> (дата обновления: 5.11.2011).

ИНФОРМАЦИЯ



С 1 по 3 октября 2013 в Москве, в ЦВК "Экспоцентр" состоится

**2-я международная форум-выставка
передовых информационных технологий**

"ИТ СТРАТЕГИИ ЛИДЕРСТВА"

Тематическая направленность выставки:

- IT решения по управлению корпоративной информацией
- Инфраструктура и системы хранения данных
- Технологии управления, обучения и подготовки кадров

Контактная информация:

Тел.: (812) 320-8098, 320-0141

Факс: (812) 320-8090

E-mail: itcom@restec.ru, ict-dep@restec.ru

Л. Р. Бекирова, канд. техн. наук, доц., докторант,
Государственная нефтяная академия,
г. Баку, Азербайджан,
e-mail: Bekirova_Lale@rambler.ru

Вопросы минимизации влияния изменчивости атмосферного аэрозоля на информативность спутниковых мультиспектральных изображений

На основе фактического материала показано, что многомодальность функции распределения аэрозоля приводит к существенным отклонениям от экспоненциального вида функции зависимости оптической толщины аэрозоля от длины волны.

Получено математическое выражение для вычисления оптимальной функциональной зависимости оптической толщины аэрозоля от длины волны в смысле минимизации предложенного интегрального энтропийного критерия.

Отмечено, что минимизация интегрального энтропийного критерия позволяет уменьшить влияние атмосферного аэрозоля на многоспектральные космические изображения.

Ключевые слова: мультиспектральные изображения, атмосфера, аэрозоль, информативность, энтропийный критерий

Хорошо известно, что спутниковые аэрозольные измерения базируются на зондировании рассеянного аэрозодем Солнечного света. В спутниковом дистанционном зондировании поверхности Земли роль атмосферного аэрозоля такова, что благодаря существованию аэрозоля спутник измеряет не действительное отражение земной поверхности, а кажущееся отражение [1]. Таким образом, высокоточное спутниковое дистанционное зондирование в любом случае предполагает изучение физических свойств атмосферного аэрозоля, а также оценку влияния аэрозоля на достоверность космических изображений.

Спутниковые аэрозольные измерения обычно реализуются путем взаимодействия следующих компонентов:

- солнечное излучение;
- объект исследования — поверхность Земли;
- измеритель излучения, установленный на спутнике;
- компоненты атмосферы — аэрозоль, воздух и облака.

Солнечная радиация достигает поверхности Земли в виде прямого или рассеянного излучения.

Часть излучения обратно рассеивается в космос и детектируется спутниковой аппаратурой. Часть радиации, попадающая на поверхность Земли, отражается, далее ослабляется атмосферой, но также усиливается путем аэрозольного рассеяния в поле зрения сенсоров.

Как указывается в работе [1], суммарная радиация, поступающая на вход спутника, содержит две компоненты: радиация, отраженная от поверхности Земли; радиация, отраженная от аэрозоля. Как было указано в работе [2], решение уравнения переноса радиации дает следующее выражение для результирующей радиации поверхностного отражения радиации:

$$R = [R_0 + (e^\tau - 1)P]e^{-\tau}, \quad (1)$$

где R — кажущееся отражение поверхности Земли; R_0 — отражение от поверхности; $(e^\tau - 1)$ — отражение аэрозоля; $e^{-\tau}$ — ослабление при прохождении через аэрозольный слой атмосферы; τ — оптическая толщина атмосферного аэрозоля; P — фазовая функция рассеяния аэрозоля.

Таким образом, аэрозоль влияет на радиацию, отраженную от поверхности двояким образом. С одной стороны, в результате рассеяния света аэрозольными частицами регистрируемая яркость поверхности усиливается. Этот эффект зависит от оптической толщины аэрозоля, а также от фазовой функции аэрозольного рассеяния P . С другой стороны, аэрозоль экспоненциально ослабляет радиацию, поступающую от поверхности, а также радиацию, рассеянную аэрозодем. На рис. 1 приведены графики зависимости, регистрируемой на спутнике, сигнала кажущегося отражения от отношения P/R_0

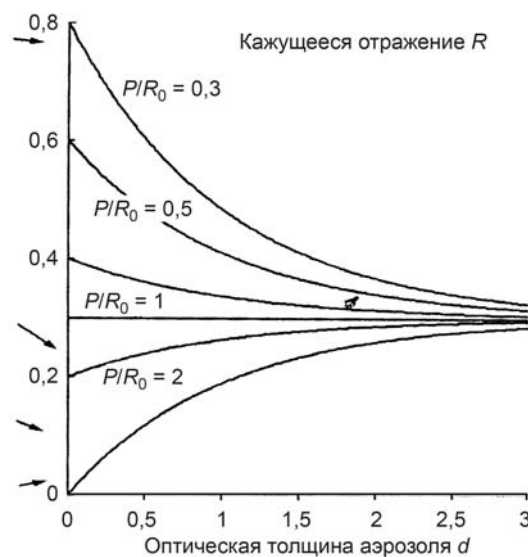


Рис. 1. Зависимость кажущегося поверхностного отражения от параметра P/R_0 и оптической толщины аэрозоля

и оптической толщины аэрозоля. Зависимость кажущегося отражения с поверхности исследована достаточно подробно и основные результаты этих исследований приведены в работе [1]. Цель настоящей статьи — информационная оценка влияния вариаций оптической толщины аэрозоля, возникающих при появлении эффекта многомодальности функции распределения массы аэрозоля по размерам.

Подробно рассмотрим результаты экспериментальных исследований влияния многомодальности функции распределения аэрозоля по размерам на волновую зависимость оптической толщины аэрозоля.

В работе [3] приведена спектральная серия измеренных значений оптической толщины аэрозоля с помощью солнечного фотометра типа IP-RAD (рис. 2). На рис. 3 приведены графики соответствующих функций распределения аэрозоля по размерам, вычисленные на основе метода инверсии Кинга [4]. Как видно из приведенных графиков, двухмодальность распределения аэрозоля по размерам значительно влияет на форму кривой волновой зависимости оптической толщины аэрозоля.

Отметим, что аналогичные результаты экспериментальных исследований приведены во многих работах, например [5—7].

Влияние указанного эффекта изменения волновой зависимости оптической толщины аэрозоля на спутниковые колориметрические измерения бесспорно. При этом в широком смысле слова можно говорить об искажении цветности во всех спутниковых мультиспектральных изображениях.

С учетом таких искажений задачу дальнейших исследований можно сформулировать следующим образом: вычислить такую функцию волновой зависимости оптической толщины аэрозоля, при которой прирост информативности вследствие влияния аэрозоля был бы минимальным. Для математической формулировки поставленной задачи воспользуемся выражением (1), интегральная энтропийная оценка задачи будет иметь следующий вид:

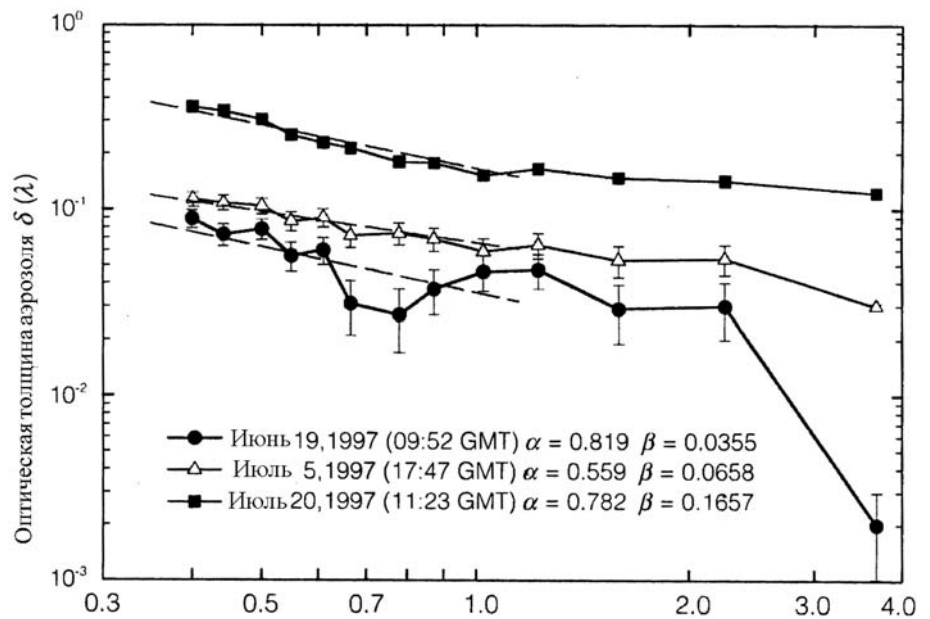


Рис. 2. Спектральная серия измеренных значений оптической толщины аэрозоля, 19 июня, 5 июля и 20 июля 1997 г. [3]

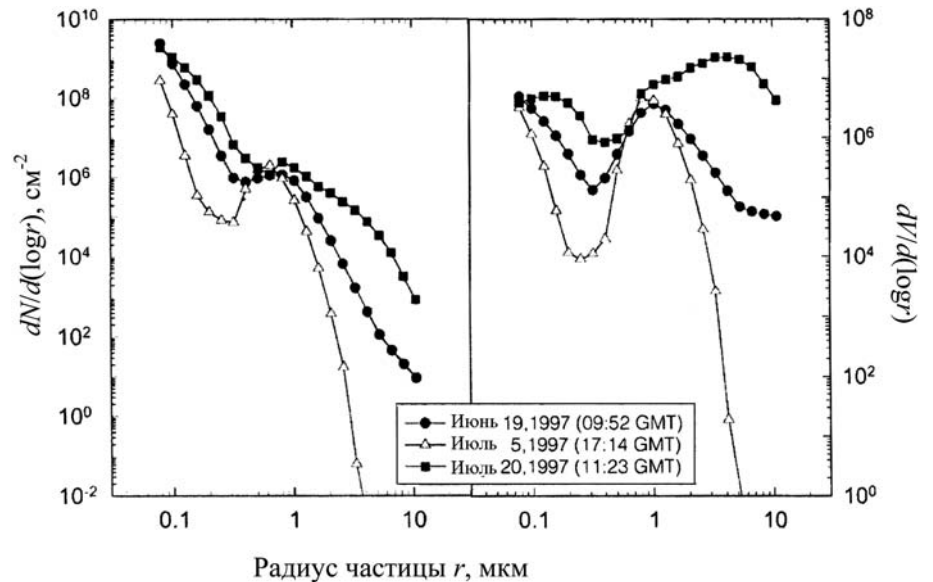


Рис. 3. Кривые распределения по радиусу общего количества аэрозольных частиц (слева) и объемного распределения (справа) [3]

$$F_1 = \int_0^{\lambda_{\max}} \ln[R_0 + (e^{\tau(\lambda)} - 1)P]e^{-\tau(\lambda)} d\lambda, \quad (2)$$

где λ — длина волны.

Требуется вычислить такую оптимальную функцию $\tau(\lambda)$, при которой функционал цели (2) достигает минимального значения. В качестве ограничительного условия примем следующее выражение:

$$\int_0^{\lambda_{\max}} e^{-\tau(\lambda)} d\lambda = C = \text{const}. \quad (3)$$

С учетом выражений (2) и (3) функционал безусловной вариационной оптимизации имеет следующий вид:

$$F_2 = \int_0^{\lambda_{\max}} F_3[R_0, \tau(\lambda), P] d\lambda = \int_0^{\lambda_{\max}} \ln[R_0 + (e^{\tau(\lambda)} - 1)P] e^{-\tau(\lambda)} d\lambda + \gamma \int_0^{\lambda_{\max}} e^{-\tau(\lambda)} d\lambda, \quad (4)$$

где γ — множитель Лагранжа.

Согласно методу Эйлера оптимальная функция $\tau(\lambda)$ должна быть выбрана из условия

$$\frac{dF[R_3, \tau(\lambda), P]}{d\tau(\lambda)} = 0. \quad (5)$$

С учетом выражений (4) и (5) получаем

$$\frac{e^{\tau(\lambda)} P}{R_0 + (e^{\tau(\lambda)} - 1)P} - 1 - \gamma e^{\tau(\lambda)} = 0. \quad (6)$$

Из выражения (6) находим

$$e^{\tau(\lambda)} = \frac{P\gamma + R_0 - P}{P\gamma - R_0\gamma}. \quad (7)$$

С учетом выражений (3) и (7) имеем

$$\gamma = \frac{\int \frac{R_0}{P - R_0} d\lambda - \int \frac{P}{P - R_0} d\lambda}{C - \int \frac{P}{P - R_0} d\lambda} = C_1, \quad (8)$$

где $C_1 = \text{const}$.

С учетом выражений (6) и (8) нетрудно получить следующее выражение для $\tau(\lambda)$:

$$\tau(\lambda) = \ln \left[\frac{C_1 P - C_1 R_0}{C_1 P - P + R_0} \right]. \quad (9)$$

Таким образом, полученное решение (9) позволяет вычислить оптимальный вид функции $\tau(\lambda)$, при которой функционал (4) достигает экстремального значения.

Вычислим условие, когда функционал (4) достигает минимального значения.

Взяв вторую производную функционала (4), можно показать, что она положительна при выполнении условия

$$PR_0 + P^2 > \gamma. \quad (10)$$

Таким образом, показано существование оптимальной функции (9), при которой функционал (4) достигает минимального значения при выполнении условия (10). Минимальная интегральная энтропия, достигаемая при решении (9), означает, что функционал претерпевает минимальные изменения при изменении $\tau(\lambda)$, это минимизирует влияние изменчивости аэрозоля к многоспектральной космической информации.

Следовательно, полученное решение позволяет определить условие минимизации влияния изменчивости атмосферного аэрозоля на информативность спутниковых мультиспектральных изображений.

В заключение сформулируем основные выводы и положения проведенного исследования:

1. На основе фактического материала показано, что многомодальность функции распределения аэрозоля приводит к существенным отклонениям от экспоненциального вида функции зависимости оптической толщины аэрозоля от длины волны.

2. Получено математическое выражение для вычисления оптимальной функциональной зависимости оптической толщины аэрозоля от длины волны в смысле минимизации предложенного интегрального энтропийного критерия.

3. Отмечено, что минимизация интегрального энтропийного критерия позволяет уменьшить влияние атмосферного аэрозоля на многоспектральные космические изображения.

Список литературы

1. Husar R. B. Satellite Measurements of Atmospheric Aerosols. Aerosol measurement: Principles, Techniques, and Applications // URL: http://capita.wustl.edu/CAPITA/CapitaPesports/090120_AerosolMeasHandbook/Reviews/Reviewer%2021.pdf
2. Husar R. B., White W. H. On the color of the Los Angeles smog // Atmospheric Environment. 1976. Vol. 10, No. 3. P. 199–204.
3. Tomasi C., Vitale V., Lupi A., Cacciari A., Marani S., Bonafe U. Marine and continental aerosol effects on the upwelling solar radiation flux in Southern Portugal during the ACE-2 experiment // Annals of Geophysics. 2003. Vol. 46, No. 2. P. 467–479.
4. King M. D. Sensitivity of constrained linear inversions to the selection of the Lagrange multiplier // Journal Atmospheric Science. 1982. Vol. 39. P. 1356–1369.
5. Dumka U. C., Sagar R., Pant P. Retrieval of columnar aerosol size distributions from spectral attenuation measurements over Central Himalayas // Aerosols and Air Quality Research. 2009. No. 9. P. 344–351.
6. Buchard V., Brogniez C., Auriol F., Bonnel B. Aerosol single scattering albedo retrieved from ground-based measurements in the UV and visible region // Atmospheric Measurement Techniques. 2011. No. 7. P. 1–7.
7. Petters J. L., Saxena V. K., Slusser J. R., Wenny B. N., Madronich S. Aerosol single scattering albedo retrieved from measurements of surface UV irradiance and a radiative transfer model // Journal of Geophysical Research. 2003. Vol. 108, No. D9.4288, doi:10.1029/2002JD002360

УДК 378.1 + 311

Б. Д. Залещанский,

д-р экон. наук, проф., зав. каф.,
Московский государственный
технический университет радиотехники,
электроники и автоматики (МИРЭА),

А. П. Свиридов,

д-р техн. наук, проф., вице-президент,
Международная академия информатизации,
e-mail: prof_sviridov@mail.ru,

О. А. Шалобина, аспирант,

Е. А. Шалобина, аспирант,

Московский государственный
технический университет радиотехники,
электроники и автоматики (МИРЭА)

Вероятностно-статистические стратегии обеспечения качества подготовки персонала социотехнических систем

Рассматриваются периодические стратегии управления качеством профессиональной подготовки обучаемых (персонала) с учетом интенсивности забывания и экономических показателей (затрат на обучение, контроль и потерь вследствие ошибок персонала и др.), основанные на оптимальном планировании во времени повторений или предупредительных восстановлений знаний. Требуемая профессиональная готовность обучаемого (персонала) обеспечивается путем контроля и последующего управления качеством подготовки.

Ключевые слова: стратегии, переподготовка персонала, потери вследствие ошибок, интенсивность забывания

Введение

Эффективность функционирования современных социотехнических систем определяется не только техническими и программными средствами, но и деятельностью людей в их составе. В настоящее время для обеспечения надежности технических и программных средств широко используют методы технической диагностики и стратегии проведения профилактических работ с учетом характеристик надежности. Важнейшей проблемой обеспечения эффективности социотехнических систем в настоящее время является управление профессиональной

надежностью персонала предприятий и организаций на основе количественных характеристик, динамики знаний, умений, навыков и компетенций, экономических показателей и др. [1–5]. Следует отметить, что интенсивность отказов вследствие ошибок персонала значительно превышает интенсивность отказов собственно технических и программных средств. В американских рекомендациях WASH-1400 предполагается, что интенсивность отказов, связанных с человеческим фактором, по крайней мере в 100 раз больше интенсивности отказа технических средств атомной станции. На немецких атомных станциях ошибки персонала присутствуют в более чем 60 % всех несчастных случаев [1].

В настоящее время весьма актуальной является проблема оптимального управления процессами динамики знаний обучаемых (персонала) с учетом интенсивности забывания (потери работоспособности), экономических показателей (затрат на обучение, контроль и потери вследствие ошибок персонала и др.), рисков недооценки и переоценки при компьютерном контроле знаний и др. Особо актуальна эта проблема для процессов и производств с высокой "ценой" ошибки персонала.

В данной работе предлагается подход к обеспечению качества подготовки обучаемых и профессиональной надежности персонала предприятий и организаций для эффективного функционирования социотехнических систем на основе результатов статистической теории обучения [4–7].

Приведем содержание понятий "знание" и "обучаемые". Термин "знание" применяется для обозначения знаний, умений, навыков и компетенций [8, 9]. Под "обучаемыми" в данной работе понимаются:

- учащиеся школ, студенты вузов и университетов, слушатели курсов подготовки, переподготовки и повышения квалификации;
- обслуживающий персонал технических и программных средств информационно-телекоммуникационных систем;
- персонал атомных, тепловых и гидроэлектростанций, транспорта;
- операторы химико-технологических и иных процессов и производств (в первую очередь, операторы процессов и производств с "высокой ценой" ошибки);
- личный состав спасательных служб (например, по оказанию первой медицинской и психологической помощи населению и организации взаимодействия);

— отдельные индивиды (население) для подготовки поведения в случае неблагоприятных событий (проблемы безопасности и сокращения рисков).

1. Проблема оптимизации

На основе количественных характеристик динамики знаний можно оптимизировать управление учебным процессом. При этом проблема оптимизации в общем случае ставится следующим образом. Усвоение, забывание, повторение и восстановление знаний образуют некоторый случайный процесс. Имеется некоторое множество управлений этим процессом. Для каждой комбинации из реализации случайного процесса и соответствующего управления можно определить функционал цели (степень полезности). Этот функционал содержит некоторые известные и неизвестные параметры.

Примеры известных параметров: 1) интенсивности усвоения и забывания; 2) различные затраты; 3) риски недооценки и переоценки знаний при (компьютерном) контроле знаний и др.

Примеры неизвестных параметров: 1) интервал восстановления знаний, под которым понимается промежуток времени a до следующего восстановления знаний; 2) объем и глубина повторения или восстановления знаний по учебному материалу и др.

Проблема оптимизации состоит при этом в синтезе стратегии управления случайным процессом, экстремизирующей (максимизирующей или минимизирующей) функционал цели или качества.

2. Стратегия повторения или восстановления знаний в календарные моменты времени при известной функции распределения времени забывания

2.1. Периодические стратегии без учета рисков недооценки и переоценки знаний

Повторения или восстановление знаний, умений и навыков организуются следующим образом. Непосредственно после усвоения учебного материала через некоторый промежуток времени a планируется проведение контроля знаний и последующего повторения учебного материала (ПУМ). Если обучаемый (оператор) в течение этого промежутка времени не забывает учебный материал, то проводится его повторение со средней длительностью T_w . В случае же забывания учебного материала до момента проверки организуется восстановление знаний со средней длительностью T_e ($T_e > T_w$). По завершении повторения или восстановления знаний, умений или навыков планируется следующая проверка и повторение через интервал времени a и т. д. При этом обучаемый (оператор) в произвольный момент времени может находиться в следующих состояниях знания: E_0 — состояние усвоения знаний, когда он может выполнить требуемые задачи; E_1 — состояние необнаруженного (скрытого) забывания; E_2 — состояние контроля и последующего восстановле-

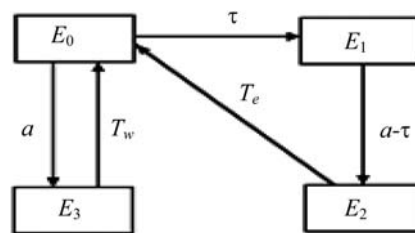


Рис. 1. Граф переходов обучаемого (оператора) из одного состояния в другое

ния знаний; E_3 — состояние контроля и последующего повторения знаний.

Соответствующий граф переходов обучаемого (оператора) из одного состояния знания в другое по положению (ТЕД) приведен на рис. 1.

На его основе можно получить соотношения для коэффициента профессиональной готовности $K(a)$ обучаемого (оператора) и его максимума:

$$K(a) = \frac{\int_0^a [1 - F(x)] dx}{a + T_w + (T_e - T_w)F(a)}; \quad (1)$$

$$K(a_0) = \max_a K(a) = \frac{1 - F(a_0)}{1 + (T_e - T_w)f(a_0)}, \quad (2)$$

где $F(x)$ — функция распределения времени забывания; a — интервал между проверками и повторениями; T_w — время повторения; T_e — время восстановления знаний.

Если время забывания имеет экспоненциальное распределение $\text{Exp}(\lambda)$, то оптимальный интервал a_0 определяется в виде решения уравнения $dK(a)/da = 0$ или

$$T_w = e^{\lambda a} - 1 - \lambda a = g(\lambda a), \quad (3)$$

где λ — интенсивность забывания.

Пример 1. Пусть для некоторого обучаемого (оператора) определено $\lambda = 0,018$ 1/мес; $T_w = 8$ ч, $T_e = 16$ ч.

Определить оптимальный интервал между повторениями и максимальное значение коэффициента профессиональной готовности.

Решение. На рис. 2 приведена зависимость коэффициента профессиональной готовности от интервала между проверками: $a_0 = 108$ ч, $K(a_0) = 0,78$.

Учет затрат при управлении профессиональной готовностью. Пусть c_i — затраты в единицу времени, когда обучаемый (оператор) находится в состоянии E_i , $i = 1, 2, 3$. При этом средние удельные затраты будут равны [3—6]

$$C^* = K^{-1} \sum_{i=1}^3 c_i k_i, \quad (4)$$

где k_i — стационарная вероятность того, что обучаемый (оператор) находится в состоянии знания E_i , $i = 1, 2, 3$, а K — коэффициент профессиональ-

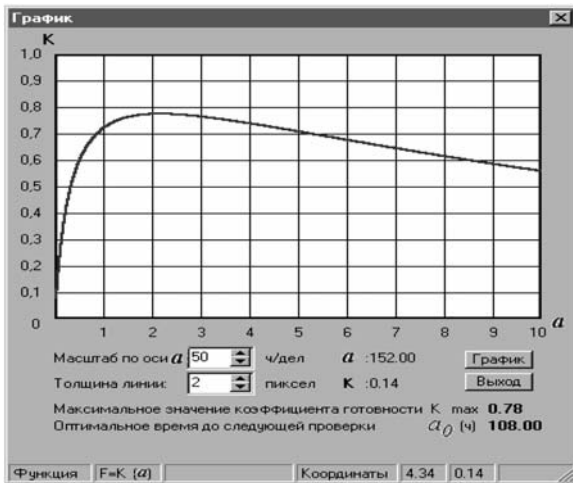


Рис. 2. Зависимость коэффициента профессиональной готовности от интервала между проверками

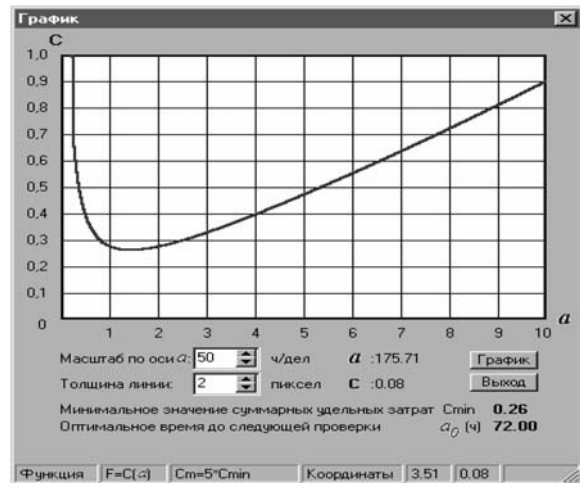


Рис. 3. Зависимость средних удельных потерь от интервала между проверками

ной готовности. Подставив значения k_i и K в соотношение (4), получим

$$C^*(a) = \frac{a \int_0^a F(x) dx + c_3 T_w + (c_2 T_e - c_3 T_w) F(a)}{a \int_0^a [1 - F(x)] dx}.$$

Оптимальный интервал a_0 , обращающий в минимум средние потери, определяется в виде решения уравнения $dC^*(a)/da = 0$ или

$$\frac{c_3 T_w}{c_2 T_e - c_3 T_w} = -F(a) + \lambda(a) \int_0^a [1 - F(x)] dx + \frac{c_1 \int_0^a x dF(x)}{(c_2 T_e - c_3 T_w) F(a)}. \quad (5)$$

Минимум средних удельных затрат при этом

$$\begin{aligned} \min_a C^*(a) &= C^*(a_0) = \\ &= C_1 \frac{F(a_0)}{1 - F(a_0)} + \lambda(a_0)(c_2 T_e - c_3 T_w). \end{aligned}$$

Если время забывания имеет экспоненциальное распределение $\text{Exp}(\lambda)$, то из уравнения (5) следует

$$\frac{C_3}{C_1} \lambda T_w = e^{\lambda a} - 1 - \lambda a = g(\lambda a). \quad (6)$$

Пример 2. Дано: $\lambda = 0,0015 \text{ ч}^{-1}$, $T_w = 8 \text{ ч}$, $T_e = 16 \text{ ч}$, $c_1 = 2$, $c_2 = 2$, $c_3 = 1$.

Определить: 1) оптимальный интервал a_0 между проверками и повторениям или восстановлениями знаний; 2) минимум средних потерь $C(a_0)$.

Решение. На рис. 3 приведена зависимость $C(a)$ средних удельных потерь от интервала между проверками: $a_0 = 72 \text{ ч}$, $C_{\min} = C(a_0) = 0,26$.

2.2. Строго периодическое повторение или восстановление знаний при учете рисков недооценки и переоценки знаний

Повторения и восстановления знаний организуются подобно предыдущей стратегии. Непосредственно после усвоения учебного материала через промежуток времени a планируется проверка знаний. Если контроль проводится с помощью компьютерной системы, то следует учитывать риски недооценки и переоценки знаний α и β . Средняя длительность проверки знаний равна t_k .

Если обучаемый (оператор) сохраняет знания до момента проверки, то при проверке знаний он с вероятностью (риск недооценки) α получает неудовлетворительную оценку. При этом ему предлагается восстановить знания. Средняя длительность восстановления знаний равна t_e ($t_e > t_k$). С вероятностью $(1 - \alpha)$ он получает заслуженную положительную оценку. В этом случае проводится повторение учебного материала.

Если обучаемый (оператор) забывает учебный материал до момента проверки знаний, то с вероятностью $(1 - \beta)$ он получает заслуженную неудовлетворительную оценку. При этом ему предлагается восстановить знания. С вероятностью (риск переоценки) β он получает незаслуженную удовлетворительную оценку.

В произвольный момент времени обучаемый (оператор) может находиться в следующих состояниях: E_0 — состояние усвоения знаний, когда он может выполнить требуемые задачи; E_1 — состояние необнаруженного (скрытого) забывания; E_2 — проверка обучаемого (оператора), забывшего учебный материал; E_3 — состояние восстановления знаний, умений или навыков; E_4 — проверка знаний обучаемого (оператора), не забывшего учебный материал; E_5 — состояние скрытого забывания, необнаруженного в результате риска переоценки при контроле знаний.

Матрица вероятностей перехода из одного состояния знаний в другое имеет вид:

$$P(a) = \begin{pmatrix} 0 & F(a) & 0 & 0 & 1 - F(a) & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 - \beta & 0 & \beta \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 1 - \alpha & 0 & 0 & \alpha & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix}.$$

Граф переходов обучаемого (оператора) из одного состояния знаний в другое приведен на рис. 4.

Динамика знаний обучаемого (оператора) описывается при этом полумарковским процессом. Коэффициент профессиональной готовности обучаемого равен [4]

$$K(a) = \frac{\int_0^a [1 - F(t)] dt}{(a + t_k) \left(1 + \frac{\beta F(a)}{1 - \beta}\right) + t_e (1 - \beta) [F(a) + \alpha (1 - F(a))]} \quad (7)$$

Если время забывания имеет экспоненциальное распределение $\text{Exp}(\lambda)$, то из соотношения (7) следует:

$$K(a) = \frac{(1 - \beta)(1 - e^{-\lambda a})}{\lambda \{ (a + t_k)(1 - \beta e^{-\lambda a}) + \lambda t_e (1 - \beta)^2 [1 - (1 - \alpha)e^{-\lambda a}] \}} \quad (8)$$

Оптимальный интервал между проверками и повторениями знаний можно определить по максимуму коэффициента профессиональной готовности (7) или (8).

Пример 3. Применим рассмотренную стратегию к операторам сложной и ответственной системы. Для них определены характеристики динамики знаний: интенсивность забывания $\lambda = 0,05 \dots 0,40$ 1/мес, математическое ожидание времени восстановления знаний $t_e = 3$ ч, математическое ожидание времени контроля знаний $t_k = 0,5$ ч.

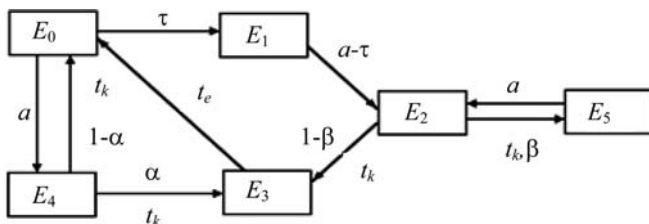


Рис. 4. Граф переходов обучаемого (оператора) из одного состояния в другое при контроле знаний с рисками недооценки и переоценки знаний

Характеристики	λ , 1/мес					
	0,05	0,10	0,15	0,20	0,30	0,40
$K_{\max} = K(a_0)$	0,98	0,97	0,96	0,955	0,945	0,94
a_0 , ч	160	100	80	70	60	50
$K(a = 600 \text{ ч})$	0,95	0,91	0,87	0,83	0,76	0,71

Предположим, что контроль знаний проводится с помощью компьютерной обучающей системы с риском недооценки $\alpha = 0,35$ и переоценки знаний $\beta = 0,40$. Необходимо определить оптимальные интервалы между проверками и повторениями знаний. Практический интерес представляет также решение следующей задачи: как влияют отклонения от оптимального значения интервала между проверками и повторениями знаний на коэффициент профессиональной готовности обучаемого (оператора).

Решение. В таблице приведены оптимальные значения a_0 и максимумы коэффициента профессиональной готовности $K(a)$, а также значения коэффициента профессиональной готовности $K(a)$ при проведении контроля знаний через интервал времени $a = 600$ ч.

Из таблицы следует, что оптимальный интервал a_0 для операторов с интенсивностью забывания $\lambda = 0,05$ 1/мес составляет 160 ч (примерно 7 дней), а для оператора с $\lambda = 0,40$ 1/мес $a_0 = 50$ ч (примерно 2 дня).

Если контроль знаний оператора с $\lambda = 0,40$ 1/мес вместо оптимального интервала $a_0 = 50$ ч проводится через интервал $a = 600$ ч, то коэффициент профессиональной готовности снижается с 0,94 до 0,71. Этот пример показывает влияние отклонений от оптимального значения интервала между проверками и повторениями знаний на коэффициент профессиональной готовности обучаемого (оператора).

Заключение

В работе предложены две стратегии обеспечения профессиональной готовности обучаемых с учетом и без учета рисков недооценки и переоценки знаний при компьютерном контроле знаний. Исходные данные для них: интенсивность забывания; среднее время контроля знаний, среднее время повторения учебного материала или восстановления знаний. Критерии оптимальности: максимум коэффициента профессиональной готовности обучаемого и минимум средних удельных затрат.

Список литературы

1. Человеческий фактор (Handbook of Human Factors) / Под ред. Сальвенди Г. М.: Мир, 1991. 599 с.
2. Birolini A. Qualitat und Zuverlassigkeit technischer Systeme. Theorie, Praxis, Management. — Berlin, Heidelberg, New York u.a.: Springer—Verlag, 1991. 507 s.
3. E-Learning deutscher Universitäten. Fallbeispiele aus der Hochschulpraxis // Kurzinformation Bau und Technik, HIS Hochschulinformationssystem. 2005. В 4. 97 s.
4. Свиридов А. П. Введение в статистическую теорию обучения и контроля знаний. Ч. 2. Элементы статистической динамики знаний. М.: Изд-во МЭИ, 1974. 152 с.
5. Свиридов А. П. Основы статистической теории обучения и контроля знаний. М.: Высшая школа, 1981. 262 с.
6. Sflsse R., Sviridov A. P. Statistische Kenntnis-Dynamik. — Pirmenau: Wissenschaftsverlag, 1998. 256 s.
7. Свиридов А. П. Статистическая теория обучения. М.: Изд-во РГСУ, 2009. 577 с.
8. Ушинский К. Д. Человек как предмет воспитания. Опыт педагогической антропологии. М.: Учпедгиз, 1950.
9. Тальзина Н. Ф. Управление процессом усвоения знаний (психологические основы), 2-е изд. М.: Изд-во МГУ, 1984. 344 с.

CONTENTS

Struchenkov V. I. Mathematical Models and Optimization Techniques in Systems Design of Roads 2

The problem of optimizing the position of the new road in a given region of variation. We propose new mathematical models and algorithms for the design of the vertical alignment for the specified variants of the plan there route. This method provides information problem to a similar problem before designing the longitudinal profile of the new railway. Design of the longitudinal profile is regarded as a multi-step design process in conjunction with other project tasks.

Keywords: route, horizontal and vertical alignment, nonlinear programming, objective function, reduced gradient

Musikhina E. A., Darmanskaya T. A. Research and the Building of a Relief Model of Triangulation. 8

Study the fractal properties of the relief of the Irkutsk region have shown that the natural system is unstable and is not in the current balance. A synthesis of information technologies of Golden Software Surfer and Auto-desk Civil 3D in building the model of the surface area in order to identify the characteristics of the relief of the proposed site of construction is used. Rendering TIN surfaces to simulate a complicated dissected relief is substantiated.

Keywords: fractal properties of the terrain, synergy, digital model, dissected relief, data array, triangulation

Chuvasheva E. S., Chuvashev S. N. Optimization of High Speed Vehicles with the Complex Mathematical Model 12

The previously described mathematical model is used which accounts for versatile physical processes in the main units of a high speed air breathing vehicle. Examples of different mission plans and apparatuses of various scales and design are analyzed. Bottlenecks are revealed, and ways of optimization are suggested. The created application program can be used by the developers of new high speed apparatuses on the stages of concept and draft design.

Keywords: high speed vehicles, heat exchange, flight dynamics, complex mathematical model

Sherstyukov O. N., Ryabchenko E. Yu., Martynchuk S. L. Radio Modem of the LPD Band for the Telemetry. . . 17

The scheme of radio modem of the LPD band for the telemetry systems development has been proposed. The modem can be used in the conditions where radio signals with frequency of 2.4 GHz have large attenuation and such wireless technologies as Wi-Fi, Bluetooth and ZigBee (802.15.4) cannot be used. The described radio modem is based on the transceiver SX1233. A channel layer protocol has been developed and a mathematical model of the protocol is described.

Keywords: radio modem, radio communication, LPD, wireless telemetry, channel layer protocol

Ivanova E. M. Computer Comparative Performance Evaluation 22

The article discusses criteria for comparative assessment of different computers. Special attention is paid to such criterion as performance. There is given analysis of the different methods advantages and disadvantages for performance measurement.

Keywords: computing, performance, peak performance, maximum performance, rated capacity, efficiency, processing power, MIPS-rate, FLOPS-rate, LINPACK, SPEC

Bobkov S. G., Chibisov P. A. Improving of Quality of High-performance Microprocessors Testing by Counter-Testing Methods with Analysis of the Functional Coverage of the Selected User Control Tasks 26

Current microprocessor testing and verification methods do not allow to avoid the numerous mistakes in the project. The proposed procedure for high-performance microprocessors is a testing quality improvement by means of "counter-testing" methodology integration in the testing route with analysis of the full functional coverage for the selected tasks for critical applications.

Keywords: counter-testing methodology, microprocessor verification, verification plan, coverage-directed test generation, test completeness metrics, tape-out readiness

Agayev B. S., Pashayev F. H. Method of Estimating Speech Quality in the Corporate VoIP Networks 34

There are considered the questions of an estimation of quality of transferred speech in package switching network in the article. There is offered hybrid method which simplifies process of an estimation of quality and reduces any uncertainty in the results, in comparison with existing methods.

Keywords: speech signals, an IP-telephony, methods of an estimation of quality, a delay of delivery, loss of packages, a hybrid method

Parshikova E. A. Eventual Approaches to the Design Automation of Decision Support Systems for Making Managerial Decisions 40

In present article the theoretical issues related to decision support systems are examined, the overview of current trends in the design of decision support systems is made and potential approaches to design automation of DSS for making managerial decisions are considered.

Keywords: decision support systems, design of decision support systems, design automation of decision support systems

<i>Bulyonkov M. A., Filatkina N. N. Situational Analysis in the System for Transportation Forecast MIX-PROSTOR</i>	45
------------------------------------------------------------------------------------------------------------------------------	----

We consider economical research automation system and its application to the problem of transportation system development forecast. The process of description, calculation and analysis of forecast variants is described. The concept of situation room is realized by the method of variant combination that was introduced in the paper. It presents a user interface for decision makers.

Keywords: research automation system, situation room, adoptive user interfaces, transportation network forecast

<i>Sadykov S. S., Bulanova Yu. A. Algorithm of Allocation of Area of a Cyst on Low-Contrast Mammograms</i>	53
----------------------------------------------------------------------------------------------------------------------	----

The article analyzes the known preprocessing algorithms applied to low-contrast mammographic images. The algorithm of histogram transformation mammograms, taking into account the average contrast image. An algorithm for selection of tumor type "breast cyst" on the mammogram.

Keywords: mammogram, cyst, preprocessing image, histogram transformation, average contrast, texture segmentation, selection of tumor, edge enhancement

<i>Solodovnichenko M. A., Zamyatin A. B. Algorithm of Forest Fire Spread with Rothermel Model and Cellular Automata</i>	58
-----------------------------------------------------------------------------------------------------------------------------------	----

Process of forest fire spreading according to model of Rothermel on the basis of the cellular automata allowing predicting the territories of the forest in fire and burnt out at a given time, on a known landscape of district, speed and a direction of a wind and an arrangement of the seat is considered. The algorithm for program realization is resulted, and also one of test data preparation variants is considered.

Results of comparative tests of the developed software with known realization of Rothermel model — FARSITE, comparing dependences of model on direction and strength of wind change, and also comparing results of modeling in territory with an identical landscape are resulted.

Keywords: forest fire spread modeling, cellular automata, Rothermel model

<i>Bekirova L. R. Questions on Minimization of Effect of Variability of Atmospheric Aerosol on Information Capability of Satellite Multispectral Images</i>	64
-----------------------------------------------------------------------------------------------------------------------------------------------------------------------	----

On the basis of known factual data it is shown, that multi-modal functions of distribution of aerosol leads to significant deviations from exponent form of function of dependence of aerosol optical depth on wavelength.

The mathematical formula for calculation of optimal functional dependence of aerosol optical depth on wavelength is given, which means that the suggested reach the minimal value.

It is noted, that minimization of integral entropy criteria allows to decrease the effect of atmospheric aerosol on multispectral satellite images.

Keywords: multispectral images, atmosphere, aerosol, informativeness, entropy criterion

<i>Zaleshchansky B. D., Sviridov A. P., Shalobina O. A., Shalobina E. A. Probabilistic and Statistical Strategy of Ensuring Quality of Preparation of the Personnel of Socio-Technological Systems</i>	67
------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	----

Periodical strategies of vocational training quality management are examined with consideration of forgetting infectivity and economical parameters (training costs, material losses due to personal errors). Strategies are based optimal planning of repetitions and preventive knowledge renewal. Required professional preparation of personnel is ensured by means of examination and consequent training quality management.

Keywords: strategies, personnel training, forgetting infectivity

Адрес редакции:

107076, Москва, Стромьинский пер., 4

Телефон редакции журнала (499) 269-5510

E-mail: it@novtex.ru

Дизайнер *Т.Н. Погорелова*. Технический редактор *Е. В. Конова*.

Корректор *Т.В. Пчелкина*.

Сдано в набор 06.06.2013. Подписано в печать 18.07.2013. Формат 60×88 1/8. Бумага офсетная.

Усл. печ. л. 8,86. Заказ ИТ813. Цена договорная.

Журнал зарегистрирован в Министерстве Российской Федерации по делам печати, телерадиовещания и средств массовых коммуникаций.

Свидетельство о регистрации ПИ № 77-15565 от 02 июня 2003 г.

Оригинал-макет ООО "Авансед солюшнз". Отпечатано в ООО "Авансед солюшнз".

105120, г. Москва, ул. Нижняя Сыромятническая, д. 5/7, стр. 2, офис 2.