

ISSN 1818-0442

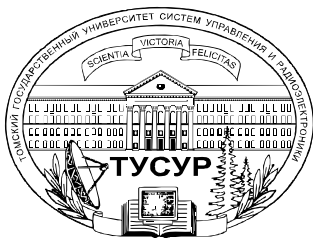
Доклады ТУСУРа. №1(27), 2013

ДОКЛАДЫ

Томского государственного университета
систем управления и радиоэлектроники

1(27) • 2013





Министерство образования и науки Российской Федерации

ДОКЛАДЫ ТОМСКОГО ГОСУДАРСТВЕННОГО УНИВЕРСИТЕТА СИСТЕМ УПРАВЛЕНИЯ И РАДИОЭЛЕКТРОНИКИ

1(27) • март 2013

Периодический научный журнал

Выходит 4 раза в год

Основан в 1997 г.

ISSN 1818-0442

Гл. редактор:
Ю.А. Шурыгин

Зам. гл. редактора:
А.В. Кобзев
А.А. Шелупанов

Редакционный совет:
Л.А. Боков
А.Г. Буймов
Ю.П. Ехлаков
А.М. Кориков
Е.М. Окс
И.Н. Пустынский
В.Н. Татарин
С.М. Шандаров
Г.С. Шарыгин

Ответственный секретарь:
В.Н. Масленников

Адрес редакции:
634050, г. Томск,
пр. Ленина, 40, ТУСУР,
тел. (382-2) 51-22-43

Свидетельство
о регистрации МНС РФ
1027000867068
от 13 октября 2004 г.

Подписной индекс 20648
в каталоге Агентства
«Роспечать»: газеты и журналы

**Издательство
Томского государственного
университета систем управления
и радиоэлектроники**
634050, Томск, пр. Ленина, 40,
тел. (3822) 51-21-21

Оригинал-макет выпуска подготовлен
и отпечатан тираж
в ООО «Издательство В-Спектр»
Техн. редактор В.М. Бочкарева
Корректор В.Г. Лихачева

Подписано в печать 14.03.2013.
Формат 60×84 1/8.
Усл. печ. л. 10
Тираж 500. Заказ 9.

Содержание

ЭЛЕКТРОНИКА, ИЗМЕРИТЕЛЬНАЯ ТЕХНИКА, РАДИОТЕХНИКА И СВЯЗЬ

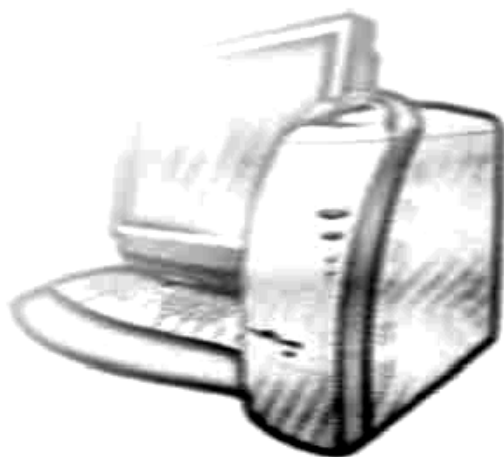
- Тюньков А.В., Юшков Ю.Г.**
Масс-зарядовый состав плазмы, генерируемой в разрядной системе с инжекцией электронов и дополнительным распыляемым электродом..... 5
- Зенин А.А., Климов А.С.**
Электронно-лучевая пайка алуминооксидной керамики с металлом с применением форвакуумного плазменного источника электронов..... 10
- Осипов А.В., Шурыгин Ю.А., Шиняков Ю.А.,
Отто А.И., Черная М.М.**
Сопоставительный анализ энергетической эффективности преобразования энергии солнечной батареи преобразователями постоянного напряжения 14
- Аржанов В.В., Мишин В.Н., Ракитин Г.А., Аржанов К.В.**
Управление позиционными электроприводами блока измерения освещенности для термобарокамеры 20
- Маморцев С.В., Михальченко Г.Я.**
Технология мягкой коммутации транзисторов в преобразователе постоянного напряжения в постоянное 24
- Якушевич Г.Н.**
Алгоритм расчета транзисторного широкополосного усилителя с двухполюсниками комбинированной обратной связи 28

УПРАВЛЕНИЕ, ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА И ИНФОРМАТИКА

- Щеголева Н.Л.**
Разработка программных средств для моделирования биометрических систем распознавания человека по изображению лица 37
- Катаев М.Ю., Широков Л.В.**
Методика определения жестов руки, наблюдаемых с помощью видеокамеры 45
- Будков В.Ю.**
Методы и программные средства обработки мультимедийных данных при сопровождении распределенных совещаний..... 50
- Силич М.П., Вахитов А.Р., Прудников И.И.**
Информационная система анализа энергетической эффективности территориальных образований 55
- Галицкая Л.В., Голубева А.А., Гриценко Ю.Б.,
Жуковский О.И., Пудуль И.Ю.**
Идентификация параметров объектов электронного генерального плана металлургического предприятия 60
- Зубов М.В., Пустыгин А.Н., Старцев Е.В.**
Применение универсальных промежуточных представлений для статического анализа исходного программного кода 64
- Ефимов Б.И., Файзуллин Р.Т.**
Вероятность принятия ложного решения под воздействием угроз информационной безопасности в системах принятия решений с привлечением экспертов 69
- Зорин Ю.А.**
Интерпретатор языка построения генераторов тестовых заданий на основе деревьев И/ИЛИ 75

Шемяков А.О.	
Причины повышения уязвимости и снижения стойкости функций безопасности автоматизированных систем вуза.....	80
Молчанов А.Л., Файзуллин Р.Т.	
Изучение свойств эффективности и робастности работы мультиплексной системы.....	83
Ле Хоай, Тузовский А.Ф.	
Поиск в семантических электронных библиотеках	87
Ефремов А.А.	
Новые операции над нечеткими числами и интервалами	95
Ехлаков Ю.П., Гриценко Ю.Б., Жуковский О.И.	
Автоматизация технологий формирования и мониторинга электронного генерального плана инженерной инфраструктуры.....	100
Шилин А.А., Букреев В.Г., Гладышева Е.И.	
Моделирование нелинейных систем на FBD-блоках с ограниченным базисом	107
Грачев В.В., Силич В.А., Силич М.П.	
Методология проектирования ситуационных центров принятия решений	114
Лазарев И.В.	
Математические модели планирования технологических регламентов реконструкции и ремонта инженерных сетей.....	119
Полищук В.Ю.	
Программный комплекс имитационного моделирования динамики полей термокарстовых озер в зонах вечной мерзлоты.....	125
Гриценко Ю.Б., Жуковский О.И., Лазарев И.В., Милихин М.М., Сенченко П.В.	
Интерфейс взаимодействия геоинформационной технологии ведения электронного генерального плана со сторонними программными системами	129
Ходашинский И.А., Горбунов И.В., Синьков Д.С.	
Алгоритмы генерации структур двухкритериальных Парето-оптимальных нечетких аппроксиматоров	135

**ЭЛЕКТРОНИКА, ИЗМЕРИТЕЛЬНАЯ ТЕХНИКА,
РАДИОТЕХНИКА И СВЯЗЬ**



УДК 533.9.07: 533.9.082

А.В. Тюньков, Ю.Г. Юшков

Масс-зарядовый состав плазмы, генерируемой в разрядной системе с инъекцией электронов и дополнительным распыляемым электродом

На основе времяпролетной методики проведены исследования масс-зарядового состава плазмы, генерируемой разрядной системой с инъекцией электронов и дополнительным распыляемым электродом. Такая электродная конфигурация разрядного устройства обеспечивает возможность создания композиционных покрытий различного функционального назначения. Показано, что доля ионов металла в объемной плазме не превышает единиц процентов.

Ключевые слова: разряд низкого давления, инъекция электронов, распыляемый электрод, времяпролетный масс-спектрометр, масс-зарядовый состав плазмы.

Многофункциональный генератор объемной газоразрядной плазмы с секционированным распыляемым электродом на основе двухступенчатой разрядной системы с инъекцией электронов из дугового контрагированного разряда успешно используется для генерации однородной газоразрядной плазмы инертных газов во время очистки и активации поверхности образцов перед нанесением покрытий [1, 2], а также для непосредственного синтеза композиционных покрытий металлов и их соединений [3, 4]. Конструкция и принцип работы систем с инъекцией электронов позволяют длительное время работать в среде углеводородсодержащих газов и в активной среде кислорода, что в свою очередь обеспечивает условия для формирования алмазоподобных [5, 6] и оксидных покрытий [7]. Отсутствуют условия образования капельной фракции, что способствует формированию однородных плотных по толщине пленок.

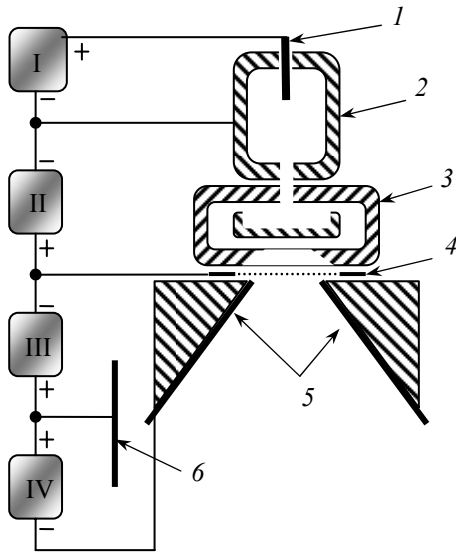
В настоящей статье представлены результаты исследования масс-зарядового состава плазмы генерируемой разрядной системой с инъекцией электронов и дополнительным распыляемым электродом.

Методика и техника эксперимента. Функционально генератор плазмы состоит из трех разрядных ступеней: эмиттера электронов, основного разряда и распылительного узла (рис. 1). При этом первые две ступени обеспечивают генерацию газовой плазмы в объеме вакуумной камеры, включая область вблизи распыляемых мишеней. Плазма генерируется в результате инъекции в объем пучка электронов с регулируемым током и энергией.

Эмиттерный разряд функционирует между полым катодом 2 с малой выходной апертурой и сеточным анодом 4, выполненным из нержавеющей стали прозрачностью 60%. Сеточный анод эмиттерного разряда в то же время является катодом основного разряда. Анодом основного разряда являются заземленные стенки вакуумной камеры 6.

При подаче рабочего газа (аргона) в катодную полость эмиттера разряд инициируется высоко-точным (30 А) газовым разрядом при подаче высоковольтного импульса (5 кВ, 35 мкс) между поджигающим электродом 1 и полым катодом 2. При подаче постоянного напряжения между электродами 2 и 4 возникает дуговой разряд с катодными пятнами. Для стабильного горения разряда в области малых токов (от 3 А и выше) полость катода выполнена из магния.

Кроме этого, между полым катодом 2 и сеточным анодом 4 установлен водоохлаждаемый фильтрующий электрод 3. Этот электрод обеспечивает необходимый перепад давления между эмиттерной и основной разрядными ступенями генератора плазмы. Он также препятствует выходу продуктов эрозии материала стенок полого катода 2 в объем вакуумной камеры. Электроны из плазмы эмиттерного разряда инжектируются в область основного разряда (вторая разрядная ступень), где производят ионизацию рабочего газа. Регулируя напряжение основного разряда (величину катодного падения потенциала), можно управлять энергией инжектируемых электронов. Изменяя ток эмиттерного разряда при постоянном напряжении горения основного разряда, можно менять количество инжектируемых электронов.



Неоднородность распределения концентрации объемной плазмы в камере в поперечном направлении на расстоянии 30 см составляет порядка $\pm 25\%$ от среднего значения при токе эмиттерного разряда 20 А. Потенциал плазмы основного разряда составляет порядка 5–15 В относительно стенок заземленной вакуумной камеры.

Рис. 1. Электродная схема распылительной системы на основе разряда с инъекцией электронов: 1 – поджигающий электрод; 2 – катод эмиттерного разряда; 3 – промежуточный фильтрующий электрод; 4 – катод основного (анод эмиттерного) разряда; 5 – распылительный узел; 6 – анод основного разряда. Источники питания: I – иницирующей системы, II – эмиттерного разряда; III – основного разряда; IV – мишеней

Распылительный узел 5 представляет собой 6 электрически изолированных друг от друга секций (мишеней), расположенных симметрично вокруг оси разрядной системы. Активная площадь распыления секций узла составляет 650 см^2 . Форма секционированного распыляемого электрода сделана в форме усеченного конуса, а сами мишени обращены в сторону подложки под большим углом. Секционирование мишени позволяет использовать несколько различных материалов и регулировать напряжение смещения на каждой из них, что в свою очередь позволяет контролировать количество распыленного материала. Разделение электрического питания позволяет существенно расширить функциональные возможности генератора. Так для очистки и активации поверхности подложки перед нанесением покрытий отрицательное электрическое смещение на мишени не подается. В этом случае генератор работает в режиме газовой плазмы, подвергая поверхность подложки интенсивной бомбардировке высокоэнергетичными ионами аргона. При подаче электрического смещения на мишени генератор работает в режиме источника газовой плазмы и потока материала мишени, позволяя проводить процессы нанесения композиционных покрытий.

Исследования масс-зарядового состава объемной плазмы генератора проводились с использованием времяпролетной методики, заключающейся, как известно, в разделении по временам пролета различных масс-зарядовых компонентов ускоренного пучка ионов [8], извлеченных из плазмы. Преимущество этой методики в отличие от оптической спектроскопии и магнитной сепарации заключается в возможности одномоментного измерения долей масс-зарядового состава пучка. Времяпролетный спектрометр имеет удовлетворительную разрешающую способность (более 10), широкий диапазон измерения масс ионов (1–300 а.е.м.) и относительно высокую чувствительность [8]. Необходимо отметить, что работа источника осуществлялась при пониженном давлении газа (не более $1 \cdot 10^{-1}$ Па). Оценки показывают, что при таком давлении длина свободного пробега составляет порядка одного метра. Следовательно, рассеяние пучка ионов на атомах остаточного газа и потери энергии при транспортировке пучка на расстояние порядка метра незначительны и использование методики вполне оправдано.

Схема измерений масс-зарядового состава плазмы на основе времяпролетного спектрометра представлена на рис. 2.

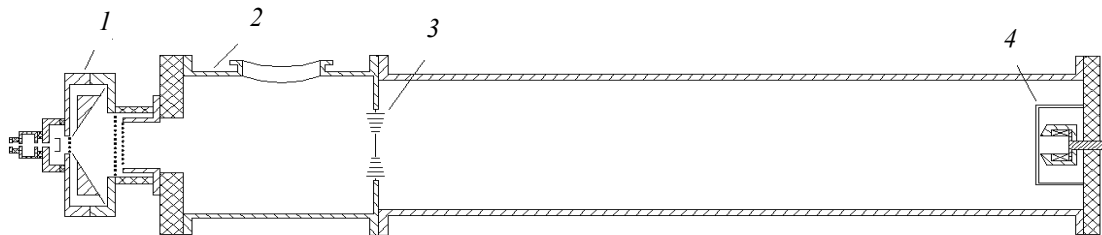


Рис. 2. Схема измерений масс-зарядового состава плазмы на основе времяпролетного спектрометра: 1 – генератор плазмы; 2 – вакуумная камера; 3 – затвор спектрометра; 4 – цилиндр Фарадея

Времяпролетный спектрометр является модификацией устройства [9], успешно зарекомендовавшего себя для измерений масс-зарядового состава пучков ионных источников [10, 11].

Величина q/M для каждого масс-зарядового компонента ионного пучка определялась из выражения (1) по величине временной задержки t относительно фронта импульса напряжения на затворе спектрометра:

$$t = \frac{L\sqrt{\frac{q}{M}}}{2eU_{\text{уск}}}, \quad (1)$$

где t – промежуток времени между приложением к затвору спектрометра отклоняющего импульса и сигналом в цепи цилиндра Фарадея; e – заряд электрона; $U_{\text{уск}}$ – ускоряющее напряжение; L – база спектрометра (пролетное расстояние), $L=1,2$ м.

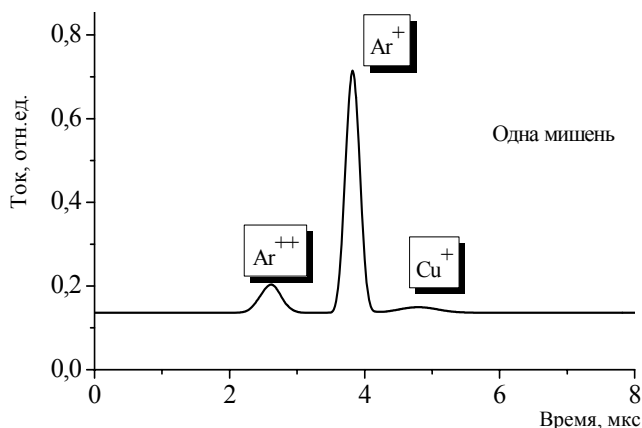
Долевое отношение компонентов ионного пучка оценивалось по величине амплитуды импульса тока цилиндра Фарадея. При этом учитывалось, что регистрируемый ток многозарядных ионов в кратность зарядов раз превышает ток частиц данной зарядности.

В экспериментах по измерению масс-зарядового состава генератор плазмы находился под высоким относительно земли потенциалом. Поэтому электропитание разрядов и мишеней осуществлялось через разделительные трансформаторы. Функционирование эмиттерного разряда осуществлялось с использованием импульсной схемы, формирующей импульс тока амплитудой 3–30 А и длительностью 400 мкс. Постоянное электрическое смещение мишеней поддерживалось на уровне 500 В, что достаточно для эффективного распыления мишеней без образования на их поверхности катодных пятен. Напряжение основного разряда изменялось в диапазоне 100–200 В.

При исследовании масс-зарядового состава объемной плазмы использовались 1 либо 3 медные распыляемые мишени. В качестве рабочего газа – аргон. Выбор распыляемого металла и рода газа обусловлен, во-первых, относительно высоким коэффициентом распыления меди (например, для энергии ионов аргона 500–600 эВ, $S_{\text{Cu}} = 3$ атома на ион). Во-вторых, масса ионов меди и аргона сильно отличается, что значительно упрощает идентификацию зарядовых составляющих этих элементов в измеряемом спектре.

Результаты эксперимента. Эксперименты показали, что при использовании одной мишени доля ионов меди в объемной плазме составляет 2% (рис. 3), а для 3 мишеней – 4% (рис. 4).

Рис. 3. Ток цилиндра Фарадея времяпролетного спектрометра при использовании 1 мишени. $U_{\text{осн}} = 200$ В, $U_{\text{миш}} = 500$ В, $U_{\text{уск}} = 20$ кВ, $I_{\text{эмит}} = 30$ А, $U_{\text{откл}} = 3$ кВ, $\nu = 5$ Гц, $p = 3,5 \cdot 10^{-2}$ Па, Ar



Расчет процентного содержания ионов материала распыляемой мишени в плазме. Доля ионов металла (меди), образованных в объемной плазме аргона, может быть оценена из выражения (2):

$$\frac{n_{i\text{Cu}}}{n_{i\text{Ar}}} \times 100\% = \frac{n_{\text{ат. Cu}} \times \sigma_{i\text{Cu}}}{n_{\text{ат. Ar}} \times \sigma_{i\text{Ar}}} \times 100\%, \quad (2)$$

где $n_{i\text{Cu}}$ – концентрация ионов меди; $n_{i\text{Ar}}$ – концентрация ионов аргона; $n_{\text{ат. Cu}}$ – концентрация атомов меди; $n_{\text{ат. Ar}}$ – концентрация атомов аргона; $\sigma_{i\text{Cu}}$ – сечение ионизации атома меди; $\sigma_{i\text{Ar}}$ – сечение ионизации атома аргона.

Значения сечения ионизации атомов меди и аргона при энергии электронов $E_e=50-150$ эВ следующие: $\sigma_{i\text{Cu}} = 4,1 \cdot 10^{-16}$ см², $\sigma_{i\text{Ar}} = 2,8 \cdot 10^{-16}$ см².

Концентрация атомов аргона определяется из соотношения (3):

$$p = n_{\text{ат. Ar}} kT, \quad (3)$$

где p – давление в вакуумной камере; k – постоянная Больцмана; T – температура газа.

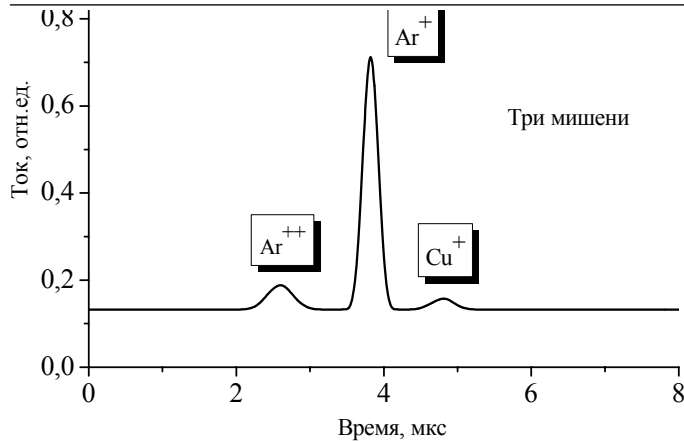


Рис. 4. Ток цилиндра Фарадея времяпролетного спектрометра при использовании 3 мишеней. $U_{\text{осн}} = 200$ В, $U_{\text{миш}} = 500$ В, $U_{\text{уск}} = 20$ кВ, $I_{\text{эмит}} = 30$ А, $U_{\text{откл}} = 3$ кВ, $v = 5$ Гц, $p = 3,5 \cdot 10^{-2}$ Па, Ar

тельного электрода 25 см, найдем плотность потока распыленных атомов меди из (5):

$$j_{\text{ат. Cu}} = \frac{N_{\text{ат. Cu}}}{S_{\text{потока}}}, \quad (5)$$

где $N_{\text{ат. Cu}}$ — число распыленных атомов меди; $S_{\text{потока}}$ — площадь потока атомов меди.

Для $N_{\text{ат. Cu}} = 3 \cdot 10^{19}$ и $S_{\text{потока}} = 500$ см² значение плотности потока составило $j_{\text{ат. Cu}} = 6 \cdot 10^{16}$ шт/см². Скорость атомов меди, эмитируемых в результате бомбардировки, определяется из (6):

$$V_{\text{ат. Cu}} = 1,38 \times 10^6 \sqrt{\frac{E_i}{A}}, \quad (6)$$

где E_i — энергия ионов, эмитируемых мишенью; A — атомная масса.

Энергия ионов, приобретаемая ионами мишени в результате ионной бомбардировки, при выходе из металла составляет приблизительно 5 эВ. Тогда значение скорости атомов меди составляет $V_{\text{ат. Cu}} = 0,4 \cdot 10^6$ см/с.

Подставляя значения $V_{\text{ат. Cu}}$ и $j_{\text{ат. Cu}}$ в выражение (4), получим, что концентрация атомов меди составляет $n_{\text{ат. Cu}} = 1,5 \cdot 10^{11}$ см⁻³.

Проведенные оценки свидетельствуют о том, что для данных условий и параметров разряда доля ионов металла в объемной плазме не превышает единиц процентов (1–5%).

Заключение. Экспериментальные результаты и проведенные оценки показали, что в разрядной системе с инжекцией электронов и дополнительным распыляемым электродом доля ионов металла в плазме объемного разряда достаточно мала и составляет единицы процентов. Следовательно, формирование покрытий с использованием такой разрядной структуры осуществляется за счет адсорбции на поверхности подложки нейтральных атомов и дальнейших плазмохимических реакций атомов распыляемой мишени и химически активных газов. Данный факт несколько сужает области применения данного источника, например в процессах плазменно-иммерсионной ионной имплантации металлов. Однако использование плазменного генератора для осаждения тонких оптически прозрачных композиционных покрытий как на непроводящие мишени (стекло, керамика и т.д.), так и на проводящие, где не требуется внедрения материала мишени в подложку, выглядит достаточно привлекательно.

Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований (грант № 12-08-31579-мол_а).

Литература

1. Генератор объемной плазмы на основе разряда с плазменным катодом / А.В. Визирь, Е.М. Окс, М.В. Шандриков, Г.Ю. Юшков // ПТЭ. – 2003. – № 3. – С. 108–111.
2. Генерация объемной плазмы на основе сильноточного газового разряда с внешней инжекцией электронов / А.В. Визирь, Е.М. Окс, М.В. Шандриков, Г.Ю. Юшков // Прикладная физика. – 2004. – № 6. – С. 115–119.

Для значений $p = 1 \cdot 10^{-1}$ Па, $k = 1,38 \cdot 10^{-23}$ Дж/К и $T = 300$ К концентрация атомов аргона в камере составляет $n_{\text{ат. Ar}} = 2,5 \cdot 10^{13}$ см⁻³.

Концентрация атомов меди определяется по формуле (4):

$$n_{\text{ат. Cu}} = \frac{j_{\text{ат. Cu}}}{V_{\text{ат. Cu}}}, \quad (4)$$

где $j_{\text{ат. Cu}}$ — плотность потока распыленных атомов меди; $V_{\text{ат. Cu}}$ — скорость распыленных атомов меди.

Учитывая, что при энергии ионов 500 эВ коэффициент распыления для меди $S_{\text{Cu}} = 3$ атома на ион, ток на мишени 1,5 А, площадь мишеней 650 см², то количество распыленных атомов меди составляет $3 \cdot 10^{19}$. Зная диаметр распылительного электрода 25 см, найдем плотность потока распыленных атомов меди из (5):

3. Генератор газометаллической плазмы на основе разряда с инжекцией электронов / А.В. Визирь, Е.М. Окс, А.В. Тюньков, М.В. Шандриков // ПТЭ. – 2011. – № 2. – С. 73–77.
4. Формирование функциональных оксидных композиционных покрытий на основе разрядной системы с инжекцией электронов / А.В. Визирь, К.П. Савкин, М.В. Шандриков, Г.Ю. Юшков // Доклады Том. гос. ун-та систем управления и радиоэлектроники. – 2011. – № 1 (23). – С. 83–88.
5. Shandrikov M.V. Formation of hard amorphous hydrogenated carbon films on metal by CVD method in acetylene plasma / A.V. Vizir, M.V. Shandrikov // Известия вузов. Физика. – 2007. – № 9. Приложение. – С. 460–463.
6. Shandrikov M.V. Deposition and Investigation of Amorphous Hydrogenated Carbon Films from Acetylene Plasma // 9th International Conference on Modification of Materials with Particle Beams and Plasma Flows: Proceedings. – Tomsk: Publishing House of the IOA SB RAS, 2008. – P. 568–571.
7. Discharge System with Electron Injection for Formation of Composite Metal-Oxide Hydrophobic Coatings / A.V. Tyunkov, A.V. Vizir, M.V. Shandrikov et al. // 10th International Conference on Modification of Materials with Particle Beams and Plasma Flows: Proceedings. – Tomsk: Publishing House of the IOA SB RAS, 2010. – P. 594–597.
8. Brown I.G. A new method for ion charge state analysis / I.G. Brown, J.C. Kelly // Applied Physics. – 1988. – Vol. 63. – P. 254–258.
9. Improved time-of-flight ion charge state diagnostic / I.G. Brown, J.E. Galvin, R.A. MacGill, R.T. Wright // Review of Scientific Instruments. – 1987. – Vol. 58. – P. 1589–1592.
10. Measurement of total ion flux from vacuum arc cathode spot / A. Anders, E.M. Oks, G.Yu. Yushkov et al. // IEEE Transaction on Plasma. Science. – 2005. Vol. 33. – № 5 (1). – P. 1532–1536.
11. Boron ion source based on planar magnetron discharge in self-sputtering mode / V.I. Gushenets, A. Hershcovitch, T.V. Kulevoy et al. // Review of Scientific Instruments. – 2010. – Vol. 81. – P. 02B303.

Тюньков Андрей Владимирович

Канд. техн. наук, доцент каф. физики ТУСУРа
Тел.: (382-2) 41-33-69
Эл. почта: andrew71@sibmail.com

Юшков Юрий Георгиевич

Канд. техн. наук, мл. науч. сотрудник лаборатории плазменной электроники НИЧ каф. физики ТУСУРа
Тел.: (382-2) 41-33-69
Эл. почта: YuYushkov@sibmail.com

Tyunkov A.V., Yushkov Yu.G.

Plasma mass-charge composition measurements generated in the discharge system with electron injection and additional sputtering electrode

Based on the time of flight technique mass - charge composition of the plasma generated in the discharge system with electron injection and additional sprayed electrode are investigated. This electrode configuration of the discharge unit enables the creation of composite coatings for various applications. It is shown that the metal ion contribution in the bulk plasma does not exceed a few percent.

Keywords: low pressure discharge, sectioned sputtering electrode, electron injection and TOF mass-spectrometer.

УДК 533.525

А.А. Зенин, А.С. Климов

Электронно-лучевая пайка алюмооксидной керамики с металлом с применением форвакуумного плазменного источника электронов

Представлены результаты исследований, направленных на создание механически прочных и вакуумно-плотных соединений алюмооксидной керамики с титаном посредством электронно-лучевой пайки. Компенсация зарядки электронным пучком поверхности электрически непроводящей керамики обеспечивалась применением для этих целей форвакуумного плазменного источника электронов. В качестве связующей прослойки использован алюминий.

Ключевые слова: электронно-лучевая пайка, форвакуум, плазменный электронный источник, диэлектрики, металлокерамические соединения.

Наиболее распространенным методом создания металлокерамических соединений является пайка [1]. Причем пайке, проводимой с использованием припоев, предшествует, как правило, операция металлизации керамики, заключающаяся в нанесении металлосодержащей пасты и ее последующем вжигании в керамику. Поскольку эта операция усложняет весь процесс, то предпринимаются попытки ее исключения [2]. Собственно пайка может осуществляться как нагревом в печи, так и использованием иных источников энергии, например лазерного луча [3]. Для предотвращения окисления металла и припоя пайку проводят в атмосфере инертного газа. Альтернативное решение состоит в проведении пайки в вакууме, причем в этом случае конкуренцию лазерному лучу может составить ускоренный электронный пучок. Компенсация зарядки электронным пучком поверхности электрически непроводящей керамики обеспечивается применением для этих целей форвакуумного плазменного источника электронов [4]. Об успешном опыте применения электронного источника такого типа для соединения керамических деталей сообщалось нами в [5]. В настоящей работе приводятся результаты исследований электронно-лучевого метода для создания соединений металла с керамикой.

Техника и методика эксперимента. Эксперименты проводились на установке, подробно описанной в [5] и схематически изображенной на рис. 1. Установка оснащена плазменным электронным источником, способным функционировать в так называемом форвакуумном диапазоне давлений (5–20 Па). Давление регулировалось напуском газа (воздух, гелий) непосредственно в рабочую камеру. Как показали ранее проведенные исследования [5], при генерации электронных пучков при столь высоких давлениях газа образующаяся в области транспортировки пучка плазма нивелирует процесс зарядки ускоренными электронами поверхности непроводящей керамической мишени. В этих условиях воздействие электронного пучка на керамику принципиально не отличается от электронно-лучевой обработки заземленных металлических образцов. В экспериментах плотность мощности электронного пучка достигала 2×10^4 Вт/см², что позволяло расплавлять и керамику, и металлы.

В качестве материалов для создания металлокерамических соединений использовались алюмооксидная керамика марки ВК94-1, керамика на основе талько-шамотной смеси, алюминий, титан ВТ-5, сталь 3, нержавеющая сталь. Соединяемые элементы представляли собой трубки. Образцы закреплялись в системе вращения с возможностью регулировки скорости в диапазоне от 1 до 30 оборотов в минуту, что позволяло осуществлять равномерный нагрев соединяемых деталей. Контроль качества соединения металла с керамикой осуществлялся измерением пределов прочности на излом и на разрыв с помощью разрывной машины 2167 Р-50, испытанием на устойчивость к термоциклированию и проверкой на герметичность, а также исследованием шлифов в растровом электронном микроскопе Hitachi TM-1000, снабженном энергодисперсионным микроанализатором QUANTAX50. Микроанализатор позволил определить относительное изменение концентрации элементов в поперечном направлении к границе металл–керамика.

Испытание на термоциклирование проводилось следующим образом – в нагретую до 350 °С муфельную печь помещались соединенные детали и выдерживались в печи 15 мин.

После чего образцы извлекались из печи и в течение 15 мин остывали при комнатной температуре, такой процесс повторялся 10 раз. Для испытания на герметичность зажатые в специальной оснастке трубчатые образцы помещались в емкость с водой, после чего во внутреннюю полость подавался газ (воздух) под давлением 6 атм. Наличие пузырьков воздуха в емкости с водой позволяло судить о герметичности шва.

Результаты экспериментов и их обсуждение. Для определения совместимых пар металл–керамика были проведены предварительные эксперименты, суть которых состояла в следующем. На очищенные плоские керамические пластины помещались кусочки металла, которые затем нагревались электронным пучком в вакууме до расплавления металла, либо керамики. Остуженные образцы извлекали из вакуумной камеры и проверяли наличие соединения. По итогам этих испытаний были выбраны пары: керамика ВК94-1 и алюминий, керамика на основе талько-шамотной смеси и титан ВТ-5.

Получение качественного соединения пары керамики ВК94-1 с алюминием оказалось затруднительным ввиду сложности образования паяного шва, несмотря на достаточно хорошую смачиваемость алюминием данного типа керамики. При достижении температуры плавления алюминий расплавлялся целиком и стекал с устройства крепления. На рис. 2 представлена пара ВК94-1 и алюминий. Данный образец был нагрет до температуры порядка 600 °С, и на нем уже заметны следы сильных деформаций.

Пайка керамики на основе талько-шамотной смеси с титаном проводилась в атмосфере гелия, поскольку данный металл активно взаимодействует (особенно в жидком состоянии) со всеми газами, составляющими атмосферу. Полученное соединение оказалось более прочным, однако паяный шов имел низкое качество ввиду наличия пор и несплавлений по всей его длине (рис. 3).

Для получения герметичного соединения было принято решение использовать алюминий в качестве связующего материала между алюмооксидной керамикой ВК94-1 и титаном ВТ-5 (рис. 1), содержащим, помимо титана, 5% алюминия. Успешность такого подхода было обусловлена хорошей смачиваемостью алюминием как керамики, так и титана.

Трубчатые образцы керамики и титана имели внешние диаметры 12 и 11 мм соответственно. Диаметр и толщина алюминиевой прокладки составляли 13 и 1 мм соответственно. Образцы керамики и титана имели внутренний диаметр 9 мм. Перед пайкой все исходные материалы обезжиривались и закреплялись в системе вращения. Дополнительной обработки поверхностей соединяемых деталей не проводилось. После вакуумирования рабочей камеры область стыка материалов прогревалась расфокусированным электронным пучком до температуры плавления алюминия, затем при неизменной температуре производилась выдержка в течение 3–5 мин, после чего за 10 мин мощность пучка плавно снижалась до нуля и спаянные детали остужались в вакууме в течение 10–15 мин. На рис. 4 представлены соединяемые образцы до и после пайки.

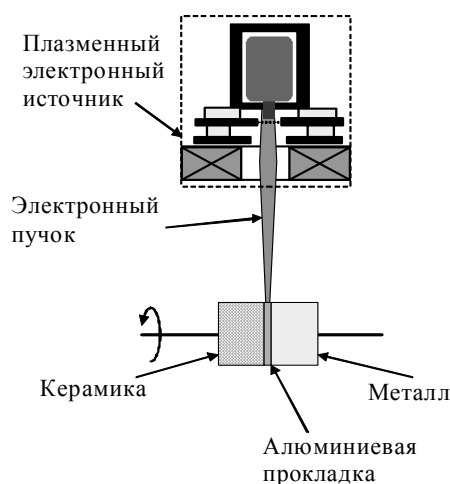


Рис. 1. Схематическое изображение экспериментальной установки



Рис. 2. Соединение керамики ВК94-1 и алюминия (слева – алюминий, справа керамика)



Рис. 3. Соединение керамики на основе талько-шамотной смеси с титаном ВТ-5

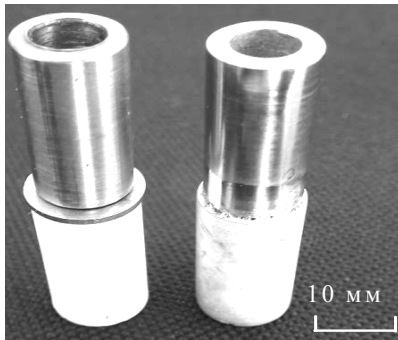


Рис. 4. Образцы до (слева) и после (справа) пайки: сверху титан ВТ-5; внизу керамика ВК94-1

Исследование полученных металлокерамических соединений методами растровой электронной микроскопии (рис. 5) и рентгеноспектрального микроанализа (рис. 6) выявили, что структура переходной зоны алюминий–керамика более резкая по сравнению с зоной алюминий–титан, что может свидетельствовать о частичном растворении титана в алюминии.

При испытаниях на термоциклирование четыре из шести испытываемых образцов сохранили свою герметичность, а в двух образцах появилась течь. Прочность на разрыв полученного соединения составила 8–20 МПа в зависимости от времени экспозиции при максимальной мощности пучка (таблица), причем разрыв происходил по алюминию.

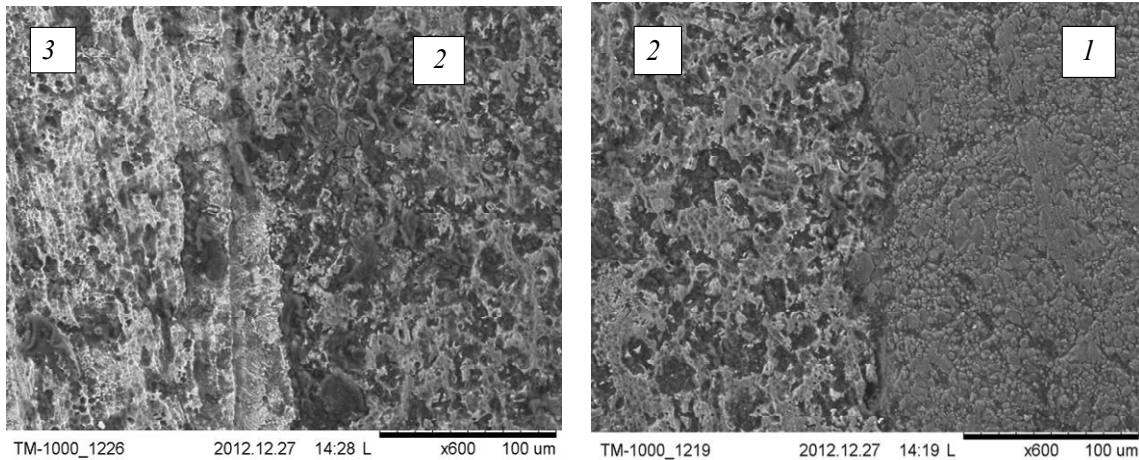


Рис. 5. Микрофотографии шлифов граничных областей соединения керамики с титаном через прослойку алюминия: 1 – керамика; 2 – алюминий; 3 – титан

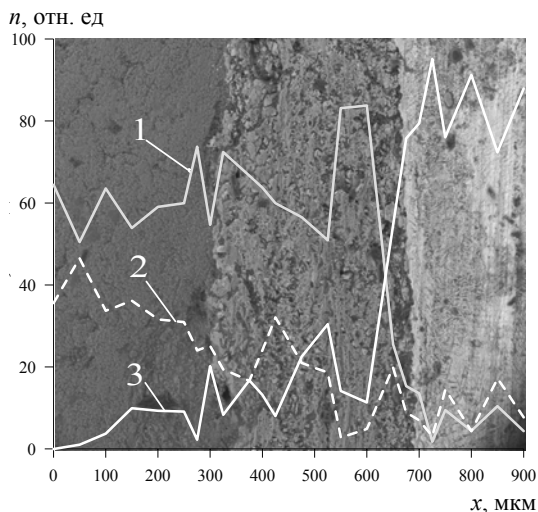


Рис. 6. Распределение концентрации n элементов в поперечном направлении к паяному соединению (на рисунке: слева – керамика, справа – титан, посередине – прослойка алюминия): 1 – алюминий; 2 – кислород; 3 – титан

Прочностные характеристики металлокерамических соединений

№ образца	Время экспозиции, мин	Предел прочности на разрыв, МПа
1	11	8
2	14	12
3	16	20

Несмотря на то, что указанные значения на порядок ниже прочности керамики, они могут оказаться достаточными для ряда металлокерамических узлов.

Заключение. Приведенные результаты указывают на перспективность использования электронно-лучевого метода для получения вакуумно-плотного металлокерамического соединения. Важной особенностью предлагаемой процедуры является отсутствие операции металлизации керамики, что заметно упрощает создание металлокерамических узлов.

Работа выполнена при поддержке РФФИ (12-08-31043 мол-а) и Министерства образования и науки Российской Федерации (проект №7.3101.2011).

Литература

1. Ерошев В.К. Конструирование и технология изготовления паяных металлокерамических узлов: справ. материалы / В.К. Ерошев, Ю.А. Козлов, В.Д. Павлова. – М.: Энергия, 1988. – Ч. 1. – 280 с.
2. Nascimento R.M. Recent advances in metal-ceramic brazing / R.M. Nascimento, A.E. Martinelli, A.J.A. Buschinelli // *Ceramica*. – Vol. 49, № 312. – P. 178–198.
3. Виноградов Б.А. Получение многослойных металлокерамических соединений лазерной пайкой / Б.А. Виноградов, Г.П. Мещерякова, А.О. Кудрявцев // *Механика композиционных материалов и конструкций*. – 2008. – № 2. – С. 250–268
4. Бурдовицин В.А. О возможности электронно-лучевой обработки диэлектриков плазменным источником электронов в форвакуумной области давлений / В.А. Бурдовицин, А.С. Климов, Е.М. Окс. // *Письма в ЖТФ*. – 2009. – Т. 35, вып. 11. – С. 61–66.
5. Гореев А.К. Электронно-лучевая сварка керамики с металлом с использованием форвакуумного плазменного источника электронов / А.К. Гореев, В.А. Бурдовицин, А.С. Климов, Е.М. Окс // *Перспективные материалы*. – 2012. – № 3. – С. 77–80.

Зенин Алексей Александрович

Мл. науч. сотрудник каф. физики ТУСУРа

Тел.: 8-952-807-62-02

Эл. почта: zenin1988@gmail.com

Климов Александр Сергеевич

Доцент каф. физики ТУСУРа

Тел.: 8-905-990-52-41

Эл. почта: klimov@main.tusur.ru

Zenin A.A., Klimov A.S.

Electron-beam soldering of alumina ceramics and metal with using plasma electron source

The results of the research aimed at the creation of mechanically strong and leakproof compound of alumina ceramics and titanium using electron beam soldering. Compensation of charging the surface of electrically non-conductive ceramic by electron beam is provided with using a forevacuum plasma electron source. As a binder layer we used aluminum.

Keywords: electron-beam soldering, forevacuum, plasma electron source, dielectric, metal-ceramic composites.

УДК 621.31

А.В. Осипов, Ю.А. Шурыгин, Ю.А. Шиняков, А.И. Отто, М.М. Черная

Сопоставительный анализ энергетической эффективности преобразования энергии солнечной батареи преобразователями постоянного напряжения

Проведены исследования энергетической эффективности применения преобразователей постоянного напряжения при реализации режима экстремального регулирования мощности солнечной батареи. Получена математическая модель, позволяющая определить регулировочные и энергетические характеристики преобразователей при работе от солнечной батареи на нагрузку в виде аккумулятора.

Ключевые слова: преобразователь энергии солнечной батареи, солнечная батарея, аккумуляторная батарея, энергетическая эффективность, экстремальное регулирование мощности.

Одним из основных способов повышения энергетической эффективности автономных фотоэлектрических энергетических установок является реализация режима отбора максимума мощности в оптимальной рабочей точке вольт-амперной характеристики солнечной батареи, т.е. реализация режима экстремального регулирования мощности солнечных батарей [1]. Эту функцию выполняют преобразователи энергии солнечных батарей (ПЭСБ), по иностранной терминологии контроллеры заряда аккумуляторных батарей с реализацией функции МРРТ (Maximum Power Point Tracker).

Одной из важнейших задач, возникающих при проектировании и создании ПЭСБ, является согласование уровней напряжения источников энергии в единой энергетической системе. Учитывая, что напряжение рабочей точки вольт-ваттной характеристики (ВВХ) соответствующей максимальной мощности нестабильно (изменяется в зависимости от освещенности и температуры фотоэлектрических панелей), ПЭСБ должен быть регулируемым. Необходимость обеспечения широкого диапазона регулирования напряжения СБ при поиске точки экстремума мощности существенным образом отражается на структуре построения ПЭСБ и затрудняет определение его энергетических характеристик, а также расчет оптимальных параметров элементов силовой схемы.

Как правило, ПЭСБ строятся на основе бестрансформаторных преобразователей постоянного напряжения, обеспечивающих высокий КПД за счет минимума силовых элементов, при этом массогабаритные показатели фильтра минимизируются высокими частотами преобразования. Однако схемы этого класса имеют ограниченный диапазон регулирования, что усложняет согласование уровней напряжения используемых источников энергии, в данном случае солнечной и аккумуляторной батарей (СБ и АБ). Согласование, как правило, достигается путем последовательного или параллельного соединения элементарных источников составляющих АБ или СБ, что позволяет настроить уровни напряжения источников таким образом, чтобы диапазон регулирования ПЭСБ охватывал точку максимальной мощности ВВХ с учетом ее возможного дрейфа. При этом не всегда удается достичь приемлемых энергетических характеристик системы, например, в случае применения преобразователя понижающего типа для того, чтобы охватить диапазоном регулирования точку максимальной мощности ВВХ часто приходится существенно уменьшать напряжение АБ или увеличивать напряжение СБ, что приводит к завышению габаритной мощности элементов ПЭСБ. Целью настоящей работы является анализ энергетической эффективности преобразователей энергии солнечной батареи на основе известных схем преобразователей постоянного напряжения по критериям коэффициента использования солнечной батареи K_p и коэффициента полезного действия преобразователя η .

Распространенным вариантом ПЭСБ на основе преобразователей постоянного напряжения является простейший понижающий преобразователь, представленный на рис. 1, а. Несомненным преимуществом схемы является стабильный ток заряда аккумулятора, обеспечиваемый большой выходной индуктивностью, однако прерывистое импульсное потребление входного тока требует установки громоздкого С-фильтра. Регулировочная характеристика такого преобразователя при условии непрерывности тока дросселя описывается соотношением

$$\frac{U_{ВЫХ}}{U_{ВХ}} = \frac{\gamma}{1 + R_{L1}/R_H}, \quad (1)$$

где γ – относительная длительность открытого состояния транзистора; R_{L1} – активное сопротивление дросселя L_1 ; R_H – сопротивление нагрузки.

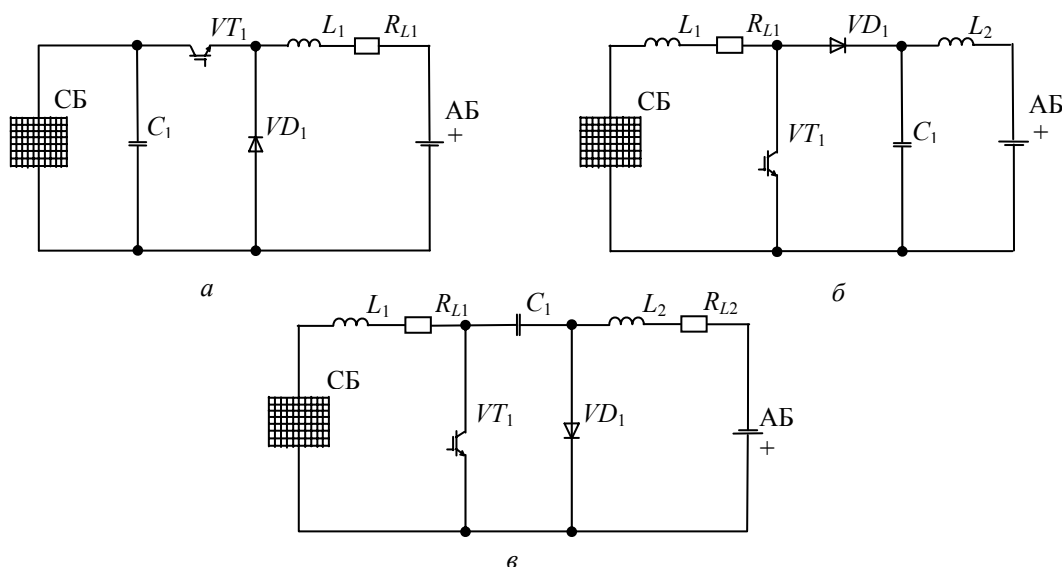


Рис. 1. Регулируемые преобразователи энергии солнечной батареи: а – понижающий преобразователь; б – повышающий преобразователь; в – преобразователь Кука

Следует заметить, что регулировочная характеристика (1) соответствует работе преобразователя от источника ЭДС. При питании преобразователя от солнечной батареи, имеющей нелинейную ВАХ и работе на АБ, которая является нагрузкой с непостоянным сопротивлением, характеристики существенно изменятся. Используя характеристику (1) и учитывая, что регулируемым параметром является входное напряжение, т.е. напряжение СБ, получим выражение

$$U_{СБ}^* = \frac{U_{СБ}}{U_{АБ}} = \frac{1 + R_{L1}/R_H}{\gamma} = \frac{1 + \frac{R_{L1}}{U_{АБ}} \cdot I_{ВЫХ}}{\gamma} = \frac{1 + \frac{R_{L1}}{U_{АБ}} \cdot \frac{I_{ВХ}}{\gamma}}{\gamma} = \frac{1}{\gamma} + \frac{R_{L1}}{U_{АБ}} \cdot \frac{I_{СБ}}{\gamma^2}.$$

В связи с существенной сложностью выражений, описывающих ВАХ реальной СБ, воспользуемся достаточно простой математической моделью СБ предложенной в [2], где ВАХ задана тремя характерными точками: напряжением холостого хода $U_{ХХ}$, током короткого замыкания $I_{КЗ}$, оптимальными значениями тока I_0 и напряжения U_0

$$I_{СБ} = I_{КЗ} \left[1 - \left(1 - \frac{I_0}{I_{КЗ}} \right) \frac{U_{ХХ} - U_{СБ}}{U_{ХХ} - U_0} \right], \quad (2)$$

используя которое получим квадратное уравнение

$$U_{СБ}^* = \frac{1}{\gamma} + \frac{R_{L1}^*}{\gamma^2} \left[1 - \left(1 - \frac{I_0}{I_{КЗ}} \right) \frac{U_{ХХ} - U_{СБ}}{U_{ХХ} - U_0} \right], \quad (3)$$

связывающее регулируемый параметр $U_{СБ}$ с относительной длительностью импульса управления γ . Сопротивление потерь в дросселе удобно выразить относительной величиной $R_{L1}^* = R_{L1} \cdot I_{КЗ} / U_{АБ}$.

Важным параметром, характеризующим энергетическую эффективность ПЭСБ, является коэффициент использования СБ K_p , определяемый соотношением фактически потребляемой мощностью от СБ и мощностью оптимальной точки ее ВАХ

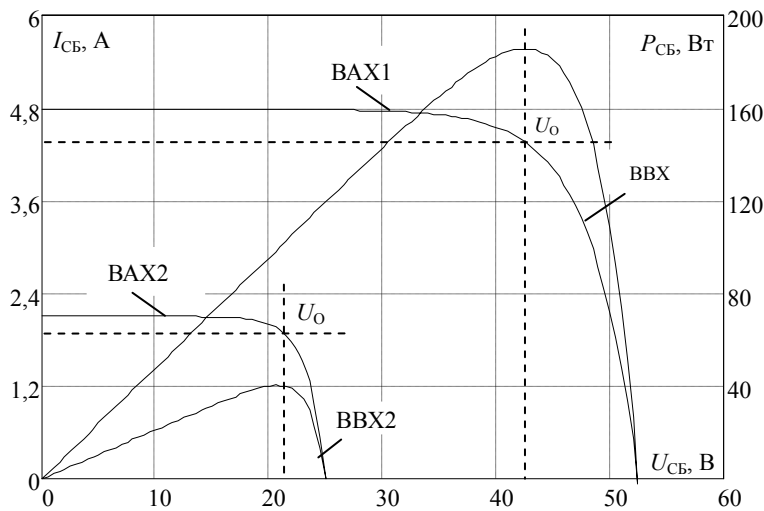
$$K_p = \frac{P_{СБ}}{P_{СБ_max}} = \frac{U_{СБ}}{U_0^*} \cdot \frac{I_{СБ}}{I_0}.$$

Другой рассматриваемый критерий энергетической эффективности – коэффициент полезного действия преобразователя η , который при пренебрежении потерями в транзисторе VT_1 , т.е. при учете только потерь в дросселе L_1 можно получить по выражению

$$\eta = \frac{1}{U_{CB}^*} \cdot \frac{I_{ВЫХ}}{I_{ВХ}} = \frac{1}{U_{CB}^*} \cdot \frac{1}{\gamma}.$$

Таким образом, задавая ВАХ СБ с помощью выражения (2), можно определить зависимости энергетических параметров преобразователя от управляющего воздействия γ .

Целесообразным является определение энергетических характеристик ПЭСБ в крайних точках его диапазона регулирования, т.е. при ВАХ соответствующих крайним значениям напряжения U_0 . В результате проведенных исследований установлено, что для панели КСМ-160 максимальное значение напряжения оптимальной точки соответствует солнечной погоде и температуре -30°C , при этом ВАХ1 имеет параметры $U_{XX} = 52,46\text{ В}$, $I_{K3} = 4,78\text{ А}$, $U_0 = 43,1\text{ В}$, $I_0 = 4,3\text{ А}$. Минимальное значение



напряжения оптимальной точки соответствует слабой освещенности панелей при температуре от $+60$ до $+70^\circ\text{C}$. ВАХ2 в этом случае имеет параметры $U_{XX} = 25,2\text{ В}$; $I_{K3} = 2,1\text{ А}$; $U_0 = 21,8\text{ В}$; $I_0 = 1,8\text{ А}$. Графически характеристики СБ в обоих случаях представлены на рис. 2.

Рис. 2. Характеристики фотоэлектрических панелей КСМ-160

Путем решения уравнения (3) в системе MathCad получены характеристики понижающего преобразователя (рис. 1а) при работе на ВАХ1 и ВАХ2 при напряжении аккумулятора $U_{AB} = 24\text{ В}$ и относительных потерях в дросселе $R_{L1}^* = 0,05$. Видно, что при температуре -30°C , входная мощность преобразователя имеет максимум при $\gamma = 0,59$, что соответствует близкому к единице коэффициенту использования $K_P \rightarrow 1$. Коэффициент полезного действия η в этой точке имеет минимум составляет $0,93\%$. Следует отметить, что во всем диапазоне регулирования γ напряжение СБ превышает напряжение АБ, т.е. $U_{CB}^* > 1$. Увеличение напряжения СБ ограничено напряжением U_{XX} СБ, которое достигается при $\gamma = 0,46$, при дальнейшем уменьшении γ и наложении условия непрерывности тока дросселя ток нагрузки отсутствует $I_{ВЫХ} = 0$, что обуславливает $K_{PD} \rightarrow 1$. Однако, ввиду конечности индуктивности дросселя преобразователь переходит в режим прерывистого тока и его параметры не могут описываться характеристикой (3).

Уменьшение напряжения оптимальной точки ВАХ до значения $U_0 = 21,8\text{ В}$ при $U_{AB} = 24\text{ В}$ не позволяет эффективно применять преобразователь понижающего типа, о чем свидетельствуют характеристики, представленные на рис. 3, б. Видно, что диапазон регулирования преобразователя в этом случае очень узкий и заключен в интервале $24\text{--}25\text{ В}$, и самое главное не охватывает точку ВАХ с максимальной мощностью. Коэффициент использования в этом случае $K_P = 0,33$ при $\gamma = 1$. Результаты говорят об эффективности применения в этом случае повышающего преобразователя, позволяющего получить напряжение СБ меньше напряжения АБ $U_{CB} < U_{AB}$.

Повышающий преобразователь (рис. 1, б) имеет на входе дроссель L_1 , обеспечивающий непрерывность входного тока и стабильность рабочей точки ВАХ по току. Однако учитывая импульсный режим заряда конденсатора C_1 , необходима установка небольшого выходного дросселя L_2 , обеспечивающего заряд АБ постоянным током. Регулировочная характеристика повышающего преобразователя при работе от идеального источника напряжения при условии непрерывности тока дросселя L_1 описывается широко известным соотношением

$$\frac{U_{ВЫХ}}{U_{ВХ}} = \frac{1-\gamma}{(1-\gamma)^2 + R_{L1}/R_H}. \quad (4)$$

Характеристика идеального преобразователя показывает увеличение напряжения на всем диапазоне регулирования, однако наличие потерь во входном дросселе приводит к ее искажению при больших значениях γ , что связано с резким ростом входного тока по отношению к выходному [3]. Для определения регулировочной характеристики, соответствующей режиму работы от СБ на нагрузку в виде АБ, произведены преобразования аналогичные проведенным для понижающего преобразователя и получено выражение

$$U_{СБ}^* = R_{L1}^* \left(1 - \left(1 - \frac{I_0}{I_{КЗ}} \right) \frac{U_{XX} - U_{СБ}}{U_{XX} - U_0} \right) + (1 - \gamma). \quad (5)$$

Характеристики повышающего преобразователя, полученные по формуле (5) при тех же параметрах представлены графически на рис. 3. Результаты показывают, что преобразователь повышающего типа имеет линейную регулировочную характеристику, при этом обусловленная потерями входного дросселя нелинейность, свойственная уравнению (4) в уравнении (5) отсутствует, что объясняется тем, что ВАХ СБ в этом случае находится на ветви стабилизации тока. Практическим выводом является невозможность этой схемы настроиться на точку ВАХ СБ с максимальной мощностью при $U_0 > U_{АБ}$, в результате чего, коэффициент использования СБ составляет $K_p = 0,65$ и соответствует $\gamma = 0$, т.е. полностью закрытому транзистору. Однако эффективной областью применения этой схемы является работа при слабой освещенности панелей, что подтверждается результатами решения уравнения (5) для ВАХ2 с $U_{02} = 21,8$ В, при которой соблюдается условие $U_0 < U_{АБ}$.

Добиться успешного регулирования точки ВАХ с максимальной мощностью при любом соотношении напряжений $U_{АБ}$ и U_0 можно с помощью составного широкодиапазонного преобразователя, например, преобразователя Кука, имеющего входной каскад на основе повышающего, а выходной на основе понижающего преобразователя (рис. 1, в). Регулировочная характеристика преобразователя Кука описывается соотношением [3]

$$\frac{U_{ВЫХ}}{U_{ВХ}} = \frac{\gamma}{1 - \gamma} \cdot \frac{1}{R_{L1}/R_H \cdot \left(\frac{\gamma}{1 - \gamma} \right)^2 + R_{L2}/R_H + 1}$$

при этом для соотношения токов выполняется условие

$$\frac{I_{ВЫХ}}{I_{ВХ}} = \frac{1 - \gamma}{\gamma}$$

Получить характеристики преобразователя Кука при работе от СБ на АБ можно решая уравнение относительно переменной γ , при подстановке $\Psi = \frac{1 - \gamma}{\gamma}$ уравнение примет вид

$$U_{СБ}^* = \Psi^2 \frac{R_{L2}}{U_{АБ}} I_{СБ} + \Psi + \frac{R_{L1}}{U_{АБ}} I_{СБ}$$

и имеет один положительный корень

$$\Psi = \frac{\sqrt{U_{АБ}^2 + 4R_{L2}I_{СБ}(U_{СБ} - R_{L1}I_{СБ}) - U_{АБ}}}{2R_{L2}I_{СБ}},$$

который с учетом (2) позволяет получить искомую аналитическую связь.

Результаты решения уравнения представлены на рис. 3, видно, что преобразователь Кука позволяет регулировать напряжение СБ в точках близких к максимальной мощности, т.е. получить $K_p \rightarrow 1$ как в случае $U_0 > U_{АБ}$ при $\gamma = 0,38$ с КПД равным $\eta = 0,91$, так и при $U_0 < U_{АБ}$ при $\gamma = 0,58$ и $\eta = 0,92$. Таким образом, при прочих равных условиях преобразователь Кука является единственным бестрансформаторным преобразователем позволяющим получить близкий к единице коэффициент использования СБ при любом соотношении напряжений U_0 и $U_{АБ}$. Однако следует учитывать, что преобразователь Кука, являясь составным преобразователем, производит двойное преобразование энергии и соответственно имеет завышенную по отношению к рассмотренным схемам габаритную мощность силовых элементов. Это проявляется при выборе транзистора, коммутирующего суммарный ток АБ и СБ, а также в наличии в схеме сразу двух дросселей L_1 и L_2 . Ухудшение КПД

преобразователя, связанное с введением дополнительного дросселя можно наблюдать по данным табл. 1 и рис. 3.

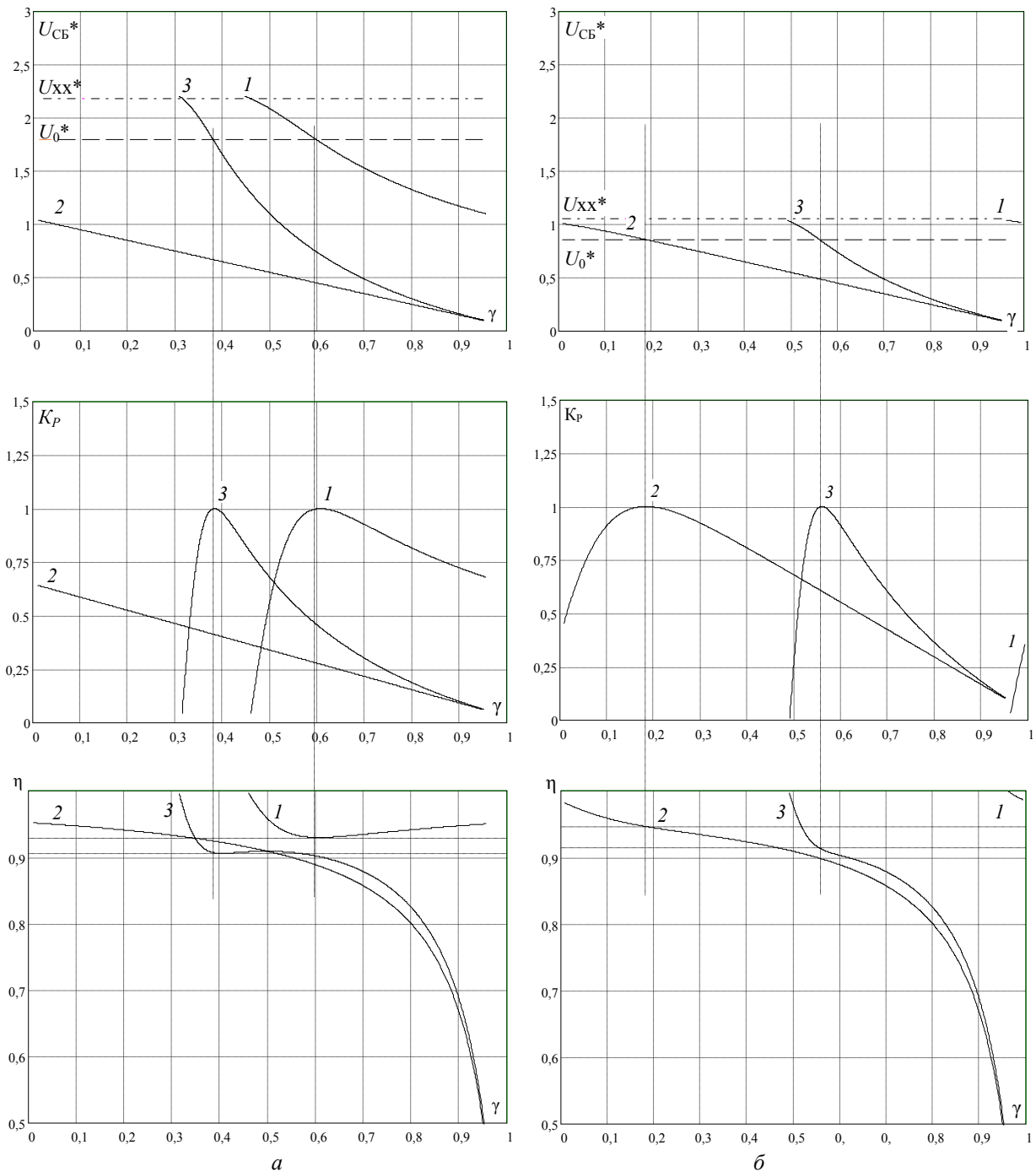


Рис. 3. Характеристики преобразователя энергии солнечной батареи на основе понижающего преобразователя (1), повышающего (2), преобразователя Кука (3) при $U_{AB} = 24В$, $R_{L1}^* = 0,05$ и параметрах: а – ВАХ1 $U_0 > U_{AB}$; б – ВАХ2 $U_0 < U_{AB}$

Для упрощения сравнительного анализа данные расчетов сведены в таблицу.

Сравнительный анализ схем преобразователей постоянного напряжения

Схема преобразователя	ВАХ1 $U_0 > U_{AB}$		ВАХ2 $U_0 < U_{AB}$	
	КПД, η	K_p	КПД, η	K_p
Понижающий преобразователь	0,93	1	0,98	0,33
Повышающий преобразователь	0,95	0,65	0,95	1
Преобразователь Кука	0,91	1	0,92	1

Характеристики построенные на основе полученной математической модели позволили произвести количественную оценку энергетических показателей преобразователей постоянного напря-

жения (таблица) сопоставительный анализ которых показывает, что преобразователь Кука за счет широкого диапазона регулирования имеет гораздо лучший K_p , фактически при любых условиях работы ($K_p \rightarrow 1$). Худшие значения КПД объясняются наличием в схеме сразу двух дросселей. При этом анализ современного состояния рынка показывает, что стоимость СБ в пересчете на единицу мощности существенно превышает стоимость силовых элементов преобразователя, это заставляет при выборе схемы отдавать предпочтение коэффициенту использования солнечной батареи K_p относительно габаритной мощности элементов силовой схемы.

Работа выполнена при поддержке Министерства образования и науки Российской Федерации, соглашения 14.В37.21.0420 и 14.В37.21.1493.

Литература

1. Шиняков Ю.А. Повышение энергетической эффективности автономных фотоэлектрических энергетических установок / Ю.А. Шиняков, Ю.А. Шурыгин, О.Е. Аркатова // Доклады Том. гос. ун-та систем управления и радиоэлектроники. – 2010. – № 2(22), ч. 2. – С. 102–107.
2. Привалов В.Д. Оценка эффективности применения экстремального регулятора в автономных СЭП / В.Д. Привалов, В.Е. Никифоров. – Куйбышев: КПИ, 1981. – 16 с.
3. Зиновьев Г.С. Основы силовой электроники. – Новосибирск: Изд-во НГТУ, 2003. – 664 с.

Осипов Александр Владимирович

Канд. техн. наук, зав. лабораторией НИИ автоматики и электромеханики ТУСУРа

Тел.: +7-8-903-914-09-67

Эл. почта: ossan@mail.ru

Шурыгин Юрий Алексеевич

Д-р. техн. наук, профессор, ректор ТУСУРа

Тел.: +7(382-2) 510-530

Эл. почта: office@tusur.ru

Шиняков Юрий Александрович

Д-р. техн. наук, директор НИИ космических технологий (КТ) ТУСУРа

Тел.: + 7(382-2) 90-01-62

Эл. почта: shua@main.tusur.ru

Отто Артур Исаакович

Аспирант каф. компьютерных систем в управлении и проектировании, мл. научн. сотрудник НИИ КТ ТУСУРа

Тел.: +7-923-415-92-32

Эл. почта: ottoai@mail.ru

Черная Мария Михайловна

Аспирант каф. компьютерных систем в управлении и проектировании, мл. научн. сотрудник НИИ КТ ТУСУРа

Тел.: +7-905-992-89-70

Эл. почта: cmm91@inbox.ru

Osipov A.V., Shurygin Yu.A., Shinyakov Yu.A., Otto A.I., Chernaya M.M.

The comparative analysis of the energy conversion efficiency of solar energy by DC converters

The article describes the research of the energy efficiency of DC converters for the extreme control of solar power. A mathematical model allows to determine the adjusting and energy characteristics of the converter when using the solar battery on the load in the form the battery.

Keywords: solar energy converter, solar battery, accumulator battery, energy efficiency, extreme power control.

УДК 621.31

В.В. Аржанов, В.Н. Мишин, Г.А. Ракитин, К.В. Аржанов

Управление позиционными электроприводами блока измерения освещенности для термобарокамеры

Приведены результаты разработки двухкоординатного позиционного электропривода блока измерения освещенности для термобарокамеры.

Ключевые слова: термобарокамера, шаговый электропривод, сервопривод, режим позиционирования, драйверы управления, микроконтроллер.

Для проведения наземных испытаний элементов спутников широко применяют термобарокамеры [1, 2]. Один из видов испытаний – это проверка воздействия солнечного излучения на элементы спутников. При данном испытании необходимо в термобарокамере (ТБК) замерять освещенность, создаваемую специальными лампами, которые имитируют солнечное излучение. Измерение осуществляется с помощью датчиков солнечного излучения, установленных на каретке, которая перемещается по всей плоскости на планшете, установленном в ТБК. Перемещение в заданную точку каретки, в которой установлены фотодатчики, измеряющие освещенность в заданном квадрате, осуществляется с помощью двухкоординатного позиционного электропривода.

В настоящее время в ТБК используются шаговые электроприводы и сервоприводы. Особенностью двигателей и механизмов является следующее: они установлены в самой термобарокамере и подвержены большим климатическим влияниям (температура, освещенность, вакуум), это приводит к изменению механических нагрузок в электроприводе. Особенно это сказывается в шаговых электроприводах, при пуске которых могут возникать сбои. Это может приводить к ошибке в перемещении к заданному положению.

В Томском государственном университете систем управления и радиоэлектроники разработана система позиционирования с шаговыми электроприводами для термобарокамеры. На рис. 1 приведена структурная схема двухкоординатного позиционного шагового электропривода, где М1, М2 – шаговые двигатели (ШД); Р1, Р2 – редукторы, ДШД1, ДШД2 – драйверы шаговых двигателей; КЭП – контроллер электропривода; КВ1–КВ4 – конечные выключатели; РС-104 – микрокомпьютер; I-7561 – конвертор. На планшете устанавливается каретка с датчиками солнечного излучения. В системе используются шаговые двигатели типа ШД-5Д и драйверы шаговых двигателей типа DM-378М.

КЭП выдает и принимает следующие сигналы с драйверов шаговых двигателей: сигналы включения (выключения) двух драйверов ШД (MF1 и MF2); сигналы готовности двух драйверов ШД (RDY1 и RDY2); число шагов, которое необходимо отработать ШД (CP1 и CP2); сигналы направления вращения ШД (DR1 и DR2). В драйвере амплитуда тока ШД и величина минимального шага перемещения для ШД задаются через канал RS 485 по протоколу MOD BUS от КЭП. КЭП имеет в своем составе микроконтроллер STM F32 103. Обмотки шагового двигателя ШД-5Д соединяются в треугольник при работе с драйвером DM-378М. Если используется драйвер KND BD3H, то обмотки двигателя ШД-5Д соединяются в звезду. Задаваемый в драйвере минимальный шаг перемещения в ШД составляет $0,01^\circ$, а максимальный задаваемый ток – до 7 А. Внешнее управление для КЭП идет через конвертор I-7561 от микрокомпьютера РС-104.

Для уменьшения рывков и качаний при перемещении в программном обеспечении КЭП заложен режим позиционирования, т.е. в электроприводе идет ограничение по скорости и ускорению при перемещении от одного заданного положения к другому. В КЭП реализован несимметричный режим разгона, движения и торможения двигателя. На рис. 2 показана диаграмма режима позиционирования в электроприводе. Несимметричный режим позиционирования в электроприводе совмещен с нелинейным заданием амплитуды тока в ШД, т.е. амплитуда тока ШД меняется во время движения. Это необходимо делать для исключения сбоев в ШД и минимизации потребляемого тока ШД. При пуске ШД с механизмом необходимо преодолеть нелинейный момент сухого трения, который меняется из-за изменения температуры и давления в термобарокамере. Поэтому целесообразно в электроприводе увеличить амплитуду тока ШД (пусковой ток увеличить до 1,5 от номинального тока), после преодоления начального момента сухого трения можно снизить амплитуду тока ШД

(рабочий ток уменьшить до 0,8 от номинального тока). Это позволяет исключить сбои при пуске ШД и в целом уменьшить потребление тока двигателем во время перемещения. Увеличение пускового тока в ШД целесообразно в режиме позиционирования при положительном значении ускорения (рис. 2).

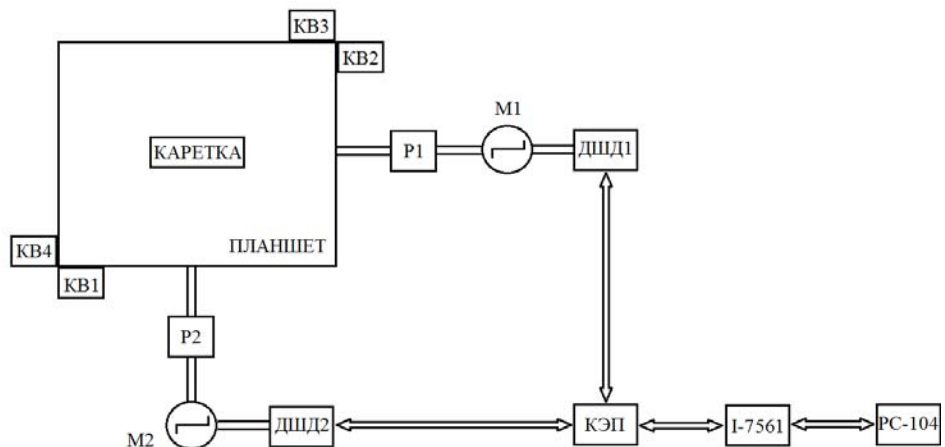


Рис. 1. Структурная схема двухкоординатного электропривода

Разработано программное обеспечение (ПО) для КЭП (программа STEP). Рабочее окно показано на рис. 3. В программе можно задавать следующие параметры: величину максимальной частоты в ШД, величину ограничения по ускорению и торможению, величину заданного перемещения, величину амплитуды пускового и рабочего тока ШД. В ПО КЭП заложены режим выхода каретки в начальное положение, режим измерения планшета и режим ограничения перемещения по двум координатам с помощью конечных выключателей.

По результатам экспериментальных исследований шаговых электроприводов получены следующие показатели: скорость вращения ШД-5Д можно поднять до 8000 Гц, ограничение по ускорению и торможению устанавливается на уровне 400 рад/с², задаваемое ограничение по току в ШД составляет от 4 до 3 А.

На рис. 4 показана фотография блока управления двухкоординатным позиционным электроприводом с шаговыми двигателями ШД-5Д.

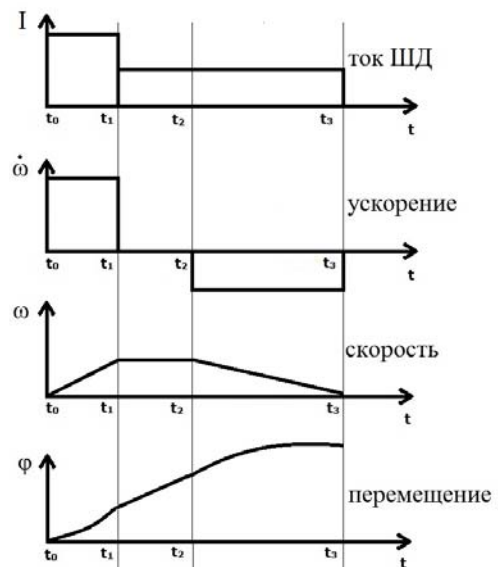


Рис. 2. Диаграмма режима позиционирования

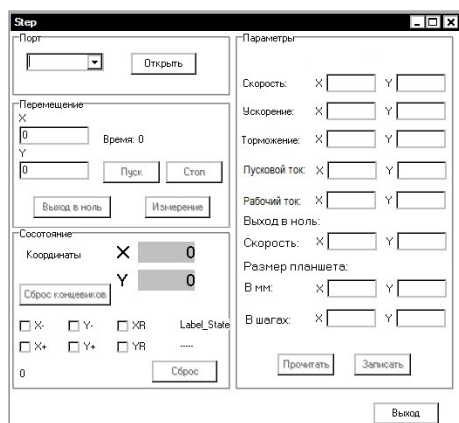


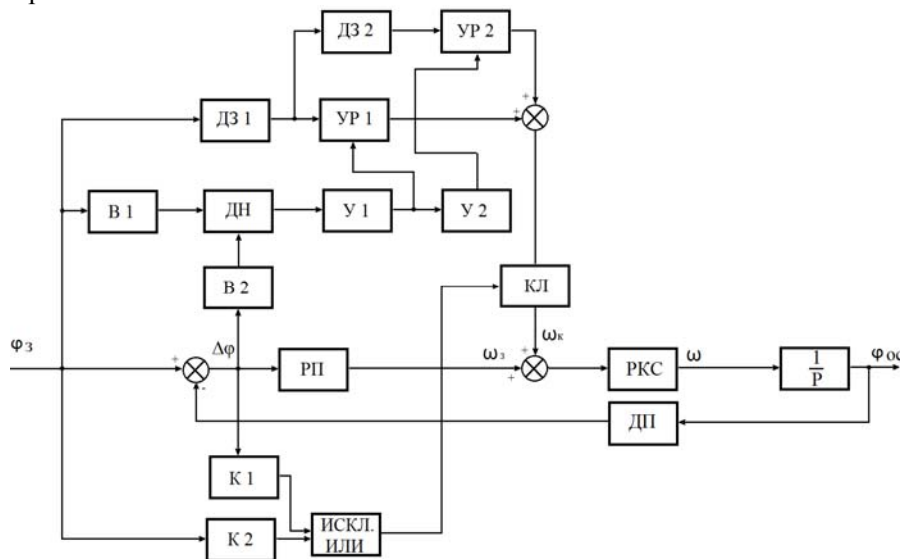
Рис. 3. Рабочее окно программы STEP



Рис. 4. Фотография блока управления двухкоординатным позиционным электроприводом с шаговыми двигателями

В некоторых термобарокамерах применяют сервопривод (регулируемый электропривод с вентильным двигателем) [3]. В ряде модификаций сервоприводов имеется программируемый внутренний контур регулирования положения. Как правило, контур регулирования положения стандартный и не имеет функции адаптивного управления или самонастраивающихся компенсационных связей. Поэтому целесообразно контур регулирования положения вынести отдельно от регулируемого сер-

воприводе, и в контур регулирования положения внести необходимые функции. В таком выполнении в структурной схеме системы управления (см. рис. 1) шаговые двигатели заменяются на вентильные двигатели с датчиками (положения ротора и положения), а драйверы шагового двигателя заменяются на блоки управления вентильными двигателями. Основное изменение – меняется программное обеспечение КЭП.



На рис. 5 показана функциональная схема разработанного контура регулирования положения с внутренним контуром регулирования скорости, реализованным на сервоприводе.

Рис. 5. Функциональная схема контура регулирования положения

На рис. 5 приняты следующие обозначения: РКС – регулируемый контур скорости сервопривода; ДЗ 1, ДЗ 2 – дифференцирующие звенья; К 1, К 2 – компараторы напряжения; КЛ – ключ; ИСКЛ. ИЛИ – элемент «исключающее или»; ДН – делитель напряжения; В 1, В 2 – выпрямители; УР 1 и УР 2 – усилители с регулируемым усилением; У 1, У 2 – усилители с ограничением; РП – регулятор положения; ДП – датчик положения; Φ_3 – задание по положению; $\Delta\Phi$ – ошибка по положению; ω_3 – заданная скорость; ω – действительная скорость; ω_k – корректирующий сигнал по скорости, $\Phi_{ос}$ – действительное положение.

Регулятор положения, реализованный в КЭП, выбран пропорционального типа. Для обеспечения инвариантности по управляющему воздействию дополнительно в контур регулирования положения вводятся две производные по сигналу задания положения. Для стабилизации динамической точности целесообразно обе производные регулировать по сигналу относительной величины ошибки по положению и в следящей системе производные целесообразно вводить только тогда, когда происходит отставание действительной траектории перемещения от заданной траектории перемещения. Режим отставания действительной траектории от заданной определяется с помощью двух компараторов, подключенных к задающему сигналу и сигналу по ошибке, и элемента «исключающее или». Выход элемента «исключающее или» подключен к ключу, который подключает или отключает регулируемые сигналы по производным во входу сервопривода. Дополнительный компенсирующий сигнал рассчитывается по формуле:

$$\omega_k = \begin{cases} 0, & \text{при } A = 0 \\ \frac{d\Phi}{dt} \cdot k_1 \cdot \frac{|\Delta\Phi|}{|\Phi_3|} + \frac{d^2\Phi}{dt^2} \cdot k_2 \cdot \frac{|\Delta\Phi|}{|\Phi_3|}, & \text{при } A = 1 \end{cases}, \quad (1)$$

где $\Delta\Phi$ – ошибка по положению; Φ_3 – задание по положению; $|\Delta\Phi|/|\Phi_3|$ – относительная ошибка по положению; k_1, k_2 – регулируемые коэффициенты передачи при первой и второй производной; A – выходной сигнал элемента «исключающее или».

Для уменьшения рывков и качаний в позиционном сервоприводе в программном обеспечении КЭП заложен режим позиционирования, причем реализован несимметричный режим разгона и торможения, т.е. торможение в электроприводе идет с пониженным ускорением по сравнению с разгоном (см. рис. 2).

Разработано программное обеспечение КЭП для управления сервоприводом (программа STEP-Serv). В программе можно задавать следующие параметры: величину максимальной скорости вращения двигателя, величину ограничения по ускорению и торможению, величину заданного переме-

шения. В программе STEP-Serv заложен режим выхода каретки в начальное положение, режим измерения планшета и режим ограничения перемещения по двум координатам с помощью конечных выключателей.

Двухкоординатный позиционный электропривод с шаговыми двигателями внедрен на термобарокамере ТБК-120 в ОАО «Информационные спутниковые системы» им. Акад. М.Ф. Решетнёва.

Заключение

1. Для блока измерения освещенности термобарокамеры ТБК-120 разработан двухкоординатный позиционный электропривод с шаговыми двигателями, имеющий несимметричный режим позиционирования, совмещенный с нелинейным заданием амплитуды тока в шаговом двигателе, обеспечивающий устойчивость и улучшение характеристик системы.

2. Для больших термобарокамер разработан двухкоординатный позиционный сервопривод, имеющий улучшенные динамические характеристики за счет введения двух производных по сигналу задания положения, которые регулируются в функции относительной величины ошибки в следящей системе, причем производные вводятся только при отставании действительной траектории по положению от заданной траектории.

Работа выполняется в рамках федеральной целевой программы «Научные и научно-педагогические кадры инновационной России» на 2009–2013 годы.

Литература

1. Решетневские чтения: матер. XIV Междунар. науч. конф., посвященной памяти ген. конструктора ракет.-космич. систем акад. М.Ф. Решетнева (10–12 нояб. 2010, г. Красноярск) / под общ. ред. Ю.Ю. Логинова: в 2 ч. – Красноярск: Сиб. гос. аэрокосмич. ун-т, 2010. – Ч. 1. – 364 с.

2. Сибирский центр российского спутникостроения [Электронный ресурс]. – Режим доступа: <http://www.elib.sfu-kras.ru>, свободный (дата посещения: 16.03.13).

3. Серводвигатели и сервоприводы [Электронный ресурс]. – Режим доступа: <http://www.dmotor.ru>, свободный (дата посещения: 16.03.13).

Аржанов Владимир Викторович

Канд. техн. наук, зав. отделом № 11 НИИ автоматики и электромеханики (АЭМ) ТУСУРа

Тел.: 8-(382-2) 55-78-43

Эл. почта: otdel11tomsk@yandex.ru

Мишин Вадим Николаевич

Канд. техн. наук, зав. отделом № 14 НИИ АЭМ ТУСУРа

Тел.: 8-(382-2) 59-00-59

Эл. почта: mishin@niiuem.tomsk.ru

Ракитин Геннадий Алексеевич

Зав. лабораторией № 141 НИИ АЭМ ТУСУРа

Тел.: 8-(382-2) 59-00-59

Эл. почта: rakitin@niiuem.tomsk.ru

Аржанов Кирилл Владимирович

Техник НИИ космических технологий ТУСУРа

Тел.: 8-(382-2) 55-78-43

Эл. почта: rbhx@yandex.ru

Arzhanov V.V., Mishin V.N., Rakitin G.A., Arzhanov K.V.

Control of positional electric drive units for measuring illumination thermal vacuum chamber

In this article there are given the results of designing two-axis positional drive unit for measuring illumination for thermal vacuum chamber.

Keywords: Space chamber, step electric drive, servo positioning mode, control drivers, microcontroller.

УДК 621.314.1:621.382.2/.3

С.В. Маморцев, Г.Я. Михальченко

Технология мягкой коммутации транзисторов в преобразователе постоянного напряжения в постоянное

Рассматриваются возможности построения преобразователя постоянного напряжения в постоянное с мягкой коммутацией полупроводниковых ключевых элементов и уменьшением пульсаций входного и выходного токов.

Ключевые слова: DC-DC-преобразователь, схема Кука, динамические характеристики, коммутирующий дроссель, мягкое включение транзисторов, пульсации.

Широкое применение находят простейшие преобразователи постоянного напряжения в постоянное (DC/DC-преобразователи) в источниках питания светодиодного освещения, сварочных аппаратах инверторного типа, зарядных устройствах аккумуляторных батарей, зарядных устройствах индуктивных и емкостных накопителей энергии, а также в технологических источниках питания, нагрузка которых может изменяться в широких пределах. Как правило, такого рода преобразователи характеризуются жесткой коммутацией ключевых полупроводниковых элементов, что сопровождается неоправданно высоким уровнем динамических потерь и электромагнитных помех [1]. Кроме того, для защиты от перегрузок и коротких замыканий в цепи нагрузки в известных устройствах, основанных на базе повышающего преобразователя, требуется применять дополнительные довольно сложные схемотехнические решения, что приводит к снижению надежности, увеличению энергопотребления и массогабаритных показателей, а также к увеличению себестоимости продукции.

Особое место среди DC/DC-преобразователей занимает преобразователь постоянного напряжения в постоянное, выполненный по схеме Кука (рис. 1), в силу непрерывности как потребляемого, так и выходного токов и меньших пульсаций выходного напряжения [2].

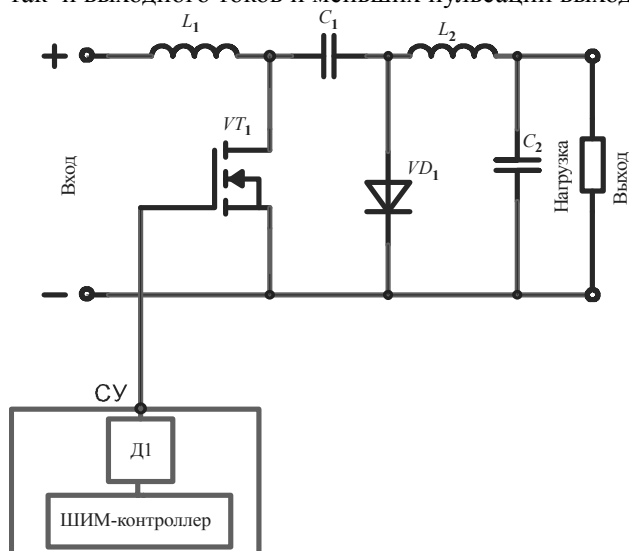


Рис. 1. Преобразователь Кука

Но и преобразователю Кука также свойственны высокие значения динамических потерь, связанные с тем, что к моменту включения транзистора его «паразитная» емкость заряжена до удвоенного напряжения источника питания, и при включении транзистора эта емкость разряжается импульсным током, в десятки раз превышающим номинальное значение тока транзистора. Кроме того, дополнительно в этот же момент транзистор нагружается еще и обратным током выключающегося диода по цепи: положительная обкладка накопительного конденсатора C , транзистор VT , диод VD , отрицательная обкладка накопительного конденсатора. Суммарное значение этого импульсного тока сопровождается выделением на кристалле транзистора мощности в

десятки киловатт за доли микросекунды, что приводит к постепенной деградации кристалла и, как следствие, к сокращению срока службы преобразователя. Традиционно снижение динамических нагрузок на кристалл транзистора достигается включением в схему снабберных RC - или RCD -цепей, что приводит к снижению коэффициента полезного действия.

Наиболее хорошие результаты по уменьшению динамических потерь достигнуты в схемах резонансных и квазирезонансных преобразователей. Но стоит отметить, что в данных решениях регулирование выходного напряжения осуществляется изменением длительности паузы при постоянстве длительности резонансного колебания, т.е. используется частотно-импульсная модуляция, что

ограничивает диапазон регулирования выходных токов и напряжений. Расширение этого диапазона способствует увеличению девиации частоты, что значительно повышает частотные требования к полупроводниковым приборам, магнитным материалам, конденсаторам, увеличиваются пульсации входного тока и выходного напряжения, а это сопровождается повышением массогабаритных показателей и стоимости фильтрующих элементов. Поэтому разработчики устройств силовой электроники ведут поиск путей построения преобразователей с широтно-импульсной или широтно-частотно-импульсной модуляцией при мягкой коммутации транзисторов [3].

Одним из путей разрешения такого рода проблемной ситуации является построение двухфазной схемы преобразователя с коммутационным дросселем (рис. 2), обеспечивающим мягкое переключение транзисторов с широтно-импульсной модуляцией. Этим достигается отсутствие всплесков выделяемой на кристалле MOSFET транзистора мгновенной мощности в моменты включения и выключения транзистора [4]. Этот эффект обусловлен переключением транзисторов при нулевом токе, что повышает коэффициент полезного действия посредством уменьшения динамических потерь.

Структурная схема двухфазного преобразователя по схеме Кука с коммутирующим дросселем представлена на рис. 2, а временные диаграммы, поясняющие принцип работы преобразователя на рис. 3, 4.

В установившемся режиме работы преобразователя в момент времени t_1 (см. рис. 2) выключается транзистор VT_1 и включается транзистор VT_2 по сигналам ШИМ контроллера выходными напряжениями драйверов $U_{д1}$ и $U_{д2}$. Ток I_{L1} начинает

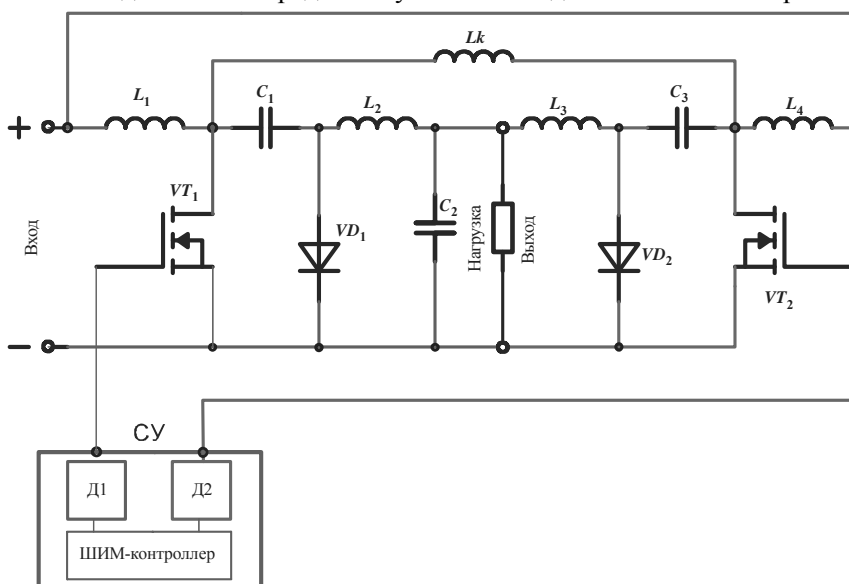


Рис. 2. Преобразователь постоянного напряжения в постоянное

нарастать, а ток I_{L4} спадает. Конденсатор C_1 подзарядается до удвоенного значения напряжения источника питания, в то время как конденсатор C_3 подзарядается током нагрузки через включенный транзистор VT_2 , который выключается в момент времени t_2 . Суммарный ток дросселей L_1 и L_4 представляет собой потребляемый от источника питания ток $I_{вх}$. Аналогично сумма токов I_{L2} и I_{L3} выходных дросселей является током нагрузки I_n и током конденсатора фильтра C_6 .

На интервале выключенного состояния транзистора VT_1 дроссель L_2 отдает накопленную энергию в конденсатор фильтра и нагрузку по цепи: вывод дросселя L_2 , диод VD_1 , конденсатор C_2 , второй вывод дросселя L_2 . Заряженный на интервале времени $t_2 - t_5$ до удвоенного напряжения источника питания конденсатор C_2 на следующем интервале разряжается через транзистор по цепи: плюс конденсатора C_3 , транзистор VT_2 , конденсатор C_2 , дроссель L_3 , минус конденсатора C_3 . Дроссель L_3 при этом накапливает энергию.

В интервале времени $t_2 - t_3$ оба транзистора выключены, питание нагрузки обеспечивается энергией конденсатора фильтра C_2 . Дроссель L_4 , конденсатор C_3 , коммутирующий дроссель Lk , конденсатор C_1 , дроссель L_2 образуют контур, обеспечивающий при максимальных углах регулирования ШИМ контроллера снижение потенциала на стоке выключенного транзистора VT_1 до отрицательного напряжения, при котором включается встроенный в него обратный диод, как показано на рис. 3. Поскольку потенциал U_{VT1} стока выключенного в интервале времени $t_1 - t_3$ транзистора VT_1 определяется суммой напряжений замкнутого контура, образованного дросселями L_1 , L_4 , Lk , и к моменту времени t_2 становится равным нулю, то и включение его в момент времени t_3 осуществляется при нулевом напряжении и отрицательном токе I_{VT1} , протекающем через встроенный обратный диод. В интервале времени $t_3 - t_5$ аналогичные процессы протекают в транзисторе VT_2 , работающем в противотактном режиме.

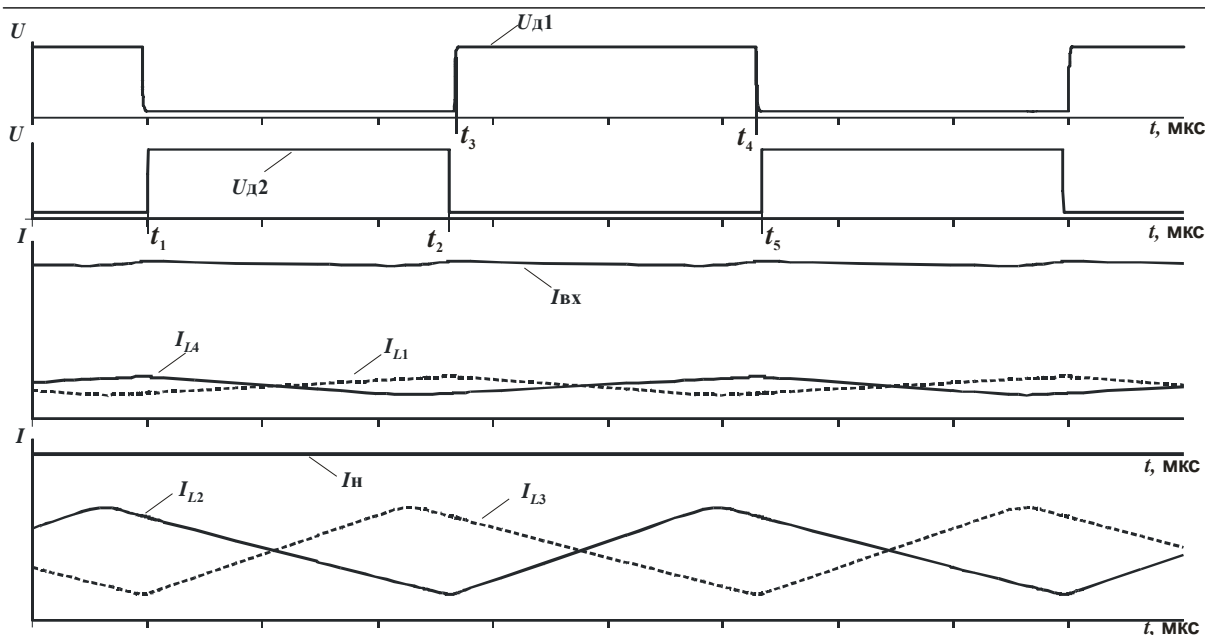


Рис. 3. Временные диаграммы работы преобразователя

Рассмотрим процессы мягкого включения транзисторов, которые пояснены диаграммами на рис. 4. К концу интервала проводимости, например транзистора VT_1 , ток дросселя L_2 убывает до нулевого значения и под действием напряжения на конденсаторе C_2 изменяет свое направление. В это же время ток подзарядки конденсатора C_1 снижается, приближаясь к значению, близкому к нулю, и как только он сравняется с током дросселя L_2 , диод VD_1 запирается. Ток дросселя L_2 продолжает протекать, теперь уже по цепи: вывод дросселя L_2 , конденсатор C_1 , паразитная емкость стока-исток транзистора VT_1 , конденсатор фильтра C_2 , второй вывод дросселя L_2 . Этим током разряжается паразитная емкость транзистора VT_1 до нулевого значения и перезаряжается до противоположной

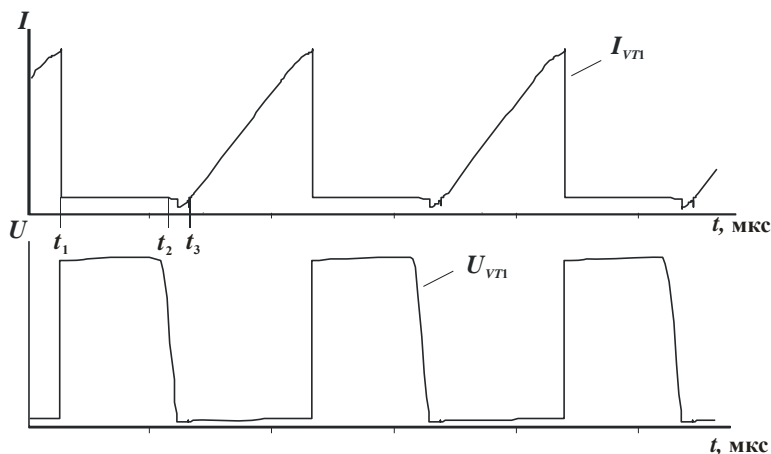


Рис. 4. Временные диаграммы работы преобразователя

полярности, при которой открывается встроенный обратный диод транзистора, который и ограничивает обратное напряжение на транзисторе. Мягкая коммутация транзистора возможна с данного момента времени и до момента, когда накопленная энергия дросселя L_2 не снизится до нулевого значения. В противотактном транзисторе VT_2 процессы включения протекают аналогично рассмотренным.

Таким образом, к моменту включения транзисторов VT_1 и VT_2 потенциалы U_{VT1} и U_{VT2} стока этих транзисторов спадают до нуля, и через встроенные обратные диоды начинает протекать отрицательный ток в направлении исток-сток, как показано на рис. 3. Это приводит к тому, что в момент включения рассеиваемая на кристалле транзистора мощность равна нулю. В момент выключения транзисторов в предлагаемом преобразователе и в преобразователе по классической схеме Кука динамическая мощность на транзисторах сравнима и находится на уровне статической рассеиваемой мощности.

Кроме того, как показано на рис. 2, токи I_{L4} и I_{L1} складываются таким образом, что потребляемый входной ток $I_{вх}$ преобразователя становится практически постоянным за счет того, что имеет удвоенную частоту пульсаций по отношению к частоте коммутации транзисторов и минимальные амплитуды пульсаций. Аналогично, за счет эффекта сложения токов I_{L2} и I_{L3} , изменяющихся противоположно, выходной ток $I_{н}$ нагрузки преобразователя имеет сглаженную форму постоянного

тока. Пиковые значения тока и рассеиваемой мощности транзистора преобразователя по классической схеме Кука во много раз выше значения тока и мощности транзистора рассматриваемого преобразователя. В частности, по результатам имитационного моделирования преобразователя, построенного по классической схеме Кука (см. рис. 1), численные значения амплитуды динамического тока разряда «паразитной» емкости транзистора в моменты включения в 22 раза превышают таковые в математической модели схемы, представленной на рис. 2. Такой эффект сопровождается тем, что пиковые значения амплитуды мгновенной мощности снижаются в сотни раз. В частности, в сравниваемых моделях снижение пиковой мощности достигает 289 раз.

Заключение. Таким образом, рассмотренный преобразователь обладает существенно меньшими динамическими потерями в момент включения транзисторов, обеспечивает повышение коэффициента полезного действия, улучшает надежность и срок службы преобразователя. Также значительно уменьшаются пульсации входного и выходного токов в количестве раз, равное количеству параллельно соединенных ячеек преобразователей по схеме Кука с коммутирующими дросселями, соединяющими стоки транзисторов. Не менее важно и то, что коммутирующий дроссель переменного тока L_k определяет строго противофазное изменение переменной составляющей токов дросселей L_1, L_2, L_3, L_4 и строгое равенство постоянных составляющих, т.е. придает каждой ячейке «естественные» свойства равномерного распределения этих токов.

Литература

1. Михальченко Г.Я. Научно-технические проблемы электропитания / Г.Я. Михальченко, С.Г. Михальченко, А.Н. Четин // Труды Всерос. научно-техн. конф. – М.: ЗАО «ММП-Ирбис», 2011.
2. U. S. Patent 4,257,087. Mar. 17, 1981. DC-to-DC Switching Converter With Zero Input And Output Current Ripple And Integrated Magnetics Circuits / Slobodan M. Cuk.
3. Кобзев А.В., Михальченко Г.Я., Музыченко Н.М. Модуляционные источники питания РЭА. – Томск: Радио и связь, Томский отдел, 1990. – 336 с.
4. Пат. 2464692 РФ, МПК Н 02 М 3/335. Преобразователь постоянного напряжения в постоянное (варианты) / Г.Я. Михальченко (РФ), С.В. Маморцев (РФ), С.Г. Михальченко (РФ). – №2 011 134 950 / 07; заявл. 19.08.2011; опубл. 20.10.2012. Бюл. № 29. – 13 с.

Маморцев Степан Валерьевич

Аспирант каф. промышленной электроники ТУСУРа
Тел.: 8 (382-2) 90-00-29, доб.107
Эл. почта: stepan_mamortsev@mail.ru

Михальченко Геннадий Яковлевич

Д-р техн. наук, профессор каф. промышленной электроники ТУСУРа
Тел.: 8 (382-2) 41-32-32
Эл. почта: kpe-tusur@yandex.ru

Mamortsev S.V., Mikhachenko G.Y.

Technology of soft switching of transistors in DC-DC converter

Opportunities of DC-DC converter designing with soft switching of semiconductor switching elements and reducing of input and output currents ripple are explored in article.

Keywords: DC-DC converter, Cuk scheme, dynamic characteristics, commutating choke, soft switching of transistors, ripple.

УДК 621.372:621.375

Г.Н. Якушевич

Алгоритм расчета транзисторного широкополосного усилителя с двухполосниками комбинированной обратной связи

Предложен алгоритм расчета транзисторного широкополосного усилителя на основе математической модели активного четырехполосника для широкополосного усилителя с двухполосниками комбинированной обратной связи (ОС). Предложена обобщенная структурная схема активного четырехполосника, представленного активным элементом, на входе и выходе которого включены фазированно-согласующие цепи (ФСЦ).

Ключевые слова: алгоритм, математическая модель, активный четырехполосник, широкополосный усилитель, двухполосник, комбинированная ОС.

В работе [1] было показано, что при использовании простейших структур цепей двухполосников комбинированной ОС необходимо с помощью фазированно-трансформирующих цепей (ФТЦ) и симметрирующих цепей (СЦ), включенных на входе и выходе активного элемента, приблизить параметры активного элемента к требуемым параметрам активного четырехполосника для каскада широкополосного усилителя с двухполосниками комбинированной ОС. Введение двух типов цепей, ФТЦ и СЦ, усложняет расчет, принципиальную схему и настройку широкополосного усилителя.

В данной работе для приближения параметров активного элемента к требуемым параметрам активного четырехполосника для широкополосного каскада с двухполосниками комбинированной ОС предложена обобщенная структурная схема активного четырехполосника, представленного активным элементом, на входе и выходе которого включены фазированно-согласующие цепи (ФСЦ), которые выполняют одновременно функции фазирования и согласования, что упрощает расчет, реализацию и настройку широкополосного усилителя.

Требуемые S-параметры активного четырехполосника для широкополосного каскада с двухполосниками комбинированной ОС. Математическая модель в виде аналитических выражений для требуемых S-параметров идеально однонаправленного активного четырехполосника для согласованного каскада широкополосного усилителя с двухполосниками комбинированной ОС, представленного на рис. 1, запишется [1]

$$S_{11АЧ} = S_{22АЧ} = \frac{(Y_{OC} - Z_{OC})(1 - S_{21OC})}{2 - (Y_{OC} + Z_{OC})(1 - S_{21OC})}, \quad (1)$$

$$S_{21АЧ} = \frac{4S_{21OC} - 2(Y_{OC} + Z_{OC})(1 - S_{21OC}^2) + 4Y_{OC}Z_{OC}(1 - S_{21OC})^2}{[2 - (Y_{OC} + Z_{OC})(1 - S_{21OC})]^2}, \quad (2)$$

$$G_{НОМАЧ} = |S_{21АЧ}|^2 / (1 - |S_{11АЧ}|^2)^2. \quad (3)$$

где $S_{11АЧ}$, $S_{22АЧ}$, $S_{21АЧ}$, $G_{НОМАЧ}$, S_{21OC} , Z_{OC} , Y_{OC} – коэффициенты отражения по входу и выходу, прямой коэффициент передачи, номинальный коэффициент передачи по мощности активного четырехполосника, прямой коэффициент передачи каскада широкополосного усилителя с комбинированной ОС, нормированные сопротивление и проводимость двухполосников комбинированной ОС.

Из выражений (1)–(3) следует, что при условии равенства $Y_{OC} = Z_{OC}$ соотношения для требуемых S-параметров идеально однонаправленного активного четырехполосника для согласованного каскада широкополосного усилителя с двухполосниками комбинированной ОС запишутся

$$S_{11АЧ} = S_{22АЧ} = 0; \quad (4)$$

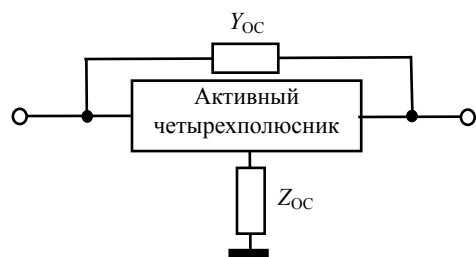


Рис. 1. Активный четырехполосник с двухполосниками комбинированной ОС

$$S_{21АЧ} = \frac{S_{21OC} + Y_{OC}(S_{21OC} - 1)}{1 + Y_{OC}(S_{21OC} - 1)}; \quad (5)$$

$$G_{НОМАЧ} = |S_{21АЧ}|^2. \quad (6)$$

По выражению (5) для заданной структуры двухполосников комбинированной ОС можно определить требуемые S -параметры активного четырехполосника для каскада широкополосного усилителя без ОС, при этом комплексный коэффициент передачи S_{21OC} задается в следующем виде:

$$S_{21OC} = |S_{21OC}| \exp(i\Theta_{21OC}), \quad (7)$$

где Θ_{21OC} – фаза комплексного коэффициента передачи.

Нормированные значения для R -двухполосников комбинированной ОС, проводимость параллельной ОС Y_{OCR} и сопротивление последовательной ОС Z_{OCR} определим из выражения

$$Y_{OCR} = Z_{OCR} = 1 / (1 + |S_{21OC}|); \quad (8)$$

для последовательного RL -двухполосника проводимость параллельной ОС Y_{OCR_L} и параллельного RC -двухполосника, сопротивление последовательной ОС Z_{OCR_C} находим из выражения

$$Y_{OCR_L} = Z_{OCR_C} = Y_{OCR} / [1 + i(1 - \Theta_{21OC} / 180^\circ)]; \quad (9)$$

для параллельного RC -двухполосника проводимость параллельной ОС Y_{OCR_C} и последовательного RL -двухполосника, сопротивление последовательной ОС Z_{OCR_L} находим как

$$Y_{OCR_C} = Z_{OCR_L} = Y_{OCR} \cdot [1 + i(1 - \Theta_{21OC} / 180^\circ)]. \quad (10)$$

На рис. 2 приведены зависимости: модуля коэффициента передачи $|S_{21OC}|$ от фазы коэффициента передачи Θ_{21OC} для каскада широкополосного усилителя с ОС; модуля коэффициента передачи $|S_{21R}|$ от фазы коэффициента передачи Θ_{21R} для активного четырехполосника с R -двухполосниками комбинированной ОС; модуля коэффициента передачи $|S_{21LC}|$ от фазы коэффициента передачи Θ_{21LC} для активного четырехполосника с RL -двухполосником параллельной ОС Y_{OCR_L} и RC -двухполосником последовательной ОС Z_{OCR_C} ; модуля коэффициента передачи $|S_{21CL}|$ от фазы коэффициента передачи Θ_{21CL} для активного четырехполосника с RC -двухполосником параллельной ОС Y_{OCR_C} и RL -двухполосником последовательной ОС Z_{OCR_L} , рассчитанные на основе выражений (5)–(10) для коэффициента передачи широкополосного усилителя $|S_{21OC}| = 12$ дБ и $\Theta_{21OC} = -90^\circ \div +90^\circ$ с шагом $\Delta\Theta_{21OC} = 22,5^\circ$.

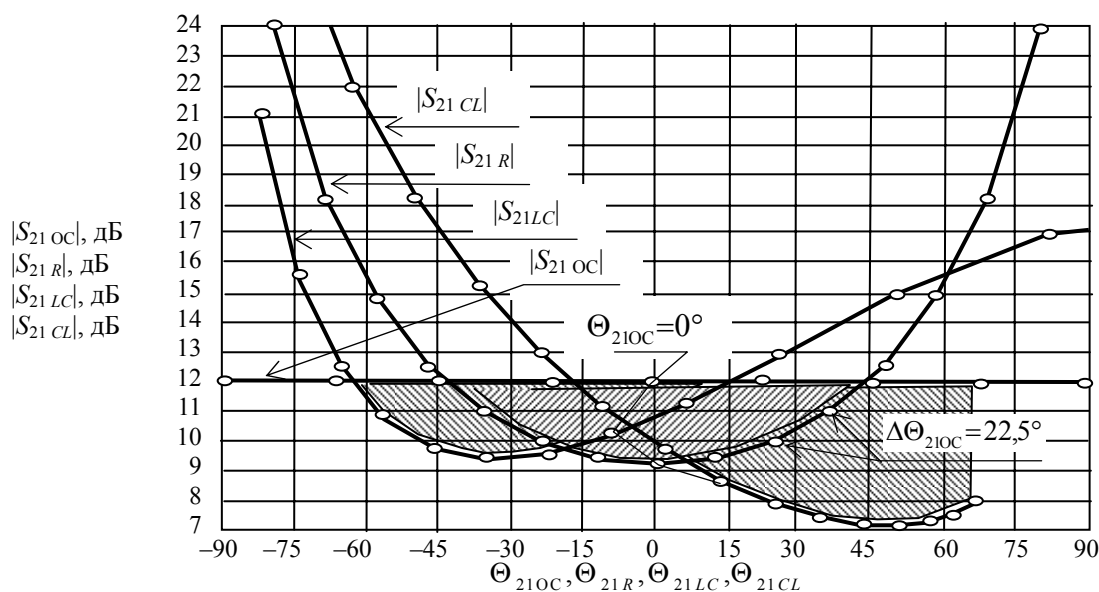


Рис. 2. Зависимости $|S_{21OC}|$, $|S_{21R}|$, $|S_{21LC}|$, $|S_{21CL}|$ активного четырехполосника для широкополосного усилителя с коэффициентом передачи $|S_{21OC}| = 12$ дБ

Анализ зависимостей, приведенных на рис. 2, показывает, что в областях ниже коэффициента передачи для каскада широкополосного усилителя с ОС $|S_{21OC}| = 12$ дБ получается выигрыш как по коэффициенту усиления, так и по верхней граничной частоте по отношению к номинальному коэффициенту передачи активного элемента.

В общем случае зависимости параметров активного элемента отличаются от требуемых параметров активного четырехполюсника, приведенных на рис. 2, и для их сближения на входе и выходе активного элемента необходимо включать ФСЦ. Обобщенная структурная схема широкополосного усилителя с двухполюсниками комбинированной ОС на основе активного элемента с ФСЦ на входе и выходе, образующими активный четырехполюсник, представлена на рис. 3.

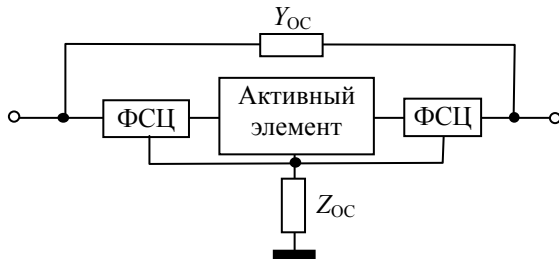


Рис. 3. Обобщенная структурная схема широкополосного усилителя с двухполюсниками комбинированной ОС

Для активного элемента при двухстороннем согласовании выражения для номинального коэффициента усиления по мощности $G_{НОМ}$ и для оптимальных проводимостей генератора $Y_{ГО}$ и нагрузки $Y_{НО}$ запишутся [2]

$$G_{НОМ} = \frac{|y_{21}|}{|y_{12}|} (K_Y - \sqrt{K_Y^2 - 1}); \quad (11)$$

$$Y_{ГО} = G_{ГО} + iB_{ГО} = \frac{|y_{12} y_{21}| \sqrt{K_Y^2 - 1}}{2\text{Re}(y_{22})} + i \left(\frac{\text{Im}(y_{12} y_{21})}{2\text{Re}(y_{22})} - \text{Im}(y_{11}) \right); \quad (12)$$

$$Y_{НО} = G_{НО} + iB_{НО} = \frac{|y_{12} y_{21}| \sqrt{K_Y^2 - 1}}{2\text{Re}(y_{11})} + i \left(\frac{\text{Im}(y_{12} y_{21})}{2\text{Re}(y_{11})} - \text{Im}(y_{22}) \right), \quad (13)$$

где $G_{ГО}$, $B_{ГО}$, $G_{НО}$, $B_{НО}$ – реальная и мнимая части оптимальных проводимостей генератора и нагрузки; y_{11} , y_{12} , y_{21} , y_{22} – y -параметры активного элемента; $| \cdot |$, Re , Im – модуль, реальная и мнимая части комплексного числа; K_Y – инвариантный коэффициент устойчивости, равный

$$K_Y = \frac{2\text{Re}(y_{11})\text{Re}(y_{22}) - \text{Re}(y_{12}y_{21})}{|y_{12}y_{21}|}. \quad (14)$$

Для безусловной устойчивости активного элемента должны выполняться следующие неравенства: $K_Y > 1$, $\text{Re}(y_{11}) > 0$, $\text{Re}(y_{22}) > 0$. Если эти неравенства не выполняются, то на входе или выходе вводят стабилизирующие проводимости.

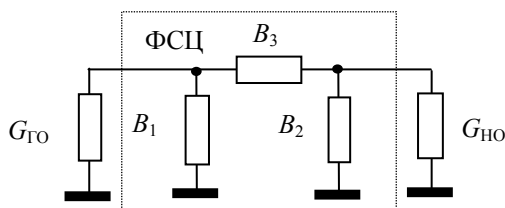


Рис. 4. Схема ФСЦ

Расчет ФСЦ. Расчет ФСЦ, приведенной на рис. 4, проведем на основе нормированной матрицы S -параметров согласованного реактивного обратимого четырехполюсника, которая запишется в следующем виде [3]:

$$[S] = \begin{bmatrix} s_{11} & s_{12} \\ s_{21} & s_{22} \end{bmatrix} = \begin{bmatrix} 0 & \exp(i\Theta) \\ \exp(i\Theta) & 0 \end{bmatrix}, \quad (15)$$

где Θ – значение фазы коэффициента передачи ФСЦ.

Используя нормированную матрицу S -параметров согласованного реактивного обратимого четырехполюсника (15), на основе соотношений перехода от S -параметров к Y -параметров получим сначала матрицу нормированных Y -параметров ФСЦ, а затем матрицу ненормированных Y -параметров ФСЦ в следующем виде

$$Y_{ФСЦ} = \begin{bmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{bmatrix} = \begin{bmatrix} G_{ГО} \frac{1 + \exp(i2\Theta)}{1 - \exp(i2\Theta)} & -2\sqrt{G_{ГО} G_{НО}} \frac{\exp(i\Theta)}{1 - \exp(i2\Theta)} \\ -2\sqrt{G_{ГО} G_{НО}} \frac{\exp(i\Theta)}{1 - \exp(i2\Theta)} & G_{НО} \frac{1 + \exp(i2\Theta)}{1 - \exp(i2\Theta)} \end{bmatrix}, \quad (16)$$

где Y_{11} , Y_{12} , Y_{21} , Y_{22} – Y -параметры ФСЦ.

Задавая значения фазы в выражении (16), можно рассчитать Y -параметры ФСЦ. Так как ФСЦ является реактивным обратимым четырехполюсником, то матрица Y -параметров ФСЦ содержит только мнимые части, и недиагональные элементы матрицы равны $Y_{12} = Y_{21}$. Тогда выражения для нормированных значений элементов принципиальной схемы ФСЦ запишутся

$$B_1 = \text{Im}(Y_{11}) + \text{Im}(Y_{12}), \quad B_2 = \text{Im}(Y_{22}) + \text{Im}(Y_{21}), \quad B_3 = -\text{Im}(Y_{12}). \quad (17)$$

Если проводимости генератора и нагрузки содержат мнимые части $B_{Г0}$ и $B_{Н0}$, то они учитываются при расчете B_1 и B_2 .

Обобщенная структурная схема активного четырехполюсника для широкополосного усилителя с комбинированной ОС на основе активного элемента с ФСЦ на входе и выходе приведена на рис. 5.

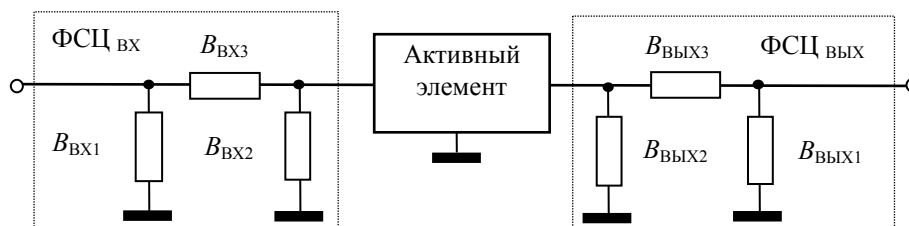


Рис. 5. Схема активного четырехполюсника на основе активного элемента с ФСЦ на входе и выходе

Алгоритм расчета каскада широкополосного усилителя с двухполюсниками комбинированной ОС. Алгоритм расчета ориентирован на максимальное приближение S -параметров активного элемента к требуемым параметрам активного четырехполюсника для широкополосного усилителя с двухполюсниками комбинированной ОС и заключается в следующем:

1) на основе S -параметров активного элемента на заданной верхней частоте рассчитываются Y -параметры активного элемента и по выражениям (11)–(14) – инвариантный коэффициент устойчивости K_Y , номинальный коэффициент усиления по мощности $G_{НОМ}$, оптимальные проводимости генератора $Y_{Г0}$ и нагрузки $Y_{Н0}$;

2) для номинального коэффициента усиления по мощности $G_{НОМ}$ на основе выражения (5) рассчитываются и строятся зависимости требуемых S -параметров активного четырехполюсника для широкополосного каскада с двухполюсниками комбинированной ОС (8)–(10), аналогичные зависимостям рис. 2;

3) для приближения S -параметров активного элемента к требуемым параметрам активного четырехполюсника задается значение фазы коэффициента передачи ФСЦ и по выражениям (16) и (17) рассчитываются Y -параметры ФСЦ, а затем S -параметры активного четырехполюсника, приведенного на рис. 5, которые сравниваются с требуемыми параметрами активного четырехполюсника, рассчитанными в пункте 2, и уточняется структура двухполюсников ОС [3];

4) рассчитываются S -параметры широкополосного усилителя с двухполюсниками комбинированной ОС.

Применим этот алгоритм к расчету каскада широкополосного усилителя.

Математическое моделирование широкополосного усилителя с комбинированной ОС. Математическое моделирование проведем на основе эквивалентной схемы кристалла, которая использовалась при математическом моделировании в работе [1].

На рис. 6–8 приведены принципиальные схемы и результаты расчета S -параметров для различных вариантов эквивалентных схем кристалла транзистора [1].

Результаты расчетов показывают, что чем точнее приближены с помощью ФСЦ S -параметры кристалла транзистора к требуемым параметрам активного четырехполюсника, рассчитанным в соответствии с п. 2, тем равномернее частотная зависимость коэффициента передачи и меньше коэффициенты отражения широкополосного усилителя с комбинированной ОС (рис. 6, б, в). Отклонение от требуемых S -параметров приводит к появлению неравномерности коэффициента передачи и увеличению коэффициента отражения на средних частотах (рис. 7, б, в и 8, б, в).

Экспериментальная проверка аналогичных теоретических исследований и математического моделирования была приведена автором в предыдущих статьях [1, 4–6].

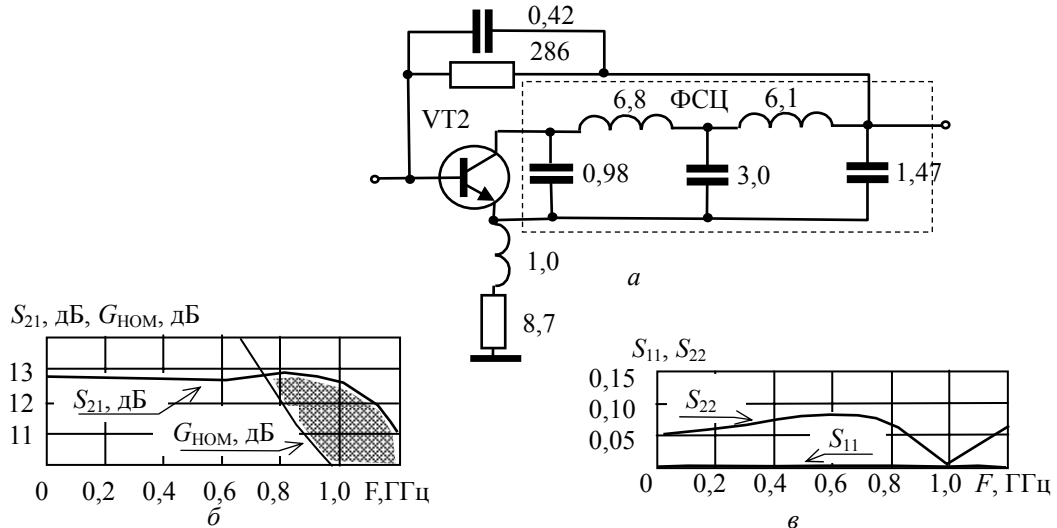


Рис. 6. Принципиальная схема (а), коэффициенты S_{21} (б), S_{11} и S_{22} (в) для кристалла типа VT1

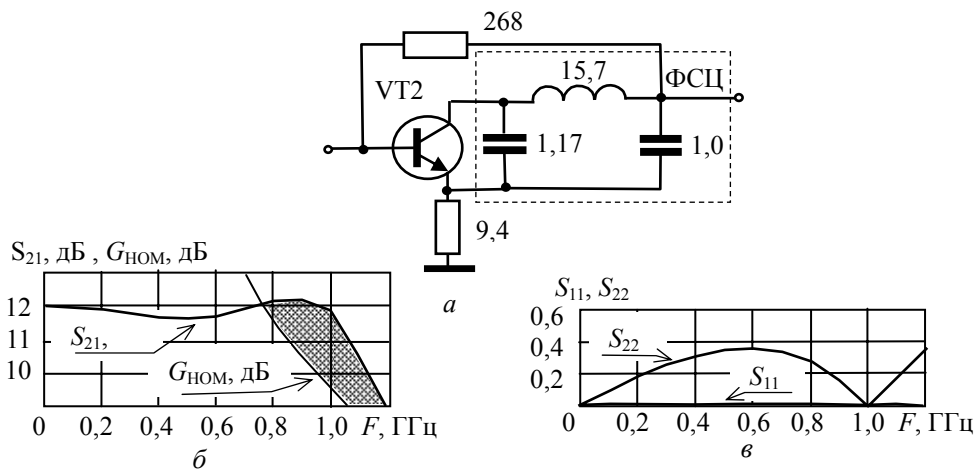


Рис. 7. Принципиальная схема (а), коэффициенты S_{21} (б), S_{11} и S_{22} (в) для кристалла типа VT2

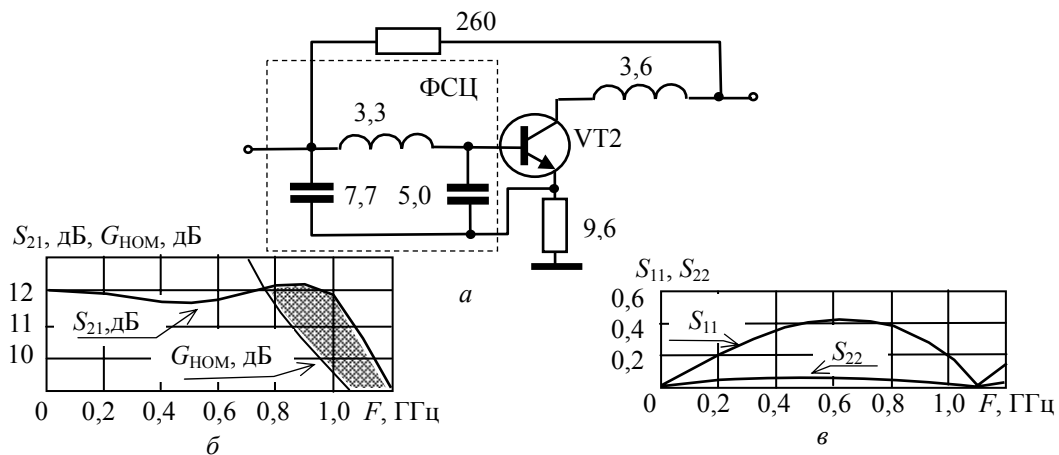


Рис. 8. Принципиальная схема (а), коэффициенты S_{21} (б), S_{11} и S_{22} (в) для кристалла типа VT3

Полученные результаты. Предложен алгоритм расчета каскада транзисторного широкополосного усилителя с двухполосниками комбинированной ОС, основанный на том, что прежде чем ввести двухполосники комбинированной ОС, необходимо приблизить S -параметры активного элемента с помощью ФСЦ к требуемым S -параметрам активного четырехполосника для каскада широкополосного усилителя с двухполосниками комбинированной ОС.

Приведены результаты математического моделирования широкополосных усилителей с комбинированной ОС для различных вариантов эквивалентных схем кристалла транзистора [1], показывающие выигрыш в верхней граничной частоте и в коэффициенте передачи по отношению к номинальному коэффициенту передачи по мощности за счет положительной ОС (см. рис. 6, б – 8, б, заштрихованные области).

Показано, что применение на входе и выходе активного элемента ФСЦ вместо двух типов цепей ФТЦ и СЦ [1] упрощает расчет, принципиальную схему активного четырехполюсника и, следовательно, реализацию и настройку широкополосного усилителя.

Литература

1. Якушевич Г.Н. Математическая модель активного четырехполюсника для широкополосного усилителя с двухполюсниками комбинированной обратной связи // Доклады Том. гос. ун-та систем управления и радиоэлектроники. – 2012. – № 1 (25). – С. 34–39.
2. Шварц Н.З. Линейные транзисторные усилители СВЧ. – М.: Советское радио, 1980. – 368 с.
3. Фельдштейн А.Л. Синтез четырехполюсников и восьмиполюсников на СВЧ / А.Л. Фельдштейн, Л.Р. Явич. – 2-е изд., перераб. и доп. – М.: Советское радио, 1971. – 388 с.
4. Якушевич Г.Н. Математическая модель активного четырехполюсника для широкополосного СВЧ-усилителя с двухполюсником параллельной обратной связи // Доклады Том. гос. ун-та систем управления и радиоэлектроники. – 2009. – № 2 (20). – С. 32–37.
5. Якушевич Г.Н. Алгоритм расчета транзисторного широкополосного усилителя с двухполюсником параллельной обратной связи // Доклады Том. гос. ун-та систем управления и радиоэлектроники. – 2010. – № 1 (21), ч. 2. – С. 63–68.
6. Якушевич Г.Н. Проектирование транзисторных широкополосных усилителей с двухполюсником параллельной обратной связи // Доклады Том. гос. ун-та систем управления и радиоэлектроники. – 2011. – № 1 (23). – С. 65–70.

Якушевич Геннадий Николаевич

Канд. техн. наук, доцент каф. средств радиосвязи ТУСУРа
Тел.: (382-2) 41-37-09
Эл. почта: mrc@main.tusur.ru

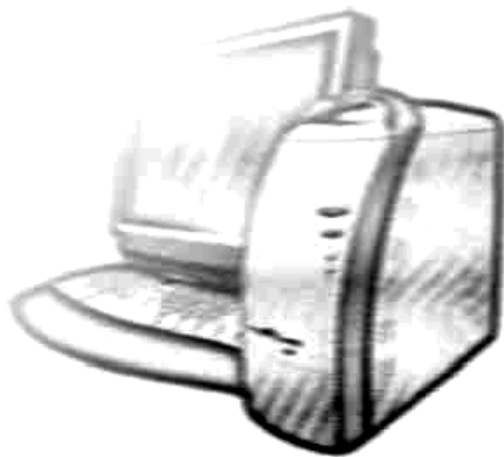
Jakushevitch G.N.

Calculation algorithm of a wideband amplifier with two-pole serial and parallel feedback

An algorithm of a wideband amplifier calculation on the basis of an active four-pole in case of a wideband amplifier with two-pole serial and parallel feedback (FB) is suggested. Also a generalized structure of the active four-pole presented by an active element with phase-matching circuits at the input and at the output is suggested.

Keywords: algorithm, mathematical model, active four-pole, wideband amplifier, serial and parallel feedback two-port.

**УПРАВЛЕНИЕ,
ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА
И ИНФОРМАТИКА**



УДК 004.932.2

Н.Л. Щеголева

Разработка программных средств для моделирования биометрических систем распознавания человека по изображению лица

Сформулированы и обоснованы требования к средствам моделирования систем распознавания изображений лиц, принципы их построения и анализа. Предложена архитектура пакета, приведены состав подсистем и их назначение.

Ключевые слова: распознавание лиц, предметно-ориентированное моделирование.

В России биометрические системы стали появляться уже в середине девяностых годов. По ряду очевидных причин все они были чрезвычайно дороги. Однако совершенствование цифровых камер, давших возможность получать все более качественные изображения, вычислительной техники, обеспечившей возможность быстрой обработки значительных объемов данных, появление новых методов обработки и анализа, активное развитие мультимедийных и цифровых технологий дали резкий толчок для развития биометрических систем. На сегодняшний день такие системы существенно подешевели, что стало причиной для их широкого распространения. Но это только одна сторона, вторая же причина популярности биометрических систем – это возможность обеспечения безопасности офиса, дома, интеллектуальных и материальных ресурсов и т.д.

Среди существующих биометрических систем стоит особо отметить системы распознавания человека по изображению лица. Они являются наиболее удобными, поскольку эти технологии *ненавязчивы (распознавание происходит на расстоянии, не задерживая и не отвлекая человека), они, как правило, пассивны (не требуют специальных действий со стороны человека), они не ограничивают пользователя в свободе перемещений*» [1].

Однако в работе [2] справедливо указывается, что факторы, влияющие на процесс обработки и анализа изображений, могут быть весьма изменчивыми и неформализуемыми. Изображения лиц, несмотря на их известную структуру, расположение глаз, носа, рта, не являются исключением. Основные факторы, которые усложняют работу системы, – это трудные условия съемки, недостаток освещения, сложный текстурированный фон, сокрытие частей лиц, поворот влево/вправо, наклон головы, а также эмоции, которые испытывает человек в разные моменты времени.

Отсутствие общепринятого формализованного описания перечисленных факторов, вносящих разного рода неопределенности в процесс обработки, приводит к тому, что в отличие от многих других областей обработки сигналов говорить о существовании единственного оптимального алгоритма для решения той или иной задачи обработки изображений (а их огромное количество) практически невозможно.

Последние десять лет в области распознавания человека по изображению лица ведутся активные разработки и предложены различные методы распознавания, использующие такие подходы, как метод главных компонент, методы с использованием гистограмм, нейросети, байесовские сети, классификатор Viola-Jones, статистические методы и т.д. При этом ни один из методов не может быть признан предпочтительнее остальных. Иными словами, ни один из существующих методов не позволяет получить 100% результат.

Однако, как показал опыт, существует некоторая совокупность методов, позволяющая добиться более качественного результата [3]. Это приводит к важной проблеме выбора некоторой совокупности методов из существующих и определения последовательности их применения к имеющимся данным для обеспечения требуемого результата распознавания, т.е. построению системы распознавания человека по изображению лица (Face Recognition Systems – FaReS).

Поэтому дальнейшее повышение качества и эффективности биометрической идентификации может быть достигнуто за счет рационального сочетания методов, оптимизации архитектуры FaReS и максимизации производительности ее аппаратного и программного обеспечения (ПО).

Следует отметить, что в настоящее время основные исследования направлены на совершенствование методов обработки изображений лиц, однако чрезвычайно мало внимания уделяется вопросам, посвященным разработке и оптимизации FaReS как взаимосвязанной последовательности методов предобработки, экстракции/селекции признаков и классификации для решения конкретной задачи. Связано это, в частности, с очень малым количеством полноценных средств моделирования FaReS, что не позволяет добиться максимальной эффективности разрабатываемых систем.

Сложность решаемых задач биометрии требует разработки концепции построения средств моделирования, в рамках которых можно было бы исследовать перспективные структуры систем и их основные модули, ставить натурные эксперименты, получая эффективные решения.

Структура FaReS. Процесс обработки изображений лиц состоит из следующих этапов [3] (рис. 1): детекции области каждого лица на исходном цифровом изображении; экстракции признаков – представления изображения выделенного лица (всей выделенной прямоугольной области) в форме вектора исходных признаков; селекции некоторых признаков из полного набора исходных признаков или редукции исходного пространства признаков (СП/РПП); сравнения признаков нового образа (изображения лица) с признаками эталонов и принятия решений о принадлежности этого образа к одному из известных классов.

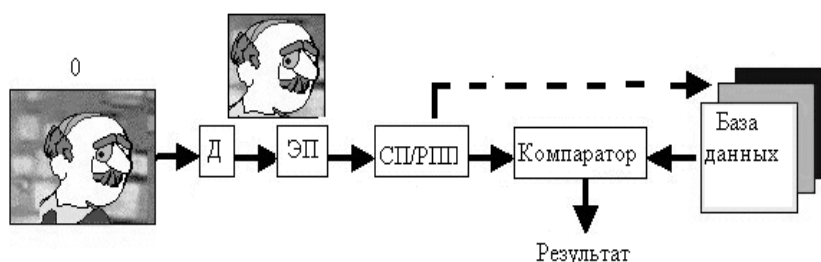


Рис. 1. Структура системы распознавания изображений лиц

Следует заметить, что каждый из перечисленных выше этапов обработки изображений может быть реализован различными способами, состоять из нескольких методов (каждый из которых зависит от ряда параметров), иметь параллельную или каскадную структуру. Все это определяет наличие очень большого количества вариантов FaReS, причем точность распознавания, вычислительная сложность, время обработки данных и т.д. полученных систем могут существенно отличаться. Поэтому основные усилия разработчиков будут сосредоточены на поиске оптимальной структуры FaReS (рис. 2).



Рис. 2. Алгоритм разработки системы [4]

Сложность построения структуры оптимальной FaReS, большое количество параметров и переменных, неполнота и недетерминированность исходной информации, разнообразие и случайный характер воздействий внешней среды – все это обуславливает необходимость моделирования.

Современные средства моделирования. Анализ существующего ПО с точки зрения возможности использования его для создания средств моделирования FaReS показывает, что его можно разделить на две группы [4]:

- универсальное ПО (среды программирования на языках высокого уровня – Microsoft Visual Studio, Piton, Eclipse и т.д.);
- специализированные пакеты программ – среды технического моделирования и вычислений (MATLAB, LabVIEW и т. д.) и специализированные программные средства с фиксированным количеством модулей (National Instruments Vision Assistant [5], FaReS-MOD (Face Recognition System Modeler) [6] и FaRetSys (Face Retrieval System) [7], Fabio [8]).

Очевидно, что использование универсального ПО будет сопряжено с большим количеством проблем: необходимо наличие грамотных специалистов по обработке изображений, времени, поскольку программирование всех необходимых алгоритмов обработки изображений лиц весьма трудоемкий процесс, написание огромного объема сервисного ПО для визуализации изображений, обеспечения интерактивного ввода параметров, векторной графики и т.п. Поэтому данный вариант представляется не самым лучшим решением.

Преимуществом большинства сред технического моделирования и вычислений являются универсальность и широкий набор сторонних библиотек математической статистики и цифровой обработки изображений, однако в них отсутствует возможность написания удобного и развитого интерфейса.

Альтернативой является использование готового специализированного ПО для моделирования FaReS. Рассмотрим существующие решения более подробно.

Среда Vision Assistant не является полноценным средством моделирования систем распознавания, т.к. ориентирована на решение задач пакетной обработки изображений, позволяет создавать в интерактивном режиме только линейные вычислительные модели без возможности классификации.

Среда FaReS-MOD является самой функциональной. Имеющийся в ней набор модулей позволяет проектировать системы, содержащие все этапы распознавания от предобработки до классификации, но только линейной структуры. В ней детально проработаны методология построения систем распознавания, настройка систем и получение результатов в процессе моделирования.

Среда FaRetSys во многом аналогична предыдущей, но позволяет реализовать каскадную и параллельную структуру для построения эффективных, в плане качества распознавания и быстродействия, систем на базе простых классификаторов. Возможно расширение функционала пользовательскими модулями. Недостатками среды являются малый базовый функционал, невозможность изменения параметров модулей системы.

Программное средство Fabio в значительной степени аналогично представленным выше средам, имеет модульную структуру, основной особенностью является использование для реализации графического редактора моделей и интерфейсной части Microsoft Visual Studio 2010, а для реализации функций, предназначенных для анализа изображений, – пакет Matlab. Возможно расширение функционала пользовательскими модулями, написанными на языке пакета Matlab. Недостатками являются ограниченный набор методов и линейная структура системы.

Описание модели эксперимента. Необходимо отметить, что в статьях, посвященных распознаванию изображений лиц, модели экспериментов авторами описываются различным образом, что часто затрудняет понимание, каким образом проводился эксперимент и какие параметры были выбраны для конкретных методов, что не позволяет исследователю повторить эксперимент и корректно сравнить результаты. В связи с этим необходимо унифицировать описание экспериментов.

Большинство исследователей предпочитают описание в форме таблицы, поэтому рассмотрим способ компактной и однозначной записи модели [9] и ее связь с табличным описанием. Пусть описание эксперимента выполнено в форме таблицы, дополнительно известно, что классификация выполнялась по минимуму расстояния, использовалась процедура кросс-валидации, размер изображения 112×92. Очевидно, что эта важная информация должна быть известна, но весьма часто авторы ее не указывают в описании эксперимента.

Описание эксперимента

База данных	Число классов (K)	Число образов в классе (Q)	Количество эталонов в классе (L)	Количество тестовых образов (Q–L)	Метод экстракции признаков	Метрика	Ранг
ORL	40	10	1	9	2DPCA	L1	1

Тогда модель эксперимента в соответствии с [9] будет иметь вид:

$$\text{ORL}(40/1/9/CV) \{2\text{DPCA}:112 \times 92 \rightarrow d \times d / \text{KMP}/L1/\text{rank}=1\},$$

где определены:

ORL(40/1/9/CV) – база лиц, структура данных и процедура кросс-валидации (CV);

2DPCA:112×92→d×d – использованный метод и размеры данных до и после редукции размерности пространства признаков;

KMP/L1/rank = 1 – тип классификатора, метрика и ранг результата.

Данное описание является более компактным и содержит все данные эксперимента.

Требования. Анализ работ посвященных разработке систем распознавания, показал, что требования, предъявляемые к программным средствам, предназначенным для моделирования FaReS, могут быть разделены на:

1. Общие:

- удобство использования,
- форматы используемых данных,
- возможность использования различных тестовых баз данных изображений,
- организация структуры программного средства,
- функциональный состав модулей,
- возможность интерактивно формировать из имеющихся готовых блоков любые системы обработки без использования какого-либо специального командного языка,
- описание модели эксперимента,
- сохранение структуры и параметров моделей систем и возможность их загрузки для повторения эксперимента или дальнейшей работы,
- возможность корректировки параметров используемых модулей после тестирования FaReS,
- возможность осуществлять автоматическую обработку данных в моделируемой системе при их обновлении,
- оценка качества разработанной системы.

2. Инженерные:

- скорость обработки данных,
- возможность использования видеозаписей,
- возможность оптимизации ранее разработанных систем,
- возможность получения готового программного продукта,
- построение гибкой структуры моделируемых FaReS.

3. Исследовательские:

- анализ состава и характеристик пространства признаков, представляющих информацию о лицах,
- поиск минимального размера пространства признаков для представления лиц конкретной базы данных,
- формирование отчетов в зависимости от целей исследования,
- возможность добавления новых функциональных модулей,
- сравнительный анализ различных структур FaReS,
- возможность использования разработанных FaReS для анализа изображений другой природы,
- автоматический выбор оптимальной структуры среди заданных пользователем.

Обоснование требований. Рассмотрим данные требования более подробно. Заметим, что удобство использования поставлено первым, т.к. разработка биометрической системы идентификации человека по изображению лица является весьма сложным процессом, требующим от разработчика максимум времени и внимания, а это означает, что эргономические показатели, предъявляемые к данному виду ПО, будут весьма высоки.

Все перечисленные выше специализированные пакеты используют визуальное программирование для моделирования FaReS, процесс построения которой [3] состоит в выборе функциональных блоков (рис. 3), реализующих методы обработки, выборе их параметров, установлении между ними связей (рис. 4). Это позволяет уделять основное внимание вопросам, связанным с разработкой и оптимизацией систем распознавания как взаимосвязанной последовательности методов предобработки, экстракции/селекции признаков и классификации. Интерактивный характер разработки должен обеспечивать возможность анализа системы, а также модификации проекта FaReS – его структуры и параметров произвольно выбранных блоков.

Сохранение в проекте FaReS иерархии блоков и соединений модели, включая их настройки и информацию для отображения, может быть реализовано с помощью файла, имеющего специальную структуру [6–8].

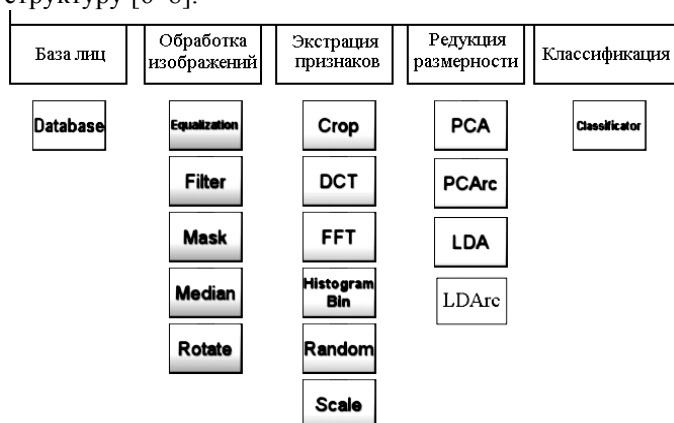


Рис. 3. Функциональные блоки, используемые в пакете FaReS-MOD

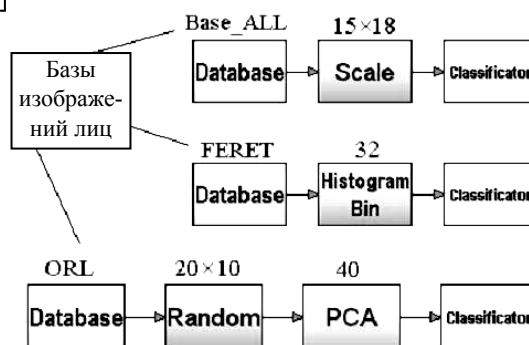


Рис. 4. Примеры FaReS, построенных в среде FaReS-MOD

Программное средство для моделирования должно обеспечивать возможность обработки наиболее популярных файловых форматов, загрузку, обработку, просмотр и сохранение бинарных, полутоновых и цветных изображений, возможность использования различных тестовых баз данных изображений, а также любых двумерных и одномерных числовых массивов, как целочисленных, так и вещественных.

В работе [4] обосновывается необходимость использования модульной структуры (рис. 5).

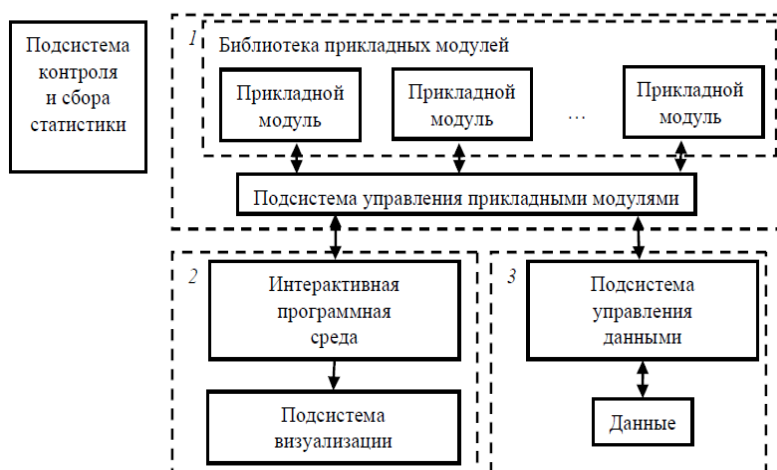


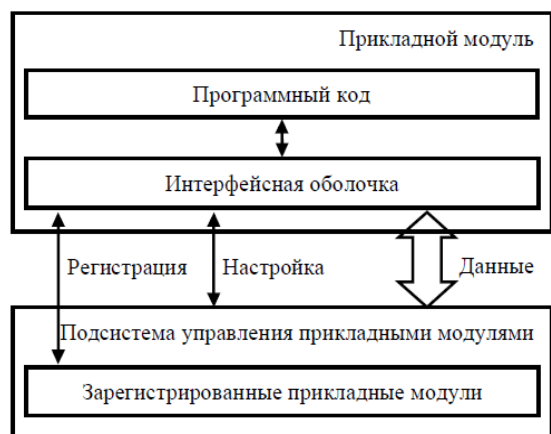
Рис. 5. Модульная структура программного средства [4]

В предложенной архитектуре выделяются три основные части, что позволяет разделить между собой систему ввода/вывода 3, интерактивную программную среду 2 и библиотеку прикладных модулей 1, что улучшает масштабируемость, сопровождаемость и надежность системы в целом. Модульная архитектура обеспечивает гибкость и расширяемость системы без существенного увеличения трудозатрат на ее реализацию.

Учитывая специфику предметной области, авторы выделяют следующие группы модулей: модули для работы с базами изображений; модули предобработки изображений; модули экстракции/селекции признаков; модули классификации; модули отображения результатов.

Все прикладные модули реализуют единообразный интерфейс взаимодействия кода модуля с другими подсистемами (рис. 6). Подсистема управления прикладными модулями осуществляет настройку модулей и предоставляет доступ к их функционалу, взаимодействует с интерактивной программной средой.

Очень важным при проектировании FaReS является оценка качества ее работы на конкретных данных. Два наиболее важных показателя: FAR (False Acceptance Rate) – процентный порог, определяющий вероятность того, что один человек может быть принят за другого (коэффициент ложного доступа) и FRR (False Rejection Rate) – вероятность того, что человек может быть не распознан системой (коэффициент ложного отказа в доступе). Однако для более полного анализа протокол результата моделирования должен сохранять информацию о структуре исходных данных, визуальную и текстовую информацию о проведенном процессе распознавания, реакцию системы (результат распознавания) на каждый предъявленный тестовый образ. Для визуализации данных должны исполь-



зоваться таблицы, 3D-графика, различного рода проекции, позволяющие оценить влияние используемого признакового пространства, его размерности, структуры системы на качество распознавания. Обязательно наличие временных оценок, которые позволяют сделать вывод о возможности работы в режиме реального времени.

Рис. 6. Интерфейс взаимодействия кода модуля с другими подсистемами [4]

Всех пользователей средств для моделирования FaReS можно условно разделить на три категории: студенты, которым необходимо изучить основы систем идентификации, инженеров, которые занимаются разработкой реальных систем, предназначенных для обеспечения безопасности, и ученых, основной задачей которых является разработка новых более эффективных и быстрых алгоритмов.

С этой точки зрения для инженеров наиболее актуальным будет следующее: возможность использования при моделировании видеозаписей, возможность получения готового программного продукта на основе смоделированной системы, т.е. фактически для них необходим модуль, который формирует уникальное приложение для отдельного заказчика и разработанного для конкретной базы данных с учетом всех ее особенностей – освещения, положения камеры и т.д.

Все FaReS можно разделить на простые, использующие быстрые алгоритмы экстракции признаков, не требующие больших объемов вычислений, и сложные – использующие методы редукции размерности пространства признаков, требующих существенных затрат вычислительных и временных ресурсов.

В настоящее время основным направлением обеспечения безопасности становится применение новых систем наблюдения, использующих значительное количество (от 30 и более) цифровых видеокамер. В средствах массовой информации регулярно появляются сообщения о разработке или испытании систем городского видеонаблюдения – сети из сотен/тысяч камер, установленных в больших городах на улицах и площадях, в аэропортах и на вокзалах, а также в общественном транспорте.

Для анализа информации, поступающей с большого количества камер, необходима разработка новых быстродействующих систем распознавания, которые смогут обеспечить работу в режиме online, ориентированных на большие объемы исходных данных, а также баз с быстрой динамикой изменения их состава.

Реализация данного рода систем возможна только при условии применения аналитики на видеокамере [10], т.к. стандартный средний поток, выдаваемый IP-камерой, 2 мбит/с. Если камера мегапиксельная, то уже оперируют потоками более чем в 10 мбит/с, причем количество камер может быть и 100, и 1000, а сетевое оборудование не может работать со 100%-й загрузкой (для видеопотоков максимум 50%). Существует еще одна причина, по которой видеоаналитика должна быть установлена на источнике видеосигнала (на IP-камере или на кодере), – любая компрессия с целью сокращения передаваемого объема данных ухудшает качество изображения, а значит, и качество работы системы идентификации.

Поскольку вычислительные ресурсы видеокамер на данный момент ограничены, то реализовать вычислительно сложные алгоритмы в самих видеокамерах не представляется возможным, однако использование в них простых алгоритмов позволит существенно уменьшить загрузку сети, используя для более точного и сложного анализа только критически важную информацию, что позволит не только повысить уровень безопасности объекта, но и снизить стоимость системы. Следует отметить, что в данном случае FaReS (или некоторая ее подсистема) будет иметь параллельную структуру.

Объем информации, хранящейся в базах данных, постоянно растет, и поиск в них нужной информации становится все более сложным, поэтому и в этом случае необходимо использовать более гибкую структуру системы – комбинацию простых и сложных методов, поскольку время поиска становится все более значительным. Решение данной проблемы предложено в [11].

Применение унимодальных систем показало, что у них есть ряд недостатков: увеличение сложности вычислений, падение качества распознавания при росте количества распознаваемых людей, относительно слабая защищенность против фальсификаций. Поэтому наиболее перспективным решением ближайшего будущего является применение мультимодальных биометрических систем. Использование различной информации об одном и том же лице человека «Multisampling / Multi-imaging» (представление изображения лица одного и того же человека через отдельные его фрагменты) [12] и «Multirepresentation» (представление изображения лица одного и того же человека в форме различных признаков: яркостных, гистограммных, спектральных или в форме изображений, полученных различными способами: в видимом свете, при инфракрасном освещении и т.д.) [13, 14] – один из самых привлекательных способов повышения эффективности биометрических систем. Однако за рост качества распознавания разработчикам приходится платить усложнением структуры системы.

Все вышеизложенное делает очень важным возможность программного средства для моделирования FaReS проектировать системы, имеющие каскадные, параллельные или комбинированные структуры, а также комплексировать информацию на разных уровнях. Отсюда же вытекает и необходимость наличия нескольких классификаторов для моделирования FaReS – с помощью функций расстояния, Байесовский, Ada-Boost, голосование и т.д.

Заметим, что исследователи, занимающиеся разработкой новых методов идентификации, при использовании моделирования более других категорий пользователей заинтересованы в наличии богатого инструментария для анализа результатов экспериментов – им необходимы самые продвинутые средства визуализации: возможность построения проекций, трехмерных графиков, таблиц. Кроме того, им необходим модуль, который освободит их от рутинных операций по непрерывному изменению параметров в некоторых пределах с целью поиска оптимальных значений, т.е. должна быть возможность описания вычислительного эксперимента для заданной структуры и определения критерия качества полученных результатов.

Заключение. Анализ существующих специализированных средств с точки зрения изложенных выше требований показывает, что в наибольшей степени им удовлетворяет пакет FaReS-MOD, который позволяет, не прибегая к классическому программированию, быстро поставить эксперимент, исследовать систему распознавания лиц и сравнить ее результативность с другими, ранее созданными системами. Однако необходимо дополнение его рядом модулей, в частности, позволяющих строить FaReS, имеющих более сложную и гибкую структуру.

В статье сформулированы и обоснованы требования к средствам моделирования систем распознавания изображений лиц, принципы их построения и анализа. Предложена архитектура пакета, приведены состав подсистем и их назначение, что позволит синтезировать системы распознавания человека по изображению лица, имеющие различную архитектуру; существенно сократить затраты и снизить требования к квалификации экспертов при проектировании систем для решения многих практических и исследовательских задач биометрии; осуществлять тестирование и параметрическую оптимизацию систем распознавания изображений лиц с использованием стандартных и специальных баз данных; выполнять тестирование новых методов обработки изображений и сравнение их эффективности с существующими; исследовать перспективные варианты построения систем и оценивать или прогнозировать их точностные характеристики.

Литература

1. Пентланд А. Распознавание лиц для интеллектуальных сред // Открытые системы. – 2000. – № 3. – С. 17 – 20.
2. Обработка и анализ изображений в задачах машинного зрения / Ю.В. Визильтер, С.Ю. Желтов, А.В. Бондаренко и др. – М.: Физматкнига, 2010. – 672 с.
3. Кухарев Г.А. Системы распознавания человека по изображению лица / Г.А. Кухарев, Н.Л. Щеголева. – СПб.: Изд-во СПбГЭТУ «ЛЭТИ», 2006. – 176 с.
4. Хомяков М.Ю. Принципы построения пакета программ для моделирования систем распознавания изображений лиц / М.Ю. Хомяков, Г.А. Кухарев // Известия СПбГЭТУ «ЛЭТИ». – 2010. – № 7. – С. 41–46.
5. NI Vision Assistant Tutorial [Электронный ресурс]. – Режим доступа: <http://www.ni.com/pdf/manuals/372228a.pdf>, свободный (дата обращения: 21.01.2013).

6. Kukharev G. An environment for recognition system modeling / G. Kukharev, A. Kuzminski // Enhanced methods in computer security, biometric and artificial intelligence systems. – New York; USA: Kluwer Academic Press, Springer, 2005. – P. 157–164.
7. Kukharev G. Strategia budowy systemow klasy «Face Retrieval» / G. Kukharev, M. Mikiasz, B. Taudal // Metody Informatyki Stosowanej. – 2007. – Т. 12. – S. 61–72.
8. Щеголева Н.Л. Программный комплекс для моделирования систем распознавания изображений лиц // Матер. XIV Междунар. конф. по мягким вычислениям и измерениям (SCM'2011). – СПб. – 2011. – С. 271–274.
9. Kukharev G. FaRetSys – the environment for modeling Face Retrieval tasks / G. Kukharev, M. Miklasz, B. Tauda // Metody Informatyki Stosowanej. – 2007. – Т. 12, № 2. – S. 61–73.
10. Видеоаналитика: потребности и возможности // Журнал ТЗ. – 2010. № 2 [Электронный ресурс]. – Режим доступа: <http://www.tzmagazine.ru/jpage.php?uid1=516&uid2=543&uid3=546>, свободный (дата обращения: 21.01.2013).
11. Кухарев Г.А. Поиск изображений лиц в больших базах данных // Мир измерений. – 2009. – № 4. – С. 22–30.
12. Price J.R. Face recognition using direct, weighted linear discriminant analysis and modular subspaces / J.R. Price, T.F. Gee // Pattern Recognition. – 2005. – № 38. – P. 209–219.
13. Rajagopalan A.N. Face recognition using multiple facial features / A.N. Rajagopalan, K. Srinivasa Rao Karri, Y. Anoop Kumar // Pattern Recognition Letters. – 2007 – № 28. – P. 335–341.
14. Kamenskaya E. Some aspects of automated recognition of person's psychological characteristics from the facial image / E. Kamenskaya, G. Kukharev // Metody informatyki stosowanej. – 2008. – Vol. 15, №2. – P. 29–37.

Щеголева Надежда Львовна

Канд. техн. наук, доцент каф. математического обеспечения ЭВМ СПбГЭТУ, г. Санкт-Петербург

Тел.: 8 (812) 234-26-82

Эл. почта: NLSchegoleva@etu.ru

Shchegoleva N.L.

Development of software for face recognition system modeling

In the work we considered functions, composition and structure of a software tool designed to graphical modeling of face recognition systems. We developed a software tool which allows you to investigate methods, algorithms and structures of a processing system.

Keywords: face recognition, domain-specific modeling.

УДК 007.61, 681.3.01

М.Ю. Катаев, Л.В. Широков

Методика определения жестов руки, наблюдаемых с помощью видеокамеры

Предложена методика определения жестов руки человека из изображений, получаемых с помощью видеокамеры. Приведены обзоры устройств и алгоритмов, где реализовано выделение жестов на изображении. Поясняется особенность предлагаемого подхода на фоне известных и приводятся первичные результаты обработки изображений и выделение жестов.

Ключевые слова: изображение, методы обработки изображений, жесты руки.

В настоящее время разработка и исследование взаимодействия человек–компьютер (человеко-машинные интерфейсы) приобретают особое внимание благодаря развитию математического и программного обеспечения. Одной из центральных задач этого взаимодействия является разработка и применение естественного для человека способа общения с компьютером, чтобы ускорить общение, сделать его более надежным и расширить круг решаемых компьютером задач и круг пользователей. К естественным способам общения можно отнести жесты (рук, головы, тела), голоса, мимики и др.

Одним из возможных вариантов управления комплексом оборудования являются статические и динамические жесты рук. Существуют алгоритмы, позволяющие выделять на изображении руку человека и распознавать жесты. К прямым аналогам можно отнести продукт компании Leap Motion, систему управления с помощью жестов – Leap 3D. Другой вариант – контроллер движений Kinect, разработанный в компании Microsoft. Приложение NPointer позволяет управлять компьютером без использования мыши. Программная система Flutter позволяет управлять музыкой и видео на Mac при помощи жестов. Фирма Pioneer разработала программно-аппаратный комплекс «Air Gesture» для управления аудиовидеооборудованием фирмы с помощью жестов. Компания «Атанор» разработала систему распознавания жестов для проведения презентаций. Система распознает жесты человека с помощью видеокамеры и специальных сенсоров глубины и позволяет ведущему управлять показом слайдов руками. Израильская компания eyeSight Mobile Technologies совместно с корейским производителем Innodigital разработала первую в мире телевизионную приставку под управлением операционной системы Android, которая оснащена интегрированной технологией распознавания жестов. Инновационный продукт получил имя nTobeBox. Японская компания NEC разработала систему управления жестами, которая использует детектор движения и проектор для отображения виртуального «пульта управления».

Помимо указанных выше разработок в сфере управления устройствами через жесты рук, которые считываются при помощи видеоустройств, показывают, что существует множество разработок, которые являются экспериментальными. Проведенный обзор [1–12] по данному направлению показывает, что данный вид взаимодействия человек–компьютер является перспективным как в научном плане, так и в коммерческом. Отметим, что указанные выше устройства имеют весьма малое распространение за счет привязанности к соответствующим техническим устройствам.

Развитие подходов детектирования и распознавания жестов особенно важно для таких устройств, как, например, STB (Set-Top-Box) и другие устройства, когда на первый план встает ограниченность в памяти и производительности вычислительного процессора. В данной статье приводится описание простой и легко реализуемой методики выделения статических жестов руки на изображении.

Типы жестов руки. На рис. 1 приведены некоторые из статических жестов, которые возможно диагностировать из анализа видео. Видно, что для диагностики жестов необходимо фиксировать положение руки и пальцев, при этом возникает множество однотипных ситуаций, которые могут быть представлены как жест, однако не будучи таковым.

При анализе жестов необходимо понять, какие может принимать положения и формы кисть, относительно стоящего/сидящего человека. Кисть практически не гнется в запястье в сторону большого пальца, в сторону мизинца гнется примерно на 70–90°. В сторону тыльной и внутренней стороны ладони кисть гнется (может гнуться, и эта величина зависит чаще всего от возраста) на 90°.

Локтевой сустав позволяет повернуть ладонь примерно на 180° в сторону большого пальца. Вращение кисти относительно линии, проходящей через центр ладони и перпендикулярной ей, производится на 270° . Если повернуть тыльной стороной к себе, то поворот осуществляется от 0 до 270° против часовой стрелки. Если повернуть внутренней стороной к себе, то поворот осуществляется от 90° до -180° по часовой стрелке. Кисть имеет пять пальцев (это определяет набор жестов и соответствующих алгоритмов распознавания).



Рис. 1. Положения руки и пальцев во время отображения жестов

Рассмотрим возможные комбинации с пальцами (см. рис. 1). Если ладонь сжать в кулак и по отдельности распрямлять каждый палец, то получится шесть вариантов: кулак, открыт большой, указательный, средний, безымянный с мизинцем, мизинец. Далее можно собирать различные комбинации, при некоторых комбинациях распрямляется третья фаланга сжатых пальцев (если это позволяют кистевые суставы). Конечно, растяжение связок и вращение в суставах индивидуально, так что здесь приведены усредненные данные. Для пальцев существует 2 позиции: сжат, разогнут. Так же углы поворота выбираются кратными 90° (это делается для облегчения обработки данных).

Итог: жест определяется положением ладони относительно видео устройства (вращением в различных плоскостях) и сгибанием/сжатием пальцев.

При детектировании жестов возникают вопросы, которые должен решать разработчик.

1. *Продолжительность времени, когда определенное положение руки и пальцев считается жестом.* Имеется в виду время фиксации руки в статичном положении, для определения жеста. Человек во время работы с устройством может выполнять произвольные или непроизвольные жесты-кулачки. Они могут возникать при общении с другим человеком выражением эмоций либо при совершении других действий, не связанных со взаимодействием с устройством, что выражается в непреднамеренных изменениях положения руки и пальцев (например, дрожание).

2. *Выбор видеоустройств.* На продолжительность времени сказывается и несовершенство воспринимающих датчиков, в нашем случае видеокамеры. Такие камеры имеют порог скорости восприятия информации (типично 25–30 кадров/с), что не позволяет уследить за всеми движениями руки (отдельные части руки могут двигаться с большей скоростью, чем другие (ситуация, когда локоть практически статичен, а ладонь движется быстро)). В такой ситуации при съемке получаем размазанный жест, который достаточно сложно интерпретировать. Конечно, можно было бы использовать дорогие камеры скоростной съемки, но система задумывалась изначально для работы с типичными камерами и их аналогами.

3. *Вычислительные алгоритмы.* Сложность алгоритмов и низкое быстродействие элементов вычислительной системы не всегда позволяют выполнить действие за доли секунды и требуют определенного времени, которое зависит от разрешения изображения, от качества съемки и других параметров.

Рассмотрим период времени, в котором человек должен зафиксировать определенный жест. Это время, по нашему мнению, лежит в интервале от 0,5 до 2 с или в среднем 1–1,5 с на жест (около 30–45 кадров при скорости съемки 30 кадров/с). В течение 1–1,5 с устройство должно провести не только съемку движения руки, но и принять решение: человек выполнил просто движение рукой или хочет выполнить команду с помощью определенного жеста.

Почему время детектирования жеста не больше 2 с? Ответ связан с физиологией человека. В теле человека существует движение крови с частотой примерно 1,2–1,5 Гц, что сказывается на статичности жестов, т.е. как бы человек не хотел удержать руку в одном положении, пульс будет создавать колебания, которые будут смещать руку. Также чем больше человек будет удерживать руку в одном положении, т.е. постоянно напрягать одну группу мышц, то в этих мышцах будет создаваться тремор, интенсивность и амплитуда которого пропорциональны времени. Свой вклад вносит и дыхание человека – при вдохе грудная клетка и брюшная полость расширяются, в момент выдоха наоборот сжимаются, подобные колебания передаются во все части тела.

Алгоритмы выделения руки человека. Метод гистограмм. Метод заключается в построении двух гистограмм изображения и их анализа. Отсекается фоновое изображение. Изображение переводится в полутоновое изображение (Gray Scale). Строится две гистограммы (по оси X и Y) оставшегося изображения руки. Принцип построения гистограммы: изображение разбивается на N частей (по горизонтали или вертикали, в зависимости от необходимой гистограммы), в каждой части получаем среднее значение оттенков серого по всем пикселям области. Нормализуем значения всех областей по максимальному значению из всех областей, далее проводим анализ гистограммы и, если надо, выводим ее на экран для наглядности.

Метод зависит от освещения, методов фильтрации, следовательно, система обработки должна иметь хорошую приспособляемость к изменению внешней среды. В некоторых случаях достаточно сложно отследить логику изменения гистограммы при изменении исследуемых частей изображения (положения руки в пространстве, относительно камеры), следовательно, анализ гистограммы предоставляется слишком трудоемкой задачей.

Метод вычисления площади. В этом подходе строится функция площади выделенного контура руки с различным количеством пальцев. Контур должен быть замкнутым. Данная функция позволяет достаточно эффективно отделять жесты один от другого, но существенно зависит от мешающих факторов (освещенность, контраст, элементы не удаленного фона). Единственное ограничение метода связано с тем, что с его помощью можно определить лишь шесть жестов.

Контурный метод. Поиск контура сводится к поиску изменения яркости полутонового изображения, также поиск контура возможен по поиску цветовых перепадов на изображении. При контурном анализе изображения можно выделить два аспекта:

1. Удаление лишних объектов (фона) с изображения, чтобы они не вносили искажение при поиске контуров (фильтрация, удаление шумов и пр.).
2. Настройка фильтра поиска контура так, чтобы незначительные детали (текстура и внутренние особенности объекта, если они нам не нужны) не выделялись фильтром поиска контуров.

При контурном анализе могут возникать проблемы: разрывность контуров, необходимо дополнительно соединять разделенные части контура, чтобы упростить сам анализ; создается проблема выделения ненужных нам контуров (остатки фона или ненужных деталей искомого объекта). Также затруднительно проводить сложный анализ изображения (определение жестов) по контуру объекта (руки). Контурный анализ чаще всего проводится для определения относительных размеров объекта, формы, сложности или проводится прямое сопоставление с шаблоном.

Анализ контура относительно центра тяжести объекта. Данный метод включает в себя предыдущие два метода (метод гистограмм и контурный метод). Анализируя гистограммы изображения, находим центр тяжести объекта. Получаем контур объекта. Строим функцию расстояния точки контура относительно центра тяжести объекта. Этот метод представляется менее трудоемким, чем представленные выше методы. При визуализации функции расстояния довольно просто увидеть логику изменения минимумов и максимумов при изменении исследуемых частей изображения (положения руки и пальцев в пространстве, относительно камеры).

Результаты работы алгоритма. Нами применялись различные численные алгоритмы для устойчивого выделения руки на изображении: фиксация центра тяжести бинарного изображения руки, выделение контура руки, достижение условия его замкнутости (при помощи морфологических операций). Также нами был организован поиск алгоритмов устойчивого детектирования положения пальцев на руке, для различных пространственных положений руки. Среди всех возможных алгоритмов нами был отобран простой, но эффективный алгоритм, который устойчив в работе, требует минимального объема памяти и быстр в работе. Идея связана с вычислением функции расстояния от контура фигуры до центра тяжести. Тогда распознавание отдельных жестов, как и отличие одного жеста от другого, производится путем анализа этой функции расстояния. Анализируя экстремумов функции расстояния, возможны распознавание типа жеста и сравнение жестов.

$$d(i) = \sqrt{(x_c - x_i)^2 + (y_c - y_i)^2},$$

где i – текущий номер точки контура; $\{x_c, y_c\}$ – координаты центра тяжести изображения руки и $\{x, y\}$ – координаты контура, начиная с нижней части контура, перпендикулярной центру тяжести.

На рис. 2 показаны фотографии руки с изображением определенного числа пальцев (рис. 2, а), контур выделенной на фоне руки (рис. 2, б) и график функции расстояния d (рис. 2, в). На рис. 2, в хорошо видно, что количество пиков, которые выделяются при помощи нашего подхода однозначно

определяют число пальцев на руке и тем самым отделяют один жест от другого. Отметим, что первый пик на изображении рис. 2, *в* соответствует левому нижнему краю руки и всегда детектируется мешающий и не учитывается при дальнейших расчетах. Поиск числа пальцев на изображении сводится к простейшей задаче определения числа максимумов кривой. Таким образом, анализ кривой расстояния d позволяет не только провести выделение типа жеста, но и легко построить алгоритм распознавания жеста и определения той или иной команды для устройства.

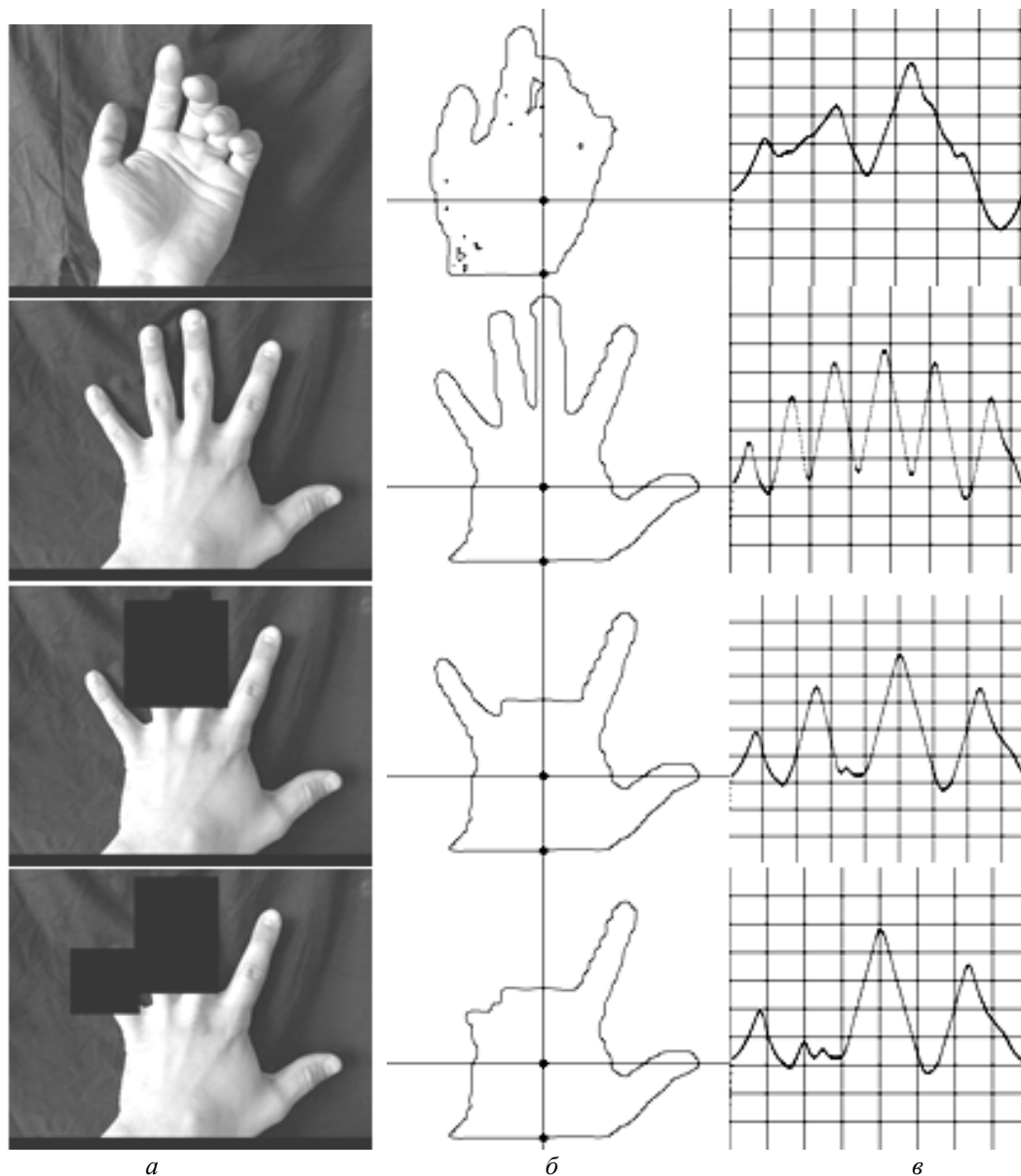


Рис. 2. Анализ жестов. *а* – изображение руки; *б* – бинарный контур руки с определенным центром тяжести и *в* – рассчитанная функция расстояния d

В настоящее время нами получены результаты, позволяющие выделять жесты руками в статическом режиме. Динамический, т.е. произвольный, режим движения руки и распознавания жестов является пока в состоянии разработки.

Заключение. Разработан и рассмотрен на примере распознавания статических жестов метод определения числа раскрытых пальцев на руке. Подобраны алгоритмы обработки изображений, позволяющие устойчиво проводить детектирование положения руки и кисти на изображении. Обработка реальных снимков показала высокую надежность обнаружения числа пальцев на руке и их последовательность открытия. В дальнейшем этот метод планируется использовать для обнаружения и распознавания динамических жестов.

Литература

1. Системы компьютерного зрения [Электронный ресурс]. – Режим доступа: <http://beta.iitvision.ru/?p=173>, свободный (дата обращения: 04.04.2013).
2. Гонсалес Р. Цифровая обработка изображений / Р. Гонсалес, Р. Вудс. – М.: Техносфера, 2006. – 1072 с.
3. Zhou H. Tracking articulated hand motion with Eigen dynamics analysis / Y. Zhou, T.S. Huang // Proc. of the Intern. conf. on computer vision, Nice (France), 14–16 Oct. 2003. – Washington DC: IEEE Computer Soc. – 2003. – Vol. 2. – P. 1102–1109.
4. Wu Y. Non-stationary color tracking for vision-based human computer interaction / Y. Wu, T.S. Huang // IEEE Trans. Neural Networks. – 2002. – Vol. 13, №4. – P. 948–960.
5. Manresa C. Hand tracking and gesture recognition for human computer interaction / C. Manresa, J. Varona, R. Mas, F.J. Perales // Electron. Lett. Comput. Vision Image Anal. – 2005. – Vol. 5, №3. – P. 96–104.
6. Mckenna S. A comparison of skin history and trajectory-based representation schemes for the recognition of user-specific gestures / S. Mckenna, K. Morrison // Pattern Recognition. – 2004. – Vol. 37. – P. 999–1009.
7. Chen F. Hand gesture recognition using a real-time tracking method and hidden Markov models / F. Chen, C. Fu, C. Huang // Image Vision Comput. – 2003. – Vol. 21, №8. – P. 745–758.
8. Ng C.W. Gesture recognition via pose classification / C.W. Ng, S. Ranganath // Proc. of the 15th Intern. conf. on pattern recognition, Barcelona (Spain). – 2000. – Vol. 3. – P. 699–704.
9. Oka K. Real-time fingertip tracking and gesture recognition / K. Oka, Y. Sato, H. Koike // Proc. IEEE. Comput. Graphics Appl. – 2002. – Vol. 22, №6. – P. 64–71.
10. Huang C. A model-based hand gesture recognition system / C. Huang, S. Jeng // Machine Vision Appl. – 2001. – Vol. 12, №5. – P. 243–258.
11. Cutler R. View-based interpretation of real-time optical flow for gesture recognition / R. Cutler, M. Turk // Proc. of the 3rd IEEE conf. on face and gesture recognition, Nara (Japan), 14–16 Apr., 1998. – P. 416–421.
12. Lu S. Using multiple cues for hand tracking and model refinement / S. Lu, D. Metaxas, D. Samaras, J. Oliensis // Proc. of the IEEE conf. on computer vision and pattern recognition, Madison (USA). – 16–22 June, 2003. – P. 443–450.

Катаев Михаил Юрьевич

Д-р техн. наук, профессор каф. автоматизированных систем управления (АСУ) ТУСУРа,
профессор Юргинского технологического института (филиала НИ ТПУ)
Тел.: 8-960-975-27-85, (382-2) 70-15-36
Эл. почта: kataev.m@sibmail.com

Широков Леонид Владимирович

Аспирант каф. АСУ ТУСУРа
Тел.: 8-962-778-33-67, (382-2) 70-15-36
Эл. почта: ngtrx89@gmail.com

Kataev M.Yu., Shirokov L.V.

The methodology of determining the hand gestures the observed with the camcorder

A procedure is proposed for determination of the hand gestures from the data processing of the images received with the help of camcorder. Provides reviews of devices and algorithms, which implements the allocation of gestures in the image. Explains the features of the proposed approach, on the background of well-known and are the primary results of processing of images and the selection of gestures.

Keywords: image, image processing methods, the gestures of the hand.

УДК 004.5

В.Ю. Будков

Методы и программные средства обработки мультимедийных данных при сопровождении распределенных совещаний

Рассмотрены методы и программно-аппаратные средства автоматического анализа аудиосигналов, записанных в ходе сопровождения распределенных совещаний и применяемых для генерации отчетных материалов по результатам мероприятия.

Ключевые слова: окружающее интеллектуальное пространство, сопровождение совещаний, телеконференции, диаризация дикторов, протоколирование мероприятий.

Основные проблемы систем сопровождения распределенных совещаний. Современные системы связи и совместной работы не позволяют полностью автоматизировать процесс информационной поддержки проведения совещаний, поэтому большая часть работы по сопровождению удаленных участников выполняется операторами-людьми. Кроме того, при проведении деловых встреч, заседаний, совещаний и других формальных мероприятий обязательной процедурой является протоколирование выступлений участников [1, 2]. Однако экспертный анализ и расшифровка аудиозаписей совещаний требуют привлечения специалистов-стенографистов и занимают длительное время обработки [3]. Современные методы анализа речи и диаризации дикторов позволяют автоматизировать процесс выделения реплик участников совещания. Одним из перспективных способов увеличения эффективности систем диаризации является применение многоканального и многомодального анализа поведения участников в зале совещаний [4, 5]. Проведенный анализ комбинированных систем диаризации показал, что использование дополнительных признаков, извлекаемых при анализе изображений и локализации источника звука, позволяет повысить точность определения момента смены диктора [6, 7].

Другим ограничением систем телеконференций являются пропускная способность коммуникационных сетей и мультимедийные возможности клиентского устройства, которые существенным образом влияют на вид пользовательского интерфейса и выбор информационных каналов, доступных для удаленных участников. Ключевым вопросом при дистанционной коммуникации является высокая неопределенность о ситуации в удаленной аудитории, вызванная пространственно-временными различиями [8]. Физические и психологические барьеры препятствуют удаленному участнику быстро присоединиться к дискуссии проблемы, обсуждаемой участниками внутри зала, и тем более предложить новое направление разговора.

Анализ направлений исследований по изучению поведения человека в интеллектуальном пространстве и развитию интуитивных многомодальных интерфейсов показал перспективность дальнейшей разработки методов и программных средств аудиовизуальной поддержки мобильных телеконференций, отличающихся применением средств автоматического анализа и оценки информационной значимости передаваемого контента, а также передачи оптимизированного аудиовизуального потока данных, обеспечивающих снижение когнитивной нагрузки на пользователя и уменьшение потребляемых ресурсов мобильным устройством [3, 9].

Таким образом, основной задачей данного исследования является разработка математических и программных решений, повышающих возможности удаленного участника при принятии решений и участии в дискуссиях во время распределенных мероприятий, а также снижении затрат на подготовку мультимедийных отчетных материалов.

Модель сопровождения участников распределенных мероприятий. Сопровождение мероприятий включает в себя три основных этапа: подготовка к совещанию, запись и трансляция, затем архивирование и анализ записанных материалов. На втором и третьем этапах для анализа многоканальных потоков мультимедийной информации, поступающей от распределенных участников мероприятия, требуется привлечение средств автоматической обработки данных. Причем более жесткие требования по времени обработки предъявляются при трансляции мероприятий, а архивирование и

подготовка протокола мероприятия могут быть выполнены после завершения совещания. На рис. 1 представлены основные методы обработки мультимедийных сигналов, применяемых для трансляции мероприятий и последующей подготовки отчета в предложенной модели сопровождения мероприятий.

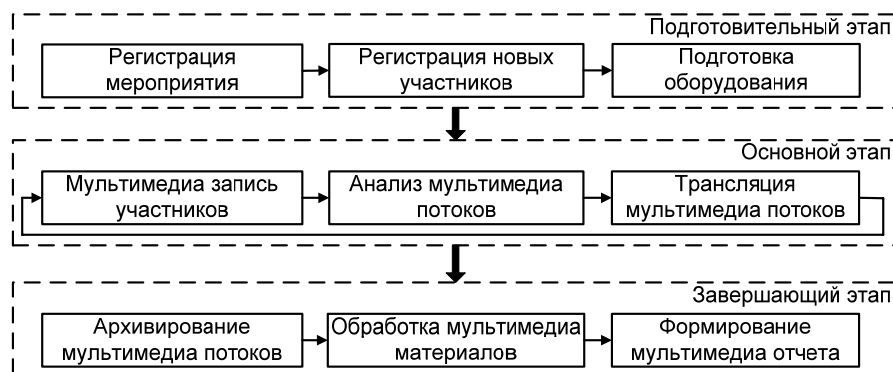


Рис. 1. Обработка мультимедийных сигналов в системе сопровождения мероприятий

Подготовительный этап включает в себя регистрацию мероприятия, подготовку оборудования и регистрацию участников. Вначале в систему сопровождения мероприятия вносятся данные о времени и месте проведения, разрешениях на доступ к трансляции мероприятия, возможности регистрации посторонних участников и др. После этого участники могут зарегистрироваться в системе и авторизоваться для участия в данном мероприятии. Перед началом мероприятия производится подготовка и настройка необходимого мультимедийного оборудования. Основной этап включает в себя трансляцию и запись аудио- и видеопотоков и других мультимедийных данных, получаемых из различных источников, расположенных в помещении проведения мероприятия и от удаленных участников. При выборе наиболее актуальной информации, которая транслируется удаленным участникам, учитываются следующие аспекты: положение текущего выступающего, время смены слайда презентации, этап мероприятия и другие параметры. Завершающий этап сопровождения мероприятия включает в себя анализ мультимедийных записей, их архивирование и создание отчета по мероприятию. Генерация отчета производится по шаблонам, которые могут редактироваться вручную для получения необходимой формы представления.

Средства генерации отчетных материалов по результатам мероприятия. При формировании отчета по мероприятию производится формирование протокола, где отмечены реплики каждого из участников. Для реализации этой задачи был разработан метод диаризации дикторов в одноканальном аудиопотоке, включающий два основных этапа: цифровую обработку аудиосигналов и диаризацию дикторов. На рис. 2 приведена схема двухэтапной процедуры диаризации дикторов, где показаны блоки, выполняющие следующие функции: 1 – разделение аудиосигнала на речевые и неречевые фрагменты; 2 – выявление параметров речевого сигнала на основе моделей слухового восприятия человека; 3 – извлечение информативных признаков голосового источника речевого сигнала; 4 – выделение речевых артефактов, характерных для разговорной речи; 5 – оценивание отношения сигнал/шум в аудиосигнале; 6 – идентификация параметров голосового источника речевого сигнала независимо от контекста и языка; 7 – разделение речи различных дикторов.

Параметрическое представление речевого сигнала и временная разметка речевых фрагментов, вычисленных на этапе обработки аудиосигнала, используются при идентификации участников среди существующих моделей дикторов. Если диктор не идентифицирован или при отсутствии моделей, параметрическое представление текущего речевого фрагмента применяется для обучения модели нового диктора. После определения диктора текущему речевому фрагменту присваивается номер модели его диктора. Результат диаризации представляет собой временную разметку речевых фрагментов по дикторам.

Программная реализация предложенного метода диаризации речи дикторов, а также предложенные ранее методы анализа и распознавания речи были использованы при разработке программного комплекса автоматического анализа, распознавания и диаризации разговорной русской речи (ПАРАД-Р) [6, 10]. На рис. 3 представлена архитектура программного комплекса ПАРАД-Р, построенная на основе трехуровневой архитектуры (клиентская часть, серверная часть, программно-математическое ядро).

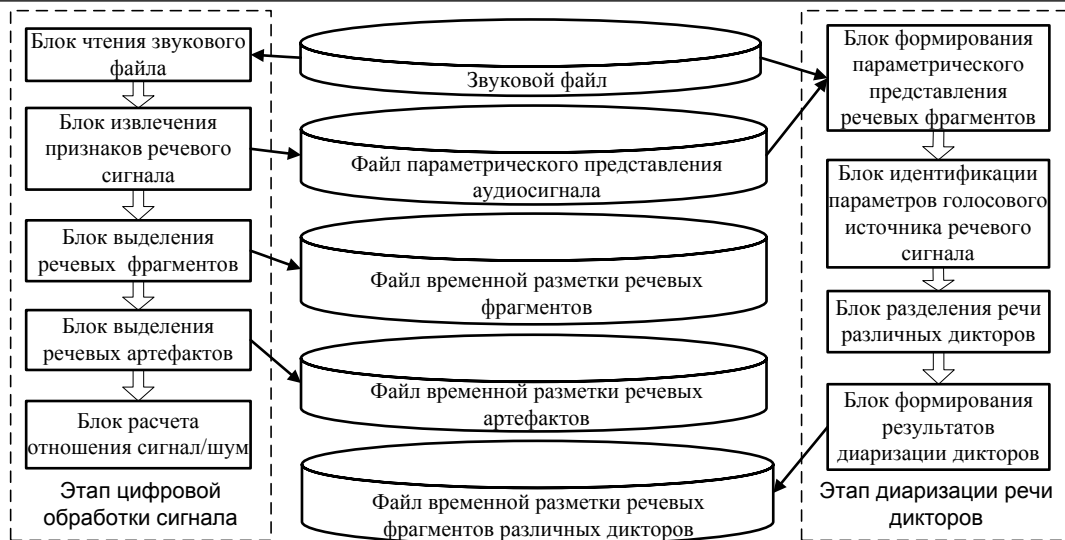


Рис. 2. Основные этапы аудиообработки при диаризации дикторов

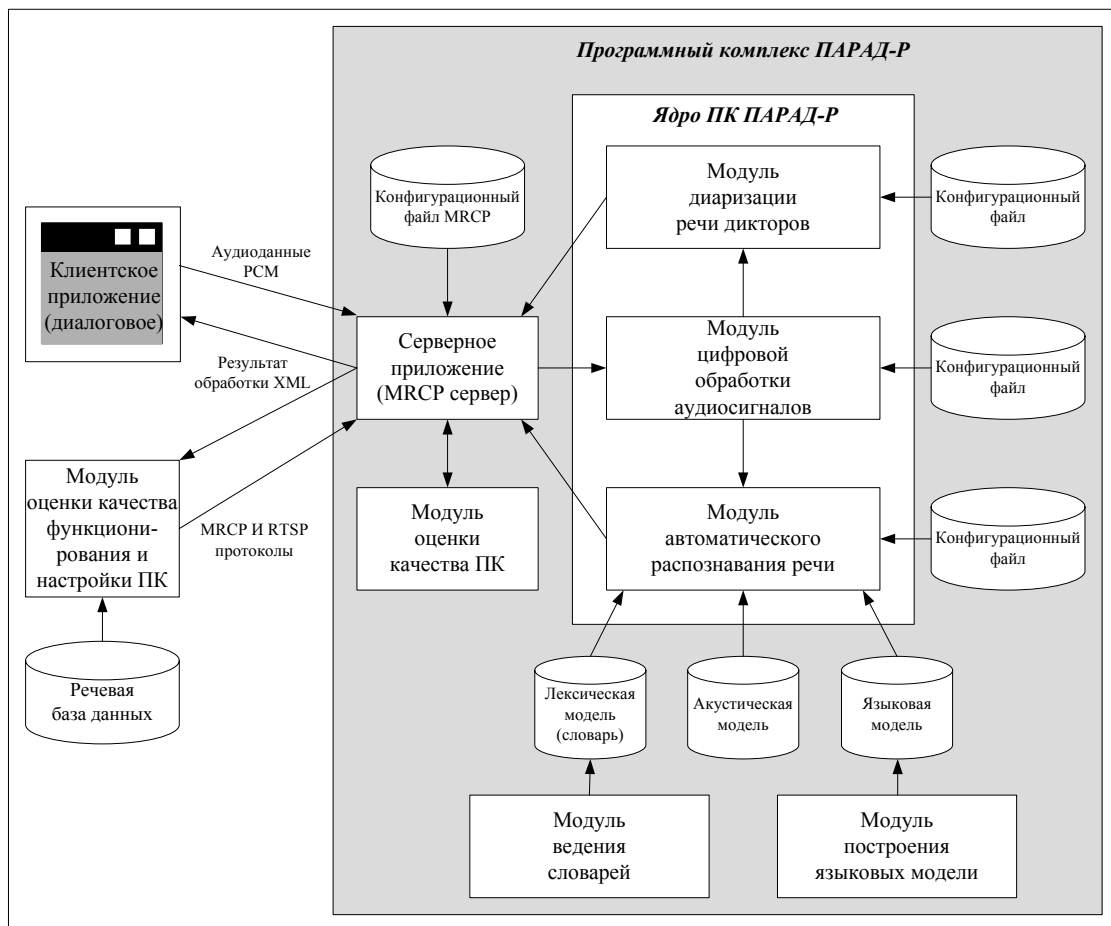


Рис. 3. Общая архитектура экспериментального образца ПАРАД-Р

Клиентская и серверная части могут располагаться как на одном компьютере, так и на разных компьютерах и взаимодействовать по компьютерной сети. Информационный обмен между клиентской частью и серверной частью реализован с использованием протокола MRCPv2 (Media Resource Control Protocol). Серверная часть состоит из следующих программных модулей:

- 1) серверное приложение – MRCP сервер;
- 2) модуль ведения словарей;
- 3) модуль построения языковых моделей;

4) модуль оценки качества ПК.

Каждый из этих модулей, кроме последнего, реализован в виде исполняемого файла, работающего под управлением ОС Microsoft Windows. Помимо этих программных модулей, серверная часть также имеет связь с программно-математическим ядром ПК, в который входят:

- 1) программная библиотека цифровой обработки аудиосигналов;
- 2) программная библиотека диаризации речи дикторов;
- 3) программная библиотека автоматического распознавания речи.

Оценка качества функционирования разработанного комплекса была проведена по методикам, учитывающим метрики WER (Word Error Rate), LER (Letter Error Rate), SWER (Speaker Attributed Word Error Rate) и DER (Diarisation Error Rate), с использованием созданного речевого корпуса слитной русской речи [11]. Разработанный метод диаризации дикторов в одноканальном аудиопотоке показал точность сегментации реплик разных дикторов свыше 85%. Для систем сопровождения распределенных совещаний комплекс может использоваться при подготовке протокола и стенограмм выступлений участников.

Заключение. Совокупность предложенных методов и программных средств автоматической обработки мультимедийных потоков данных, а также их практическая реализация представляют собой решение актуальной научно-технической задачи информационного и технологического сопровождения распределенных мероприятий на основе анализа текущей ситуации, оптимизации транслируемого контента удаленным участникам и генерации отчетных материалов по результатам мероприятия. Разработанный программный комплекс автоматического анализа, распознавания и диаризации разговорной русской речи поддерживает пакетную обработку аудиосигналов с доступом по стандартному протоколу MRCPv2 и может применяться для разработки кроссплатформенных приложений по распределению и управлению динамическими речевыми и многомодальными сервисами, в том числе по обработке архивных записей мероприятий.

Работа выполнена в рамках ФЦП «Исследования и разработки по приоритетным направлениям развития научно-технологического комплекса России на 2007–2013 годы» (ГК № 07.514.11.4139) и гранта РФФИ № 13-08-00741-а.

Литература

1. Ронжин А.Л. Проектирование интерактивных приложений с многомодальным интерфейсом / А.Л. Ронжин, А.А. Карпов // Доклады ТУСУРа. – 2010. – № 1 (21), ч. 1. – С. 124–127.
2. Мещеряков Р.В. Специализированная информационная система поддержки деятельности медицинского учреждения / Р.В. Мещеряков, Л.Н. Балацкая, Е.Л. Чойнзон // Информационно-управляющие системы. – 2012. – № 5. – С. 51–56.
3. Ронжин А.Л. Технологии поддержки гибридных е-совещаний на основе методов аудиовизуальной обработки / А.Л. Ронжин, В.Ю. Будков // Вестник компьютерных и информационных технологий. – 2011. – № 4. – С. 31–35.
4. Noulas A. Multimodal Speaker Diarization / A. Noulas, G. Englebienne, B.J.A. Krose // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2012. – Vol. 34(1). – P. 79–93.
5. Мещеряков Р.В. Сегментация и параметрическое описание речевого сигнала / Р.В. Мещеряков, В.П. Бондаренко, А.А. Конев // Изв. вузов. Приборостроение. – 2007. – Т. 50, № 10. – С. 3–7.
6. Ронжин А.Л. Особенности дистанционной записи и обработки речи в автоматах самообслуживания / А.Л. Ронжин, А.А. Карпов, И.А. Кагиров // Информационно-управляющие системы. – 2009. – Вып. 42, т. 5. – С. 32–38.
7. Ронжин Ал.Л. Формирование профиля пользователя на основе аудиовизуального анализа ситуации в интеллектуальном зале совещаний / Ал.Л. Ронжин, В.Ю. Будков, Ан.Л. Ронжин // Труды СПИИРАН. – 2012. – Вып. 23. – С. 482–494.
8. Yankelovich N. Porta-person: telepresence for the connected meeting room / N. Yankelovich, J. Kaplan, N. Simpson, J. Provino // CHI 2007. – 2007. – P. 2789–2794.
9. Ронжин Ал.Л. Система аудиовизуального мониторинга участников совещания в интеллектуальном зале / Ал.Л. Ронжин, Ан.Л. Ронжин // Доклады ТУСУРа. – 2011. – № 1 (22), ч. 1. – С. 153–157.
10. Ронжин А.Л. Топологические особенности морфофонемного способа представления словаря для распознавания русской речи // Вестник компьютерных и информационных технологий. – 2008. – № 9. – С. 12–19.

11. Будков В.Ю. Анализ современных методов и систем diarизации дикторов / В.Ю. Будков, А.Л. Ронжин // Изв. вузов. Приборостроение. – СПб.: ИТМО, 2011. – № 11. – С. 43–46.

Будков Виктор Юрьевич

Мл. науч. сотрудник лаб. речевых и многомодальных интерфейсов
Санкт-Петербургского института информатики и автоматизации Российской академии наук (СПИИРАН)
Тел.: 8 (812) 328-70-81
Эл. почта: budkov@iias.spb.su

Budkov V.Yu.

Methods and software of multimedia data processing at support of distributed meetings

Methods, software and hardware for automatic analysis of audio data recorded during support of distributed meeting and applied for creation of report materials of the meeting results.

Keywords: context-aware applications, ambient intelligent space, mobile heterogeneous devices.

УДК 621.396.41

М.П. Силич, А.Р. Вахитов, И.И. Прудников

Информационная система анализа энергетической эффективности территориальных образований

Рассматривается структура информационной системы анализа энергетической эффективности территориальных образований. Система позволяет строить иерархию когнитивных карт, отражающих связи между факторами, влияющими на состояние энергоэффективности, задавать функции принадлежности для нечеткой оценки факторов по значениям индикаторов, формировать правила и формулы для косвенного оценивания факторов. Для учета влияния внешних условий система выполняет нечеткую кластеризацию ТО. По результатам анализа может быть построено дерево причин.

Ключевые слова: энергетическая эффективность, территориальные образования, информационная система, когнитивные карты.

Экономика России на современном этапе характеризуется высокой энергоемкостью. Не случайно энергосбережение включено в список приоритетных направлений технологического развития страны. Основным инструментом такого развития служат программы в области энергосбережения и повышения энергоэффективности, которые должны быть разработаны во всех муниципальных образованиях на территории Российской Федерации. Для контроля и оценки ситуации на территории субъектов РФ разрабатываются информационные системы мониторинга и управления эффективностью энергосбережения [1–3]. Основными функциями подобных систем являются: ввод, хранение и редактирование данных об энергохозяйстве на уровне регионов, муниципальных образований, учреждений, а также мониторинг и анализ данных, внесенных в систему. Однако интерпретация пользователем формируемых отчетов зачастую вызывает трудности. Различие климатических условий, особенности инфраструктуры, сложившейся структуры потребления топливно-энергетических ресурсов, искажают результаты сравнения территориальных образований (ТО) между собой. Кроме того, для понимания ситуации необходимо учитывать, что одни показатели характеризуют созданные условия для реализации потенциала энергосбережения, другие – внешние условия, влияющие на освоение потенциала, третьи – достигнутый уровень энергоэффективности. Зачастую один и тот же фактор, определяющий эффективность использования энергетических ресурсов, может оцениваться на основе множества индикаторов, дополняющих друг друга, некоторые факторы вообще можно оценить лишь косвенно и весьма приблизительно. Для облегчения анализа показателей, собираемых в ходе мониторинга ситуации в сфере энергосбережения, информационная система должна выполнять интеллектуальную обработку данных, позволяющую оценить текущее состояние энергоэффективности территориальных образований в условиях неполноты исходных данных и неоднозначности их интерпретации, а также сформировать выводы в терминах естественного языка относительно причин сложившейся ситуации. Данная статья посвящена разработке информационной системы, обеспечивающей выполнение этих требований.

Методика анализа энергетической эффективности. В основу разрабатываемой системы положен подход, описанный в работах [4–6]. В соответствии с данным подходом анализ энергоэффективности территории включает следующие этапы.

Этап 1. Построение иерархии когнитивных карт. Карта представляет собой граф с двумя типами концептов: факторы отражают свойства системы, оцениваемые на качественном уровне, индикаторы являются измеримыми показателями, характеризующими тот или иной фактор. Примеры факторов: «уровень потребления тепловой энергии», «существенность расходов домохозяйств на электроэнергию», «суровость климатических условий»; примеры индикаторов: «удельный расход тепловой энергии в жилых домах, Гкал/м²», «доля расходов на электроэнергию по отношению к доходам населения», «средние расчетные температуры наружного воздуха за отопительный период, °С». Индикатор «привязывается» к фактору, который он характеризует, путем установления отношения ассоциации. Данному отношению может быть сопоставлен коэффициент соответствия (число

в интервале от 0 до 1). Чем выше значение коэффициента, тем точнее индикатор характеризует соответствующий фактор.

Среди множества факторов выделяют подмножества: целевых факторов, состояние которых позволяет оценить ситуацию в целом; управляемых факторов, состояние которых можно сознательно изменять; внешних возмущений, на состояние которых в рамках системы невозможно повлиять. Между факторами устанавливаются отношения влияния – причинно-следственные связи. Влияние может быть как положительным, так и отрицательным. Кроме того, сила влияния может быть оценена в виде числа от 0 до 1. Отношения влияния позволяют выявить причины достигнутого уровня того или иного фактора и оценить важность той или иной причины. Если влияющих факторов много, они могут быть вынесены на поддиаграммы – когнитивные карты нижнего уровня. Фактор, к которому «прикреплена» дочерняя карта, становится целевым на этой карте.

Этап 2. Определение способа оценивания факторов. Каждому фактору сопоставляется лингвистическая переменная, значениями которой являются нечеткие оценки типовых состояний фактора, например: «низкая», «средняя», «высокая». Чтобы оценить фактор на основе индикатора, строятся функции принадлежности на базовом множестве его значений. Функции могут строиться на основе распределения значений индикатора для всех сравниваемых территориальных образований. Параметры функций определяются исходя из плотности обучающих примеров.

Для того чтобы учесть влияние внешних факторов, отражающих природно-климатические, социально-экономические, градостроительные и другие особенности территориальных образований, предлагается использовать подход типологической группировки. Формируются кластеры территорий с похожими внешними условиями и функции принадлежности строятся отдельно для каждого кластера. Кластеризация проводится по индикаторам, «привязанным» к соответствующему внешнему фактору, на основе алгоритмов нечеткой кластеризации, в частности алгоритма Густафсона–Кесселя [7]. Предварительно рекомендуется осуществить выбор статистически независимых признаков с использованием нейронной сети с самоорганизацией корреляционного типа.

Оценки факторов, которым не сопоставлены индикаторы или нет возможности получить значения индикаторов, могут быть получены на основе оценок других факторов. Одним из возможных подходов является использование нечетких продукционных правил, содержащих в antecedенте возможные состояния оцененных факторов и в консеквенте – соответствующие состояния оцениваемого фактора. Альтернативным способом косвенного оценивания факторов являются операции с четкими или нечеткими числами, например процедура аккумуляции влияния нескольких концептов на один концепт, используемая в картах Коско, или операция свертки с нечеткими треугольными числами. В последнем случае используются процедуры преобразования нечетких значений лингвистических переменных в нечеткие числа, а также интерпретации вычисленного значения выходного фактора.

Этап 3. Анализ когнитивных карт. Анализ гибридной когнитивной карты позволяет дать оценку каждому фактору. При получении нечеткой оценки фактора по значению индикатора учитываются коэффициенты соответствия индикатора фактору. Оценки, полученные для одного фактора по нескольким индикаторам, агрегируются. Если фазсификация осуществляется с учетом внешних условий, степень уверенности в оценке корректируется, принимая во внимание степень принадлежности территории к соответствующему кластеру. В случае если территориальное образование попало одновременно в несколько кластеров, оценки, полученные для каждого кластера, агрегируются. Для получения оценок факторов, не связанных с индикаторами, выполняется логический вывод на множестве продукционных правил либо осуществляются преобразования в нечеткие числа с использованием универсальной шкалы и вычисления по формулам свертки.

По результатам анализа гибридной когнитивной карты может быть построено дерево причин. Оно показывает текущие состояния факторов, влияющих на энергоэффективность, для конкретного территориального образования и объясняет, почему, под влиянием каких причин сложились эти состояния.

Структура информационной системы. ИС включает три основных модуля (рис. 1): «Когнитивное моделирование», «Методы оценки факторов», «Анализ ситуации». Первые два модуля предназначены для создания проекта по анализу энергоэффективности территориальных образований заданного типа (например, субъектов РФ, входящих в Сибирский федеральный округ, или муниципальных образований Томской области), третий модуль используется для анализа ситуации с энер-

госбережением в конкретном территориальном образовании, а также сравнения оценок, полученных для различных ТО.



Рис. 1. Структура информационной системы

Модуль когнитивного моделирования отвечает за построение и редактирование иерархии когнитивных карт. Он представляет собой графический редактор специального назначения (рис. 2), формирующий диаграммы в форме векторного изображения, что позволяет просматривать их на мониторах с различным разрешением, растягивать и сжимать. Пользователь может добавлять / перемещать / удалять элементы карты, связывать элементы, создавать спецификации элементов, присоединять поддиаграммы. На диаграммах факторы изображаются в виде эллипсов (внешние ограничения – в виде эллипсов с пунктирным контуром), индикаторы – в виде круга, отношения ассоциации – в виде отрезков пунктирной линии, отношения влияния – в виде отрезков сплошной линии со стрелками. Если к фактору присоединена поддиаграмма, то рядом с символом соответствующего фактора помещается специальная пиктограмма. Иерархию поддиаграмм можно просмотреть в окне браузера. Для элементов диаграмм и связей можно задать атрибуты в окне спецификации. В спецификации фактора указываются его наименование, тип, описание и способ оценивания, в спецификации индикатора – наименование, источник данных, способ расчета.

С помощью модуля «Методы оценки факторов» можно специфицировать метод оценивания фактора. Для внешних факторов, для которых в качестве способа оценивания задана кластеризация, используется подпрограмма, выполняющая типологическую группировку территориальных образований. Она выбирает индикаторы, связанные с фактором, загружает из базы данных значения индикаторов для различных ТО, осуществляет предварительную обработку данных (определяет корреляцию между индикаторами, пополняет недостающие данные), формирует кластеры.

Для факторов, оценки которых определяются путем фаззификации, запускается программа построения функций принадлежности. При этом если на фактор влияют внешние ограничения, функции принадлежности строятся для каждого кластера, выделенного по соответствующему внешнему фактору. Тип функции принадлежности выбирается пользователем, параметры либо задаются пользователем, либо определяются автоматически на основе распределения значений индикатора.

Для факторов, оцениваемых косвенно на основе состояний других факторов, в зависимости от заданного способа вычисления оценки необходимо либо сформировать совокупность продукционных правил с помощью подпрограммы построения правил, либо выбрать способ преобразования нечеткого значения в числовую оценку и способ обратного преобразования и определить формулу вычисления оценки.

Модуль анализа ситуации позволяет конечному пользователю определить состояния факторов энергоэффективности для выбранного территориального образования в заданный период времени (год). По желанию пользователя оценки могут быть получены только для факторов определенной когнитивной карты. Значения внешних факторов определяются путем выявления кластеров, содержащих заданное территориальное образование. Для получения оценок по функциям принадлежности выполняются процедуры фаззификации значений индикаторов. Оценки, получаемые на основе методов косвенного оценивания, выводятся по правилам или вычисляются по формулам.

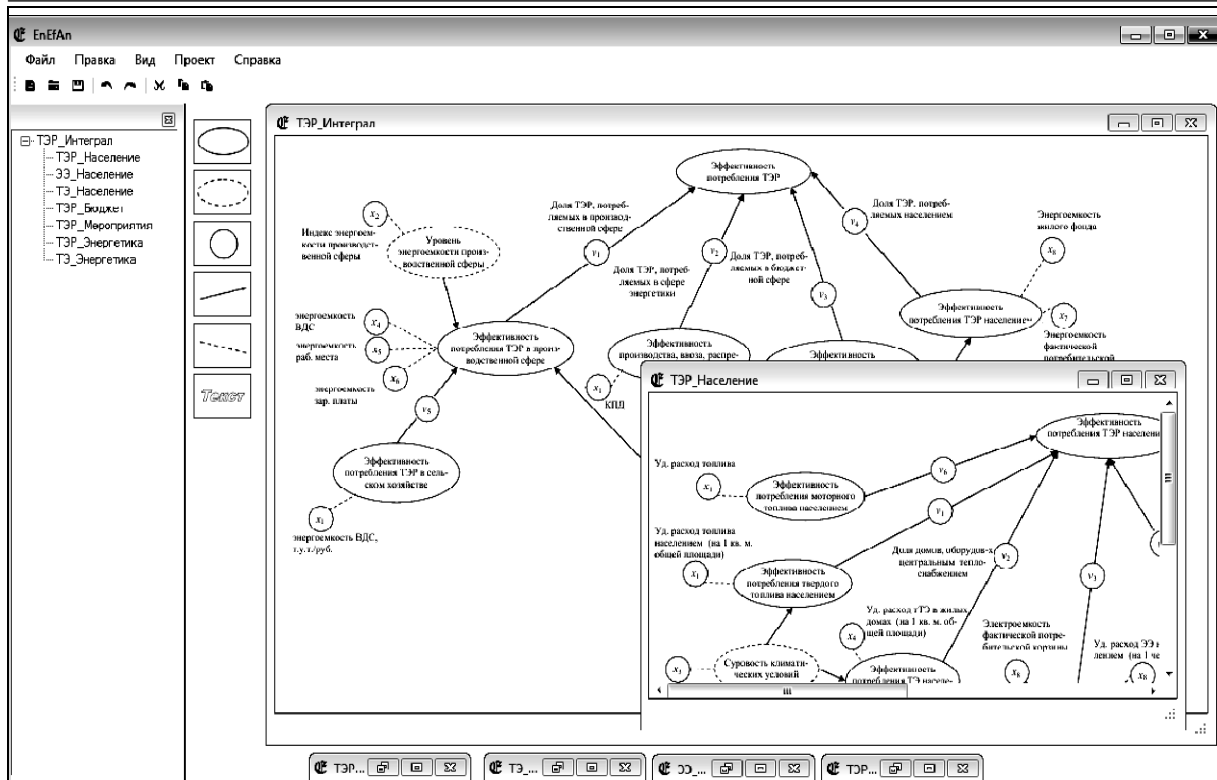


Рис. 2. Интерфейс модуля когнитивного моделирования

Результаты анализа могут быть представлены как в табличном виде, так и в виде дерева причин. Дерево причин формируется на основе соответствующей когнитивной карты путем сопоставления каждому фактору его оценки. Состояния факторов, влияющих на некоторый фактор, являются причинами его состояния, причем силы влияния показывают важность той или иной причины. Пользователь может выделить любое состояние и просмотреть для него объяснение, включающее оценку фактора и степень уверенности, способ получения оценки, значения индикаторов.

Для сравнения различных территориальных образований по тому или иному фактору могут быть использованы таблицы, гистограммы и картограммы. Гистограмма (столбиковая диаграмма) показывает числовые оценки фактора для каждого ТО, вычисленные путем дефаззификации нечетких оценок. Картограмма формируется путем закрашивания соответствующих территорий на географической карте региона определенным цветом, соответствующим полученной оценке. Визуализация результатов анализа позволяет наглядно показать различия в оценках, выделить группы территориальных образований с приблизительно одинаковым уровнем энергоэффективности, выявить образования с резко отличающимися оценками.

В качестве платформы для реализации программного продукта выбрана .NET FrameWork 4.0, использующая компонентный подход к разработке. Ее достоинства – возможность совмещать различные языки программирования, поддержка управления версиями и четкая организация программных компонентов, благодаря чему программу можно разделить на отдельные слабо зависящие друг от друга модули. Средой разработки выбраны Microsoft Visual Studio 2012 и объектно-ориентированный высокоуровневый язык программирования C#. Для файловой базы данных используется Microsoft Access, в качестве сервера базы данных для групповой разработки выбран Microsoft SQL Server 2008.

Заключение. Разрабатываемая информационная система может быть полезна органам власти субъектов РФ и муниципальных районов для обоснования принятия решений по повышению энергетической эффективности при формировании программ энергосбережения. Повышение качества решений происходит за счет интеллектуального анализа, позволяющего учитывать: климатические и социально-экономические особенности различных территориальных образований; причинно-следственные связи между факторами, влияющими на положение дел; неопределенности, возникающие при интерпретации данных, характеризующих состояние энергоэффективности.

Работа выполнена при финансовой поддержке гранта РФФИ № 13-07-00397а.

Литература

1. Энергосбережение и энергоэффективность // Челябинский ЦНТИ. Филиал ФГБУ «Российское энергетическое агентство» Минэнерго РФ [Электронный ресурс]. – Режим доступа: http://www.csti.ru/our_activities/energy_saving/, свободный (дата обращения: 11.02.2013).
2. Единая информационная система для мониторинга и управления эффективностью энергосбережения // Энергосбережение. – 2011. – № 1. – С. 4–6.
3. Сляров Е.В. О создании информационной системы мониторинга и управления эффективностью энергосбережения на объектах города Москвы // Промышленная энергетика. – 2011. – № 6. – С. 2–6.
4. Оценка ситуации с энергетической эффективностью в муниципальных образованиях на основе диаграммы влияния факторов / М.П. Силич, В.А. Силич, С.В. Аксенов, В.С. Ахмедов // Энергобезопасность и энергосбережение. – 2012. – № 3. – С. 9–14.
5. Оценка энергетической эффективности муниципальных образований на основе нечеткой интерпретации данных / М.П. Силич, С.В. Аксенов, В.С. Ахмедов, В.И. Скрыбин // Информационные и математические технологии в науке и управлении: труды XVII Байкальской Всерос. конф. – Иркутск: ИСЭМ СО РАН, 2012 – Ч. 3. – С. 129–134.
6. Municipalities' energy efficiency estimation approach based on fuzzy models / S.V. Axyonov, V.S. Ahmedov, M.P. Silich, V.I. Skryabin // Proceedings of 7th International Forum on strategic technology 2012 (IFOST 2012), IEEE. – 2012. – Tomsk, 2012. – P. 333–336.
7. Fuzzy Cluster analysis: methods for classification, data analysis, and image recognition / F. Hoepfner, F. Klawonn, R. Kruse, and T. Runkler. – New York: John Wiley & Sons, Inc., 1999. – 300 p.

Силич Мария Петровна

Д-р техн. наук, профессор каф. автоматизации обработки информации ТУСУРа
Тел.: 8 (382-2) 70-15-91
Эл. почта: smp@muma.tusur.ru

Вахитов Александр Робертович

Канд. техн. наук, нач. отд. информационно-технического обеспечения Института кибернетики ТПУ
Тел.: 8-909-538-17-09
Эл. почта: argv@tpu.ru

Прудников Илья Иванович

Студент гр. 8591 Института кибернетики ТПУ
Тел.: 8-913-822-12-69
Эл. почта: iprudnikov@yahoo.com

Silich M.P., Vakhitov A.R., Prudnikov I.I.

Information system of territories' energy efficiency analysis

The structure of information system of territories' energy efficiency analysis is considered. System allows building cognitive maps hierarchy that represents the connections between factors that have an influence on the energy efficiency state, defining the membership functions for fuzzy estimation of factors through indicators' values and generation rules and formulas for indirect estimation of the factors. The system performs the territories' fuzzy clustering to take into account the influence of external factors. Tree causes can be created as the result of analysis.

Keywords: energy efficiency, territorial formations, information system, cognitive maps.

УДК 004.827

Л.В. Галицкая, А.А. Голубева, Ю.Б. Гриценко, О.И. Жуковский, И.Ю. Пудуль

Идентификация параметров объектов электронного генерального плана металлургического предприятия

Приведено описание метода решения задачи идентификации параметров объектов электронного генерального плана металлургического предприятия на основе теории нечеткого моделирования с использованием методов наименьших квадратов, роящихся частиц и гибридного метода.

Ключевые слова: электронный генеральный план, геоинформационная система, нечеткое моделирование, идентификация.

Постановка задачи. Современные промышленные металлургические предприятия имеют в своем составе большое количество разнообразных производств: доменное, мартеновское, коксохимическое, литейное и др., которые в свою очередь занимают большое территориальное пространство, плотно насыщенное разнообразными производственными объектами. В настоящий момент времени на предприятиях используют такой подход к автоматизации, при котором основными задачами, требующими внедрения компьютерных технологий, являются те, эффект от успешного решения которых оценивается наиболее явным образом. При использовании данного подхода управление производственной инфраструктурой предприятия является одним из основных автоматизируемых процессов.

Решение задачи управления производственной инфраструктурой металлургического предприятия во многом зависит от эффективности разработки и использования генерального плана. В этом случае применение технологии ГИС – закономерный этап на пути перехода к безбумажной технологии обработки информации. Географическая информационная система (ГИС) – это современная компьютерная технология для картирования и анализа объектов реального мира и событий, происходящих с этими объектами [1].

В состав этой технологии входят традиционные операции работы с базами данных, такие как запрос и статистический анализ, с преимуществами полноценной визуализации и пространственного анализа, которые предоставляет карта [2].

На сегодняшний день применение ГИС-технологий, лежащих в основе комплексных информационных систем, все чаще ограничивается решением задач информационного сопровождения, а их роль в решении задач анализа недооценивается. Причиной подобного обстоятельства является либо наличие в современных ГИС слаборазвитого аппарата анализа данных, либо полное его отсутствие. Такое положение дел требует совершенствования подходов к созданию информационных систем на базе ГИС с учетом реализации функций анализа данных. Одной из ключевых задач анализа данных на металлургическом предприятии является задача идентификации (восстановления) параметров объектов электронного генерального плана (ЭГП). Решение задачи идентификации позволит узнать причину, свойства взаимосвязей объектов, а также место нахождения связанных объектов. Получение этой информации может помочь более глубоко понять ситуацию в данном месте, сделать правильный выбор или лучше подготовиться к предстоящим событиям и условиям.

Описание метода решения. Для решения задачи идентификации параметров объектов ЭГП авторы статьи предлагают использовать теорию нечеткого моделирования. Использование данной теории обусловлено тенденцией увеличения сложности математических и формальных моделей реальных систем и процессов управления, связанной с желанием повысить их адекватность и учесть все большее число различных факторов, оказывающих влияние на процессы принятия решений [3]. Авторами статьи разработан обобщенный алгоритм решения задачи идентификации (восстановления) пространственно-временных данных ЭГП металлургического предприятия, представленный на рис. 1. Данный алгоритм был реализован в программном обеспечении интегрированным в веб-ориентированную геоинформационную технологию формирования и мониторинга электронного генерального плана инженерной инфраструктуры [4] в качестве модуля анализа данных.

Важным этапом диагностики параметров объектов ЭГП является сбор априорной информации об их характеристиках. Важно помнить, что объекты ЭГП весьма разнообразны. Например, в определенном производстве металлургического предприятия могут пролегать самые разнообразные типы инженерных коммуникаций (водопроводные, канализационные сети, паропроводы, продуктопроводы и др.), и их параметры (толщина труб, материал, глубина залегания и др.) могут существенно отличаться. Проводя обучение системы (реализованной в модуле анализа данных) на существующих данных, нужно помнить о том, что применять задачу идентификации можно только к такому же типу данных. Другими словами, информация должна принадлежать одной генеральной совокупности.

После того как данные были собраны, можно приступить к процессу обучения системы. На первом этапе необходимо определить входные и выходные параметры системы, которыми являются характеристики объектов, и проинициализировать их, задав соответствующее количество термов и выбрав подходящую функцию принадлежности.

На данном этапе развития исследований в области нечеткого моделирования практически невозможно заранее точно сказать, как повлияют на результат вычислений выбор функций принадлежности и количество термов. В силу этого необходимо проводить эксперименты по инициализации переменных и наблюдать за результатами вычислений. Выбрав удовлетворяющую точность вычислений, определить, какая функция принадлежности наиболее пригодна и показывает адекватные результаты.

Функция принадлежности может быть представлена одним из следующих типов:

- треугольный – рационально использовать при необходимости быстрой классификации и отсутствии требований высокого качества (необходим экспресс-анализ системы), т.е. пользователю важна в первую очередь скорость обработки данных;

- гауссовское распределение (функция имеет гладкий вид) – применение обусловлено в случае, если важна устойчивость распределения и есть необходимость уточнения данных экспресс-анализа. Функция Гаусса является оптимальной и дает практически приемлемые результаты по скорости обучения, по скорости классификации и по эффективности классификации;

- трапеция (функция имеет линейный и местами ступенчатый вид) – использование трапециевидной функции обусловлено наличием специфической выборки (специфичность выборки определяется тем, что анализируемые значения выходят за рамки обучающей выборки) и отсутствием жестких ограничений по времени;

- парабола – параболическая функция принадлежности является наилучшей по скорости обучения. Эффективность классификации также находится на высоком уровне [5].

После инициализации переменных следует перейти ко второму этапу обучения системы, настройке антецедентов и консеквентов правил. Для настройки базы правил применяется ряд методов: метод наименьших квадратов (МНК), метод роящихся частиц (РЧ), гибридный алгоритм (настройка методом роящихся частиц, последующая подстройка методом наименьших квадратов).

При выборе метода настройки, также как и при выборе функции принадлежности, на данном этапе развития исследований невозможно с уверенностью сказать, какой метод будет наиболее адекватен. Но следует учесть тот факт, что метод РЧ – наиболее часто успешно применяемая эвристика для обучения нечетких систем, так как данный метод защищен от «застывания» в локальных решениях.



Рис. 1. Обобщенный алгоритм решения задачи идентификации пространственно-временных данных ЭГП металлургического предприятия

МНК является классическим алгоритмом оптимизации и находит оптимальное решение за одну итерацию, поэтому в совокупности он менее затратный по ресурсам, чем аналоги при достижении той же точности. Рекомендацией применения МНК является необходимость получить хорошее и быстрое решение.

Рекомендацией применения гибридного алгоритма является желание улучшить полученное решение, но данный метод требует больших временных затрат.

Одним из заключительных этапов процесса настройки обучения нечеткой системы является выбор типа нечеткой системы, так называемого алгоритма вывода. Для выбора типа нечеткой системы следует произвести предварительный статистический анализ (рассчитать коэффициент регрессии) [6]:

1) если коэффициент регрессии более 0,7, необходимо использовать систему нечеткого вывода типа Такаги–Сугено;

2) если коэффициент регрессии не превышает 0,2, то используем систему нечеткого вывода типа Мамдани;

3) если коэффициент от 0,2 до 0,7, то используем систему нечеткого вывода типа Сингтон.

После того как подобраны все необходимые параметры и методы, произведена настройка системы, можно запустить процесс восстановления данных.

Проведение эксперимента. Для выполнения эксперимента по аппроксимации данных смоделируем ситуацию отсутствия данных по объектам слоя «продуктопровод», для чего из таблицы наблюдений, состоящей из 104 записей и 4 полей, удалим девять значений атрибута «толщина стенки» в записях, представленных в табл. 1.

Таблица 1
Перечень записей в экспериментальной таблице, для которых необходимо удалить значение атрибута «Толщина стенки»

Object Id	Диаметр, мм	Толщина стенки, мм	Рабочая температура, °С
119465	127,5	3,5	14
119980	110	3,2	13
119495	125	3,5	13
118025	150	2,9	45
79113	192,5	2,8	67
119267	125	2,3	33
121518	162,5	3	52
121717	127,5	3,5	14
119045	155	2,9	43

(диаметр и рабочая температура). Для настройки параметров нечеткой модели (антецедентов правил) будем использовать метод роящихся частиц. Для осуществления настройки параметров нечеткой модели (консеквентов правил) применим метод наименьших квадратов (МНК).

Далее непосредственно запустим процедуру восстановления экспериментальных данных. Результат работы процедуры восстановления данных отображен в табл. 2. Восстановлены отсутствующие значения у объектов «продуктопроводы» в свойстве «толщина стенки». Среднеквадратичное отклонение составило 0,016.

Таблица 2
Перечень записей в экспериментальной таблице, для которых восстановили значение атрибута «Толщина стенки»

Object Id	Диаметр, мм	Толщина стенки, мм	Рабочая температура, °С
119465	127,5	3,500	14
119980	110	3,229	13
119495	125	3,509	13
118025	150	2,910	45
79113	192,5	2,791	67
119267	125	2,312	33
121518	162,5	2,990	52
121717	127,5	3,500	14
119045	155	2,893	43

Далее в модуле анализа данных осуществим задание входных и выходных параметров. Входными параметрами будут являться «диаметр» и «рабочая температура», а параметр «толщина стенки» обозначим как выходной. Выполним инициализацию и обучение на основе рассчитанного коэффициента корреляции, превышающего 0,7. В этом случае система рекомендует выбрать тип нечеткой системы Такаги–Сугено из реализованных в системе. Для осуществления инициализации переменных выберем функцию принадлежности «треугольная» и зададим по 5 термов для каждого входного параметра

(диаметр и рабочая температура). Для настройки параметров нечеткой модели (антецедентов правил) будем использовать метод роящихся частиц. Для осуществления настройки параметров нечеткой модели (консеквентов правил) применим метод наименьших квадратов (МНК).

Далее непосредственно запустим процедуру восстановления экспериментальных данных. Результат работы процедуры восстановления данных отображен в табл. 2. Восстановлены отсутствующие значения у объектов «продуктопроводы» в свойстве «толщина стенки». Среднеквадратичное отклонение составило 0,016.

Заключение. Разработанный авторами метод решения задачи идентификации параметров объектов ЭПП металлургического предприятия и реализованный в модуле анализа данных интегрированного в веб-ориентированную геоинформационную технологию формирования и мониторинга электронной генеральной плана инженерной инфраструктуры показал адекватный результат по восстановлению данных и может быть использован для решения задач управления производственной инфраструктурой предприятия.

Выполнение данной работы проводилось при финансовой поддержке Министерства образования и науки Российской Федерации в рамках мероприя-

тия 2.4 федеральной целевой программы «Исследования и разработки по приоритетным направлениям развития научно-технологического комплекса России на 2007–2013 годы», проект «Разработка Web-ориентированных геоинформационных технологий формирования и мониторинга электронного генерального плана инженерной инфраструктуры», государственный контракт № 07.524.11.4013 от 03 ноября 2011 г.

Литература

1. Ехлаков Ю.П. Архитектура веб-ГИС-сервера для ведения электронного генерального плана инженерной инфраструктуры / Ю.П. Ехлаков, О.И. Жуковский, Н.Б. Рыбалов // Известия Томского политехнического университета. – 2006. – Т. 309, № 7. – С. 146–152.
2. ДеМерс Майкл Н. Географические информационные системы. Основы. – М.: Дата+, 1999. – 490 с.
3. Леоненков А.В. Нечеткое моделирование в среде MATLAB и fuzzyTECH. – СПб.: БХВ Петербург, 2005. – 736 с.
4. Гриценко Ю.Б. Архитектура веб-ГИС-сервера для ведения электронного генерального плана инженерной инфраструктуры / Ю.Б. Гриценко, О.И. Жуковский, М.М. Милихин // Качество. Инновации. Образование. – 2012. – № 12. – С. 97–101.
5. Yuen Fuzzy linguistic variable matrix and parabola-based fuzzy normal distribution / K.K.F. Yuen, H.C.W. Lau // Intelligent Information Processing III. IFIP TC12 International Conference on Intelligent Information Processing (IIP 2006), September 20–23, Adelaide, Australia. – P. 205–215.
6. Голубева А.А. Интеллектуальный анализ пространственно-временных данных объектов инженерной инфраструктуры // Доклады ТУСУРа. – 2012. – № 2 (26), ч. 2. – С. 216–219.

Галицкая Любовь Владимировна

Мл. науч. сотрудник каф. АОИ ТУСУРа
Тел.: (382-2) 90-01-80
Эл. почта: lyb_vlad@ngs.ru

Голубева Александра Александровна

Аспирант каф. АОИ ТУСУРа
Тел.: 8-913-825-84-26
Эл. почта: Sasha_karateka@mail.ru

Гриценко Юрий Борисович

Канд. техн. наук, доцент каф. АОИ ТУСУРа
Тел.: (382-2) 90-01-80
Эл. почта: ubg@muma.tusur.ru

Жуковский Олег Игоревич

Канд. техн. наук, доцент каф. АОИ ТУСУРа
Тел.: (382-2) 90-01-80
Эл. почта: ol@muma.tusur.ru

Пудуль Игорь Юрьевич

Студент каф. АОИ ТУСУРа
Тел.: (382-2) 90-01-80
Эл. почта: svigani4ok@gmail.com

Galitskaya L.V., Golubeva A.A., Gritsehko Yu.B., Zhukovskiy O.I., Pudul I.Yu.

Identification of the parameters of the objects in electronic master plan for metallurgical enterprise

The paper describes the method for solving the parameter identification of objects in an electronic master plan for the metallurgical enterprise based on the theory of fuzzy modeling using the method of Ordinary Least Squares, swarming particles and hybrid method.

Keywords: electronic master plan, geographical information system, fuzzy modeling, identification.

УДК 004.41

М.В. Зубов, А.Н. Пустыгин, Е.В. Старцев

Применение универсальных промежуточных представлений для статического анализа исходного программного кода

Описывается использование промежуточных представлений исходного кода для выполнения статического анализа. Предлагается расширить использование универсальных и многоуровневых представлений в одном общем. Представлен прототип программного инструмента, генерирующий и обрабатывающий такое представление.

Ключевые слова: статический анализ, промежуточное представление, машинные комментарии, исходный код, диаграмма классов, Java, Python.

Статический анализ облегчает выполнение типовых задач разработки. Эти задачи могут быть связаны с поиском ошибок, улучшением качества кода, внутренним тестированием [1]. Перспективным направлением является извлечение информации по исходному коду (построение машинных комментариев [2]).

Промежуточные представления исходных текстов программ. Промежуточное представление (далее – представление) программного текста – это синтаксически и семантически эквивалентный исходному коду набор данных, над которым выполняется анализ [3]. Представление может быть выполнено в виде различных наборов данных. Самый простой и наиболее изученный в области компиляции вариант – абстрактное синтаксическое дерево AST, фактически это результат синтаксического разбора. Более сложной, обеспечивающей увеличение скорости анализа является реляционная база данных. Однако затраты на получение такого представления растут. Один из наиболее эффективных видов представления – конечный автомат. Помимо AST, в компиляции широко применение нашло представление SSA (Single-State Assignment) и его развитие GSA [4].

Представления, описывающие только один входной язык, будем называть частными. Те, что пригодны для исходного кода на нескольких входных языках, – универсальными. К частным представлениям можно отнести уже упомянутое абстрактное синтаксическое дерево, описывающее исходный текст программы во всех деталях, свойственных конкретному языку программирования. К универсальным представлениям можно отнести SSA и GSA с той оговоркой, что версии переменных будут описываться в терминах конкретного языка.

Использование универсальных представлений сокращает общие затраты при анализе, так как общее количество требуемых функциональных модулей представления и реализации цели анализа для представления снижается. При использовании частных представлений потребуется общее количество $(N + N \times M)$ операций, а при использовании универсального представления – $N + M$. При положительных значениях N и M , что имеет место для количества целей и входных языков, первая величина всегда больше второй на $M \times (N - 1)$ операций, что определяет выигрыш в сокращении общего количества операций для покрытия всех входных языков и целей анализа.

Безусловно, трудозатраты зависят не только от общего количества операций но и от их сложности, которая при использовании универсального представления может быть выше (особенно в области генерации универсального представления), но при увеличении количества языков и целей анализа преимущество в $M \times (N - 1)$ операций играет все большую роль. Уже для 3 языков и 3 целей анализа, когда выигрыш достигает уменьшения операций в 2 раза, имеется серьезная причина для выбора в пользу универсального представления. Из основного достоинства универсального представления следует и недостаток: универсальные представления не позволяют отразить все детали синтаксиса каждого из языков исходного текста программ.

Применение универсальных и многоуровневых представлений. Для выполнения анализа нам вовсе не обязательно иметь полную информацию обо всех деталях исходного текста. Самым простым уровнем является наиболее детализированный. Он полностью соответствует исходному коду. В различных проектах выделяют разное количество уровней. Так, например, в анализаторе PQL предлагается применять 3 уровня: модель глобального программного дизайна, модель структуры программы, модель детального программного дизайна [5]. А в Vauhaus Project имеется 2 уровня:

низкоуровневый Inter Mediate Language, детализированный до операторов, и Resource Flow Graphs, описывающий архитектуру [6]. Нечто подобное наблюдается и в организации компилятора GCC, в котором используются представления GIMPLE, CFG (Control Flow Graph), RTL, помимо классического AST [7]. Как видно, различные подходы сходятся в том, что есть полностью соответствующее коду представление и описывающее глобальную структуру программы, архитектуру. Между этими двумя крайними уровнями можно выделить дополнительные уровни представлений. С точки зрения анализа основной интерес представляют структурные единицы используемой парадигмы программирования и языка. Например, для объектно-ориентированных языков это может быть представление уровня классов, их иерархии и взаимосвязей (похоже на диаграмму классов UML [8]).

XML-форма классового представления. Для прототипной реализации были выбраны весьма разные языки: Java и Python. Самое главное их отличие в том, что Java – это язык со строгой типизацией данных, а Python – с динамической. Это накладывает некоторые ограничения на получение связей агрегирования и композиции, потому что типы полей в Python определяются на этапе исполнения, а не на этапе компиляции, т.е., они не могут быть получены напрямую из исходного кода. Для упрощения получения прототипа откажемся от получения типов и построения связей агрегирования.

В Python присутствует множественное наследование, которого нет в Java. В Java есть интерфейсы, которых нет в Python, но в Java есть множественное наследование у интерфейсов, а также класс может реализовывать несколько интерфейсов. В данном случае самым простым решением будет приравнять интерфейсы к классам. Фактически, так и есть, интерфейс – это абстрактный класс, который не содержит никаких реализаций. Таким образом, представление будет допускать множественное наследование для всех входных языков, даже для тех, где его нет. Отказываясь от интерфейсов, также придется отказываться от отношений реализации. В Python'e такое отношение не может быть реализовано средствами языка так, чтобы оно отличалось от отношения обобщения. Однако оно может присутствовать на уровне логики. Для предлагаемого представления удобно использовать древовидные структуры. В качестве формата данных хорошо подойдет XML. В качестве основного тега выступает тег Class, описывающий отдельный класс, в качестве дочерних тегов используются теги: Attr – описывает поля класса; Method – описывает методы класса; Parent – описывает родителей класса.

Графическое представление XML-схемы [9] изображено на рис. 1. Полный текст схемы можно получить из git-репозитория [10].

Прототип системы получения универсального классового представления. Реализованный прототип поддерживает 2 входных языка программирования – Java и Python. Для Python использовался генератор на основе библиотеки logilab-astng, что позволило сразу получать информацию о классах проекта. Для Java использовался генератор, использующий AST. Дерево разбора получалось из XML-документа, генерируемого доработанным компилятором Open JDK. Реализация прототипа выложена в git-репозитории на открытый ресурс GitHub под свободной лицензией.

Результаты получения представления и его визуальной формы. Для тестирования предложенного промежуточного представления была получена визуализированная иерархия классов в dot-формате описания графов из пакета Graphviz, так как иерархия наследования может быть представлена в виде направленного графа.

Для демонстрации универсальности представления и его полноты, независимо от входного языка, возьмем типовую реализацию шаблона проектирования «Посетитель» [11]. Были написаны 2 одинаковые по смыслу реализации на Java и Python. Полные тексты примеров доступны в git-репозитории [12].

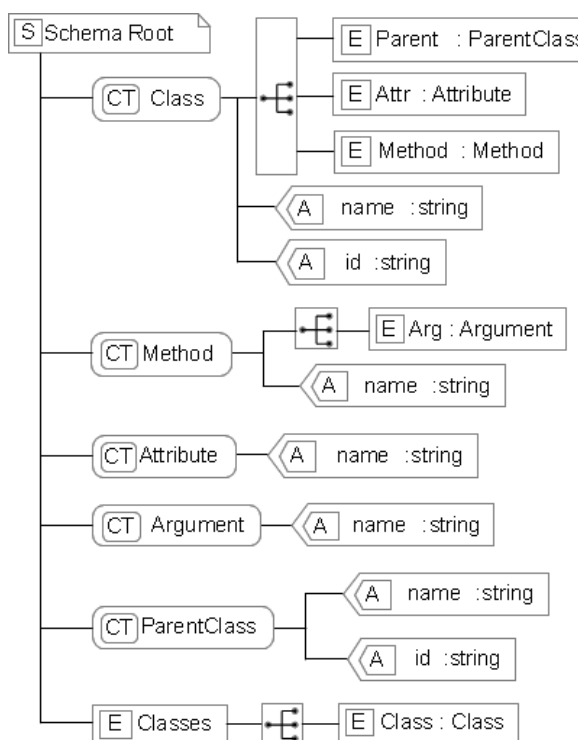


Рис. 1. Графический вид XML-схемы универсального классового представления

Листинг 1. Фрагмент промежуточного представления примера на Python.

```
<?xml version='1.0' encoding='utf-8'?>
<Classes>
  <Class id="4" name="Node">
    <Attr name="strPos"/>
    <Attr name="lineno"/>
    <Method name="accept">
      <Arg name="visitor"/>
    </Method>
  </Class>
  <Class id="5" name="VariableNode">
    <Attr name="name"/>
    <Method name="accept">
      <Arg name="visitor"/>
    </Method>
    <Parent id="4" name="Node"/>
  </Class>
  <Class id="11" name="NodeVisitor">
    <Method name="visit_variable">
      <Arg name="node"/>
    </Method>
    <Method name="visit_assignment">
      <Arg name="node"/>
    </Method>
    <Method name="visit_class">
      <Arg name="node"/>
    </Method>
  </Class>
  <Class id="12" name="Verificator">
    <Attr name="specification"/>
    <Method name="visit_assignment">
      <Arg name="node"/>
    </Method>
    <Method name="visit_variable">
      <Arg name="node"/>
    </Method>
    <Method name="visit_class">
      <Arg name="node"/>
    </Method>
    <Parent id="11" name="NodeVisitor"/>
  </Class>
</Classes>
```

Промежуточное представление иерархии классов для Python в виде XML представлено в листинге 1. Для Java визуализация такого представления в UML-подобной форме представлена на рис. 2. Анализируя данные, видно, что диаграмма и XML-представление идентичны по смыслу.

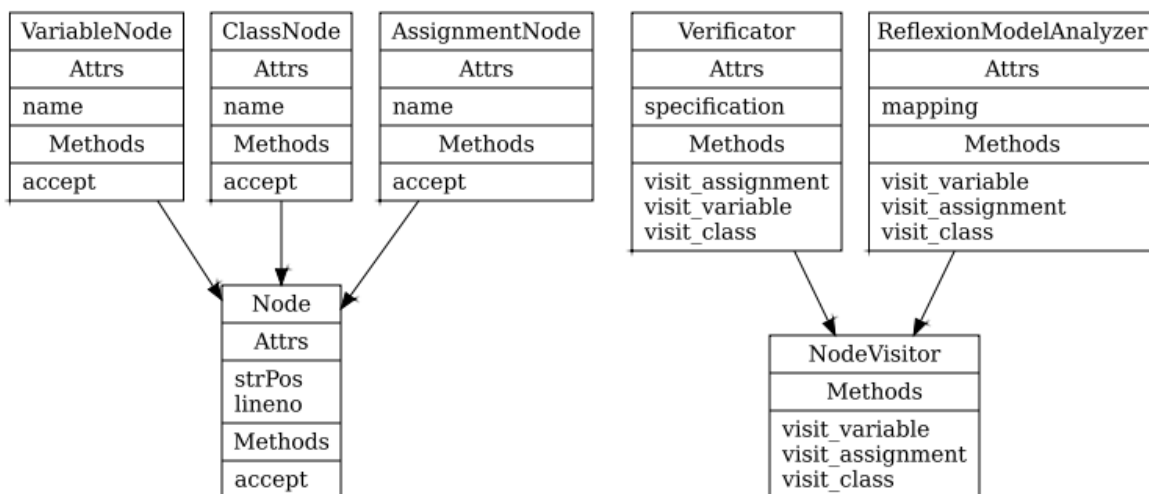


Рис. 2. UML-подобная диаграмма классов примера Java

Анализ иерархии классов с помощью построенного прототипа. Объектная структура программного продукта в соответствии с современными представлениями должна удовлетворять определенным критериям, в том числе правильно построенным линиям наследования от базовых классов.

С целью тестирования возможностей построенного прототипа было проверено выделение методов для групп объектов, которые могут быть вынесены в общий суперкласс.

Для анализа были выбраны несколько Python- и Java-проектов. На Python – библиотека для веб-разработки Django версии 1.4.1, сетевая библиотека Twisted версии 12.0.0, Python-пакет logilab, который использовался для обработки AST-деревьев (компоненты logilab-astng версии 0.23.1 и logilab-common версии 0.58.0) (табл. 1). Проблемными методами названы те, чьи имена совпадают в разных классах, но не унаследованы ими от общего суперкласса.

На Java анализировались компилятор javac, из состава OpenJDK v6u20, статический анализатор FindBugs v2.0.1, SQL-клиент SquirrelL (табл. 2).

Кроме того, был проанализирован собственный код. Для анализа был выбран построитель AST для Java. В результате анализа было найдено 20 «похожих» методов без родительской реализации из 102. Как оказалось, класс, выполняющий обход узлов, получаемых от парсера, делегировал вызовы зарегистрированным в нем классам, выполняющим вывод в XML или другую обработку. Причем делегировал он их, повторяя интерфейс этих обработчиков в 15 из 17 методов. При такой реализации стоило в самом классе, выполняющем обход, реализовать интерфейс обработчика, так как его поведение совпадало с поведением обработчика.

Таким образом, был проведен рефакторинг собственного кода, сокративший число проблемных методов с 20 до 3, причем, с точки зрения чистоты кода, этот рефакторинг требовался. Из оставшихся трех два также подлежат будущему рассмотрению. С рефакторингом можно ознакомиться по ссылке <https://github.com/exbluesbreaker/csu-code-analysis/commit/64c6974517fdc700c2dc9f0c0173208ab3352af5> (информация о коммите в git).

Полученные результаты дают хорошую почву для возможного рефакторинга исходного кода проанализированных программных продуктов, это может быть как выделение новых суперклассов или интерфейсов, так и добавление новых методов к уже существующим. Доля имен методов с одинаковыми именами, реализованными в разных классах, более 10% для всех исследованных продуктов, а для некоторых она достигает почти 20%. Разработанный прототип не предназначен для автоматизации такого рефакторинга, он лишь извлекает из исходного кода полезную для разработчика информацию.

Выводы. В результате проделанного исследования была разработана и описана модель промежуточного представления, позволяющая универсальным инструментом выполнять статический анализ исходного кода Java и Python. Представлен прототип такого инструмента. Для другого языка следует только получить универсальное промежуточное представление по предлагаемому шаблону.

Литература

1. Зубов М.В. Статический анализ ПО с помощью его промежуточных представлений и технологий с открытым исходным кодом / М.В. Зубов, А.Н. Пустыгин, Е.В. Старцев // Матер. 2-й Междунар. конф. «FOSS. Lviv-2012», Львов. – Львів: Сорока, 2012. – С. 165–168.
2. Автоматический синтез комментариев к программным кодам: перспективы развития и применения / А.Н. Пустыгин, А.И. Иванов, Ю.К. Язов, С.В. Соловьев // Программная инженерия (Москва). – 2012. – № 3. – С. 30–34.
3. Зубов М.В. Подходы к статическому анализу открытого исходного кода / М.В. Зубов, А.Н. Пустыгин, Е.В. Старцев // Матер. 8-й Междунар. конф. «Linux Vacation» Eastern Europe. – Брест: Альтернатива, 2012. – С. 36–40.
4. The Formation and Simulation of a «Whole Program» Gated Singular Assignment Program Dependence Graph / J. Cook, D. Gottlieb, B. Greskamp, R. Kujoth. – Unconventional Computer Architecture

Group, Final Report. – 2008 [Электронный ресурс] – Режим доступа: <http://iacoma.cs.uiuc.edu/~greskamp/pdfs/497mf.pdf>, свободный (дата обращения: 15.01.2013).

5. Jarzabek S. Design of Flexible Static Program Analyzer with PQL // IEEE Transactions on software engineering. – 1998. – Vol. 24, № 3. – С. 197–215.

6. Bauhaus project [Электронный ресурс]. – Режим доступа: mtc.epfl.ch/software-tools/blast/index-epfl.php, свободный (дата обращения: 02.12.2012).

7. Basic gcc Intermediate Representation Dumps [Электронный ресурс]. – Режим доступа: www.cse.iitb.ac.in/~uday/courses/cs324-05/gccProjects/node4.html, свободный (дата обращения: 02.12.2012).

8. Буч Г. Язык UML. Руководство пользователя / Г. Буч, Дж. Рамбо, А. Якобсон. – СПб.: Питер, 2004. – 432 с.

9. W3C XML Schema Definition Language (XSD) 1.1 [Электронный ресурс]. – Режим доступа: www.w3.org/TR/xmlschema11-1, свободный (дата обращения: 02.12.2012).

10. XML-схема классового представления [Электронный ресурс]. – Режим доступа: <https://github.com/exbluesbreaker/csu-code-analysis/blob/master/data/classes.xsd>, свободный (дата обращения: 02.12.2012).

11. Приемы объектно-ориентированного проектирования. Паттерны проектирования / Э. Гамма, Р. Хелм, Р. Джонсон, Дж. Влиссидес. – СПб.: Питер, 2007. – 366 с.

12. GitHub. Репозиторий CSU-code-analysis [Электронный ресурс]. – Режим доступа: <https://github.com/exbluesbreaker/csu-code-analysis>, свободный (дата обращения: 02.12.2012).

Зубов Максим Валерьевич

Аспирант каф. компьютерной безопасности и прикладной алгебры (КБиПА) ЧелГУ, г. Челябинск

Тел.: +7-961-784-45-31

Эл. почта: zubovmv@gmail.com

Пустыгин Алексей Николаевич

Канд. техн. наук, д. доцент каф. КБиПА

Тел.: +7-905-835-98-68

Эл. почта: p2008an@rambler.ru

Старцев Евгений Владимирович

Аспирант каф. КБиПА

Тел.: +7-961-789-69-23

Эл. почта: slayer-gurgen@yandex.ru

Zubov M.V., Pustygin A.N., Startsev E.V.

Use of the intermediate software representations for static analysis of source code

This article describes using of intermediate representations of source code for static analysis. It is proposed to extend using of generic and multi-level representations in one common. Software tool prototype that can make and process this kind of representation is introduced.

Keywords: static analysis, intermediate representation, machine comments, source code, class diagram, Java, Python.

УДК 004.056

Б.И. Ефимов, Р.Т. Файзуллин

Вероятность принятия ложного решения под воздействием угроз информационной безопасности в системах принятия решений с привлечением экспертов

Предложено решение задачи вычисления вероятности принятия ложного решения в системах принятия решения с привлечением экспертов под воздействием угроз информационной безопасности, направленных на изменение ответов экспертов. Получены оценки, показывающие целесообразность или нецелесообразность увеличения количества экспертов для уменьшения влияния действий злоумышленника на принимаемое решение.

Ключевые слова: информационная безопасность, системы принятия решений, эксперты, угрозы, теория вероятностей.

В любой сфере деятельности человека постоянно возникает огромное количество проблем, требующих принятия решений. Во многих случаях, особенно когда по принимаемому вопросу не разработана необходимая база первичных данных, оптимальным вариантом может являться использование экспертов-аналитиков. Для принятия сложных управленческих решений разработаны и математически описаны различные модели и методы [1, 2].

При разработке систем принятия решений с привлечением экспертов обеспечение информационной безопасности является одной из важных задач. Вопросы информационной безопасности указанных систем подробно изложены в [3, 4].

В соответствии с [4] угрозы безопасности систем принятия решений с привлечением экспертов могут быть направлены на следующие объекты: экспертов, принимающих участие в опросе, узлы коммутации, линии связи, лицо, принимающее решение.

Виды воздействий на объекты сети могут быть как преднамеренными, т.е. осуществляемыми злоумышленником, так и случайными, обусловленными отказами оборудования, программ и каналов связи. Преднамеренные угрозы направлены, в конечном итоге, на принятие «нужного» для злоумышленника решения. Отказы оборудования носят случайный, непредсказуемый характер и могут повлиять на принятие решения в пользу любой из существующей альтернатив. В данной работе рассматриваются вопросы устойчивости объективного решения экспертов при некоторой вероятности реализации угрозы, насколько оправдано увеличение числа экспертов.

Исходные положения. Для вычисления вероятности принятия ЛПР решения, отличного от решения, которое было бы принято по результатам опроса экспертов при отсутствии угроз информационной безопасности, определим следующие моменты:

- опрос экспертов проводится по выбору одной из двух альтернатив: «0» и «1»;
- действия злоумышленника направлены на изменение ответов экспертов в пользу альтернативы «0»;
- угрозы по блокированию ответа эксперта в процессе передачи до ЛПР являются случайными угрозами и могут быть реализованы в отношении любого эксперта.

Введем следующие обозначения:

m – общее количество экспертов, принимающих участие в голосовании;

n – количество экспертов, проголосовавших за альтернативу «1»;

$(m - n)$ – количество экспертов, проголосовавших за альтернативу «0»;

$P_{change\ i}$ – вероятность изменения злоумышленником ответа i -го эксперта;

$P_{bloc\ i}$ – вероятность блокирования ответа i -го эксперта в процессе передачи до ЛПР.

В данной статье рассматривается частный случай, когда угрозы по блокированию ответов экспертов отсутствуют ($P_{bloc\ i}=0$) и вероятности изменения ответов эксперта под действием угроз ИБ $P_{change\ i}$ одинаковы для всех экспертов.

Варианты голосования экспертов. Возможны следующие варианты голосования экспертов:

- 1) $n > m/2$.

Количество экспертов, проголосовавших за альтернативу «1», превышает количество экспертов, проголосовавших за альтернативу «0». В этих случаях из-за реализации угроз безопасности решение, принимаемое ЛПР по выбору одной из альтернатив, может быть изменено.

2) $n = m/2$.

Количество экспертов, проголосовавших за альтернативы «0» и «1», равны. В этом случае из-за реализации угроз безопасности ЛПР может быть выбрана альтернатива «0», а не назначено повторное голосование.

Указанный вариант возможен при четном количестве экспертов, принимающих участие в голосовании.

3) $n < m/2$.

Количество экспертов, проголосовавших за альтернативу «0», превышает количество экспертов, проголосовавших за альтернативу «1». ЛПР будет выбрана альтернатива «0», вне зависимости от того, существуют угрозы ИБ по изменению ответов экспертов в пользу альтернативы «0» либо подобных угроз не существует, т.е. действия злоумышленника в данном случае не приводят к изменению выбранной экспертами альтернативы, а могут приводить лишь к изменению процентного соотношения ответов, полученных за разные альтернативы.

Необходимо отметить, что данное утверждение верно в случае отсутствия угроз по блокированию ответов экспертов ($P_{block} = 0$), т.к. существует вероятность блокирования ответов экспертов, поданных за альтернативу «0», учитывая, что угрозы по блокированию (отказы оборудования, разрывы линий связи, «зависание» программного обеспечения) носят случайный характер.

Возможные события по выбору альтернативы ЛПР. Далее будут рассматриваться только случаи голосования, когда количество экспертов, проголосовавших за альтернативу «1», равно или превышает количество экспертов, проголосовавших за альтернативу «0» ($n \geq m/2$), т.е. те случаи, когда действия злоумышленника могут привести к изменению выбранной альтернативы. При этом возможны следующие события:

$Ans w1$ – количество дошедших до ЛПР ответов за альтернативу «1» больше, чем за альтернативу «0»; лицом, принимающим решение, выбирается альтернатива «1»;

$Ans wEq$ – количество дошедших до ЛПР ответов за альтернативы «1» и «0» равно, назначается повторное голосование;

$Ans w0$ – количество дошедших до ЛПР ответов за альтернативу «1» меньше, чем за альтернативу «0»; лицом, принимающим решение, выбирается альтернатива «0».

Событие $Ans wEq$ возможно при четном количестве экспертов m , принимающих участие в опросе (в случае отсутствия угроз по блокированию ответов экспертов).

События $Ans w1$, $Ans wEq$, $Ans w0$ образуют полную группу событий.

Условия возникновения событий $Ans wEq$, $Ans w0$. Событие $Ans w0$ возникает только под действием угроз информационной безопасности, действия злоумышленника достигают своей цели полностью (при условии $n \geq m/2$).

Событие $Ans wEq$ может возникнуть как под действием угроз информационной безопасности (при $n > m/2$), так и в случае, когда эксперты проголосовали поровну за альтернативы «0» и «1» ($n = m/2$) и их ответы в процессе передачи не изменялись. При возникновении события $Ans wEq$ возможно назначение повторного голосования.

Количество экспертов (из n экспертов, проголосовавших за альтернативу «1»), ответы которых необходимо изменить в процессе передачи для того, чтобы произошло событие $Ans wEq$:

$$n - m/2, \quad m - \text{четное.}$$

Количество экспертов (из n экспертов, проголосовавших за альтернативу «1»), ответы которых необходимо изменить в процессе передачи для того, чтобы произошло событие $Ans w0$:

$$n - m/2 + 1, n - m/2 + 2, \dots, n, \quad m - \text{четное;}$$

$$n - (m+1)/2 + 1, n - (m+1)/2 + 2, \dots, n, \quad m - \text{нечетное.}$$

Вероятности наступления событий

Вероятность наступления события $Ans wEq$:

$$P(Ans wEq) = (P_{change})^{(n-m/2)} \cdot (1 - P_{change})^{(n-(n-m/2))} \cdot C_n^{n-m/2}, \quad (1)$$

где $C_n^{n-m/2}$ – число сочетаний из n по $n-m/2$.

$$C_n^{n-m/2} = \frac{n \cdot (n-1) \cdot \dots \cdot (n-(n-m/2+1))}{1 \cdot 2 \cdot \dots \cdot (n-m/2)} = \frac{n!}{(n-m/2)! \cdot (n-(n-m/2))!} = \frac{n!}{(n-m/2)! \cdot (m/2)!}, \quad (2)$$

$$P(\text{AnswEq}) = (P_{\text{change}})^{(n-m/2)} \cdot (1-P_{\text{change}})^{(n-(n-m/2))} \cdot \frac{n!}{(n-m/2)! \cdot (n-(n-m/2))!}, \quad (3)$$

$$P(\text{AnswEq}) = (P_{\text{change}})^{(n-m/2)} \cdot (1-P_{\text{change}})^{(m/2)} \cdot \frac{n!}{(n-m/2)! \cdot (m/2)!}. \quad (4)$$

Вероятность наступления события Answ0 :

$$P(\text{Answ0}) = \sum_i P(A_i), \quad (5)$$

где P_i – вероятность появления события A_i .

Событие A_i – изменение при передаче до ЛПП ровно i ответов экспертов из n экспертов, проголосовавших за альтернативу «1».

$$i = \{n-m/2+1, n-m/2+2, \dots, n\}, \quad \text{если } m - \text{четное};$$

$$i = \{n-(m+1)/2+1, n-(m+1)/2+2, \dots, n\}, \quad \text{если } m - \text{нечетное}.$$

$$P(A_i) = (P_{\text{change}})^i \cdot (1-P_{\text{change}})^{(n-i)} \cdot C_n^i, \quad (6)$$

где C_n^i – число сочетаний из n по i

$$C_n^i = \frac{n \cdot (n-1) \cdot \dots \cdot (n-i+1)}{i!} = \frac{n!}{i! \cdot (n-i)!}, \quad (7)$$

$$P(\text{Answ0}) = \sum_i ((P_{\text{change}})^i \cdot (1-P_{\text{change}})^{(n-i)} \cdot C_n^i) = \sum_i ((P_{\text{change}})^i \cdot (1-P_{\text{change}})^{(n-i)} \cdot \frac{n!}{i! \cdot (n-i)!}), \quad (8)$$

$$i = \{n-m/2+1, n-m/2+2, \dots, n\}, \quad \text{если } m - \text{четное};$$

$$i = \{n-(m+1)/2+1, n-(m+1)/2+2, \dots, n\}, \quad \text{если } m - \text{нечетное}.$$

Вероятность наступления события Answ1 :

$$P(\text{Answ1}) = 1 - (P(\text{AnswEq}) + P(\text{Answ0})), \quad (9)$$

так как события Answ1 , AnswEq , Answ0 образуют полную группу событий.

При возникновении события AnswEq должно назначаться повторное голосование до тех пор, пока не возникнет событие Answ1 или Answ0 . При этом условная вероятность событий Answ1 и Answ2 при условии, что произошло одно из этих событий, составляет:

$$P(\text{Answ1} | \text{Answ1} \cup \text{Answ0}) = \frac{P(\text{Answ1})}{P(\text{Answ1}) + P(\text{Answ0})}, \quad (10)$$

$$P(\text{Answ0} | \text{Answ1} \cup \text{Answ0}) = \frac{P(\text{Answ0})}{P(\text{Answ1}) + P(\text{Answ0})}. \quad (11)$$

Необходимо отметить, что исключением является случай возникновения события AnswEq , когда угрозы ИБ отсутствуют и эксперты голосуют поровну за альтернативы «0» и «1». При дальнейшем описании указанный случай ($n = m/2$, m – четное) рассматриваться не будет.

Вероятности выбора альтернатив «0» и «1». Учитывая повторные голосования, назначаемые при возникновении события AnswEq , вычислим полные вероятности выбора альтернативы «0» (событие $A0$) и альтернативы «1» (событие $A1$).

$$P(A0) = P(\text{Answ0}) + P(\text{AnswEq}) \cdot P(\text{Answ0} | \text{Answ1} \cup \text{Answ0}), \quad (12)$$

$$P(A1) = P(\text{Answ1}) + P(\text{AnswEq}) \cdot P(\text{Answ1} | \text{Answ1} \cup \text{Answ0}). \quad (13)$$

Вероятность события $A0$ – вероятность того, что действия злоумышленника приводят к изменению выбранной экспертами альтернативы.

Вероятность события $A1$ – вероятность того, что действия злоумышленника не приводят к изменению выбранной экспертами альтернативы.

События $A0$ и $A1$ образуют полную группу событий.

$$P(A0) + P(A1) = 1. \quad (14)$$

Следовательно, для нахождения вероятностей их наступления достаточно найти вероятность наступления только одного события, например $P(A0)$. Из формулы (12), используя формулу (11), получаем:

$$P(A0) = P(\text{Answ0}) + P(\text{AnswEq}) \cdot \frac{P(\text{Answ0})}{P(\text{Answ1}) + P(\text{Answ0})} = P(\text{Answ0}) \cdot \left(1 + \frac{P(\text{AnswEq})}{P(\text{Answ1}) + P(\text{Answ0})}\right). \quad (15)$$

Так как события $Answ1$, $AnswEq$, $Answ0$ образуют полную группу событий, получаем

$$P(A0) = P(Answ0) \cdot \left(1 + \frac{P(AnswEq)}{1 - P(AnswEq)}\right) = \frac{P(Answ0)}{1 - P(AnswEq)}. \quad (16)$$

Вероятности $P(AnswEq)$, $P(Answ0)$ находятся по формулам (4), (8).

Программная реализация. Для вычисления вероятности события $A0$ (изменение выбранной экспертами альтернативы в процессе передачи) в зависимости от вероятности изменения ответов экспертов P_{change} , количества экспертов m и количества экспертов, проголосовавших за альтернативу «1», разработан модуль на языке программирования MATLAB.

Результаты вычислений. На рис. 1 представлена зависимость вероятности $P(A0)$ от вероятности изменения отдельного ответа эксперта P_{change} , но при постоянном соотношении количества ответов экспертов за альтернативы «0» и «1» ($n/m = 0,6$). На графике приведены зависимости при $m = \{10, 20, 30, 40, 50\}$.

Из рис. 1 видно, что существует определенное значение P_{change} , при котором вероятности наступления события $A0$ для различных значений m (при $n/m = \text{const}$) одинаковы.

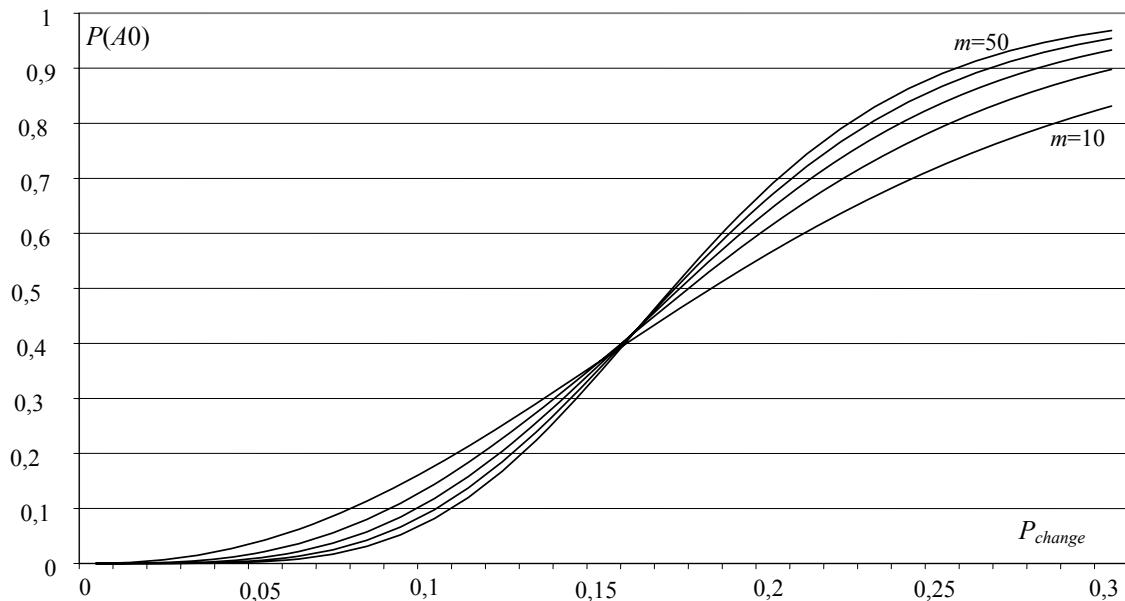


Рис. 1. Вероятность события $A0$ в зависимости от P_{change} при разных значениях m ($n/m = 0,6 = \text{const}$)

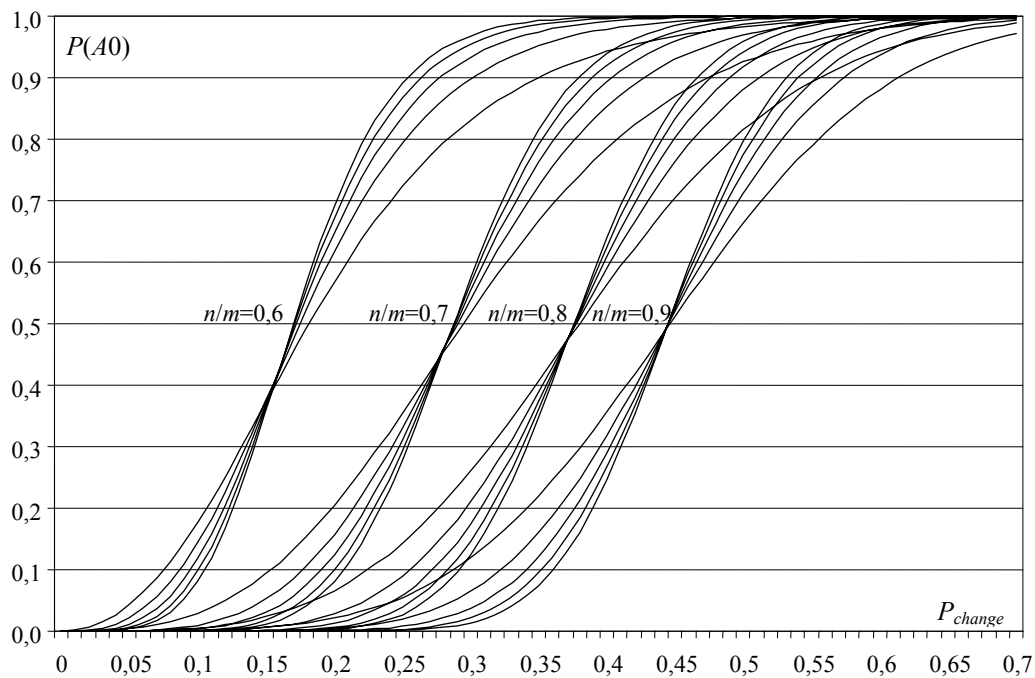


Рис. 2. Группы диаграмм зависимости $P(A0)$ от P_{change}

При $P_{change} < P'_{change}$ увеличение количества экспертов m приводит к уменьшению вероятности изменения выбранной экспертами альтернативы.

При $P_{change} > P'_{change}$ увеличение количества экспертов m приводит к увеличению вероятности изменения выбранной экспертами альтернативы.

На рис. 2 построены группы диаграмм, отображающих зависимость вероятности $P(A0)$ от вероятности изменения отдельного ответа эксперта P_{change} . Каждая группа графиков построена для различных значений n/m , т.е. для различных распределений ответов экспертов.

Как видно из рис. 2, для разных групп графиков (т.е. для разных соотношений n/m) значения P'_{change} будут различны.

На рис. 3 представлена зависимость P'_{change} от отношения n/m . То есть для предполагаемого соотношения n/m (доля экспертов, проголосовавших за альтернативу «1») и ожидаемой вероятности P_{change} изменения злоумышленником ответа эксперта в пользу альтернативы «0» можно дать рекомендацию о необходимости увеличения или уменьшения количества экспертов для уменьшения влияния действий злоумышленника на решение, принимаемое ЛПП.

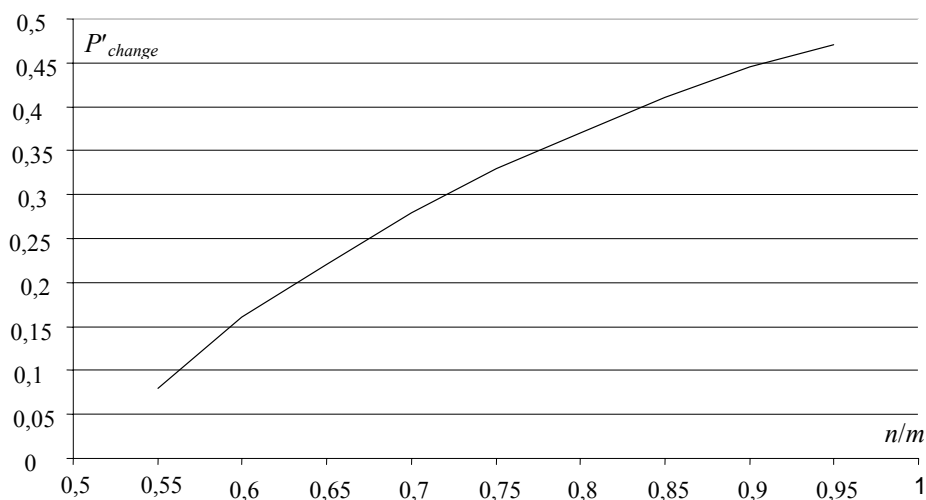


Рис. 3. Вероятность P'_{change} в зависимости от отношения n/m

Заключение. В статье показано, что для любого значения относительного количества экспертов, проголосовавших за альтернативу «1» (n/m), существует критическая вероятность изменения ответа эксперта под действием угроз информационной безопасности P'_{change} , при которой увеличение абсолютного значения количества экспертов приводит к увеличению вероятности принятия ложного решения (при $0,5 \leq n/m \leq 1$).

Так, при $n/m \rightarrow 0,5$ значение $P'_{change} \rightarrow 0$. Необходимо отметить, что при $n/m \rightarrow 0,5$ достаточно небольшой вероятности реализации угроз по изменению ответов экспертов P_{change} , чтобы было принято ложное решение. В данном случае возможным вариантом действий по проверке устойчивости результата может являться проведение ряда случайных выборок размером $m_1 < m$. При этом пространственная (сетевая) локализация экспертов, от которых лицом принимающим решение, получены результаты голосования за альтернативу «0», может служить основанием для подозрения об изменении их результатов (всех или части экспертов) вследствие реализации угроз информационной безопасности.

Литература

1. Ларичев О.И. Теория и методы принятия решений / О.И. Ларичев. – М.: Университетская книга, Логос, 2006. – 296 с.
2. Орлов А.И. Теория принятия решений / А.И. Орлов. – М.: Экзамен, 2005. – 656 с.
3. Ефимов Б.И. Обеспечение информационной безопасности систем принятия решений с использованием теории графов / Б.И. Ефимов, Р.Т. Файзуллин // Динамика систем, механизмов и машин: матер. VII Междунар. науч.-техн. конф. – Кн. 1. – Омск: Изд-во ОмГТУ, 2009. – С. 280–284.
4. Ефимов Б.И. Применение алгоритмов теории графов для решения задач, связанных с обеспечением информационной безопасности в системах принятия решений // Системы управления и информационные технологии. – 2009. – № 1.3 (35). – С. 342–346.

Ефимов Борис Игоревич

Аспирант каф. «Комплексная защита информации»
Омского государственного технического университета (ОмГТУ)
Тел.: (381-2) 79-94-22
Эл. почта: b_efimov@mail.ru

Файзуллин Рашит Тагирович

Д-р техн. наук, профессор, проректор по информатизации ОмГТУ
Тел.: (381-2) 62-87-07
Эл. почта: r.t.faizullin@mail.ru

Efimov B.I., Faizullin R.T.

The probability of making a false solutions under the influence of information security threats in decision-making systems with experts

Proposed solution to the problem of calculating the probability of a false solution in decision-making systems with experts under the influence of information security threats, aimed at changing the experts' responses. The graphs showing the feasibility of increasing the number of experts to reduce the influence of the attacker on the decision are obtained.

Keywords: information security, decision-making systems, experts, threats, probability theory.

УДК 004.4'423

Ю.А. Зорин

Интерпретатор языка построения генераторов тестовых заданий на основе деревьев И/ИЛИ

Предложен язык разработки генераторов тестовых заданий GILT, основанный на методах генерации комбинаторных множеств деревьев И/ИЛИ. Рассматривается синтаксический анализ языка и его реализация на функциональном языке F#, а также пример построения алгоритма генерации тестового задания на языке GILT.

Ключевые слова: генерация тестовых заданий, функциональный язык, дерево И/ИЛИ.

Автоматизированные системы генерации тестовых заданий являются важной частью современных систем дистанционного обучения [1]. Эти системы, как правило, обладают некоторым языком описания генераторов тестовых заданий и интерпретатором, позволяющим получить конкретный вариант задания в виде некоторого файла в заданном формате. В свою очередь, реализация алгоритмов генерации в подобных системах основана на применении различных языков программирования:

- общего назначения (C++, C, Pascal и др.) [2];
- специального назначения (Possum, «Фея») [1, 3];
- системы компьютерной алгебры (Mathematica, Maxima, MathCad).

Существует также ряд библиотек функций и классов, реализующих алгоритмы генерации для языков программирования общего назначения. К подобным относится интерпретатор языка GIL (Generation and Identification Language). Данный язык предназначен для построения генераторов комбинаторных множеств на основе деревьев И/ИЛИ [4]. Язык предусматривает различные операции над деревьями И/ИЛИ и также применим для описания генераторов тестовых заданий в виде узлов дерева И/ИЛИ, с определенными ограничениями:

- отсутствие операторов выбора варианта дерева по условию;
- отсутствие именованных узлов для идентификации варианта.

Подобные ограничения не позволяют описать генератор вопроса с различными решениями в зависимости от варианта дерева И/ИЛИ, а также корректировать условие задания в зависимости от ранее выбранных параметров. В свою очередь, метод описания генераторов тестовых заданий в виде деревьев И/ИЛИ позволит производить идентификацию генератора задания, а также его вариантов, рассчитывать мощность генератора, что позволит получать вариант задания по его номеру. Последние утверждения показывают необходимость в развитии языка GIL и создании на его основе синтаксиса языка и его интерпретатора для построения генераторов тестовых заданий на основе деревьев И/ИЛИ (GILT – Generation and Identification Language Tests).

В данной статье описывается реализация интерпретатора языка GILT и его возможности для генерации тестовых заданий.

Язык генерации тестовых заданий GILT. Язык GILT является развитием функционального языка GIL. Язык позволяет описывать узлы дерева И/ИЛИ в виде скобочной нотации. Для записи узла И предлагается использовать круглые скобки, а для узла ИЛИ – фигурные, именованные узлы содержат название перед скобками (рис. 1).

Введем понятие «условно-именованный узел» (обозначается символом решетки – «#», устанавливаемый перед именованным узлом), а также «условный узел». Условно-именованные узлы участвуют при инициализации дерева и могут переопределяться в зависимости от выбранного варианта. Условный узел является узлом И, состоящим из трех сыновей: условие, первый сын при условии True, второй – False. Узлы дерева могут быть нескольких типов:

- математическое выражение;
- строка;
- дерево;
- условное выражение (в условных узлах).

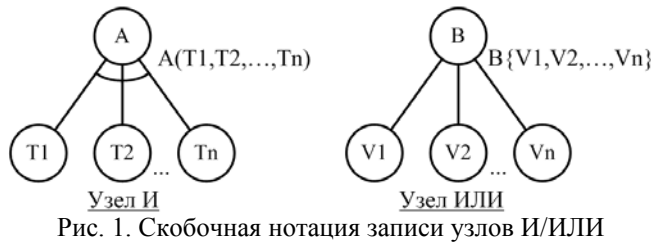


Рис. 1. Скобочная нотация записи узлов И/ИЛИ

Любой алгоритм генерации на языке GILT представляется в виде дерева И/ИЛИ. Результатом работы является получение варианта при левостороннем обходе дерева и выборе варианта (случайным образом либо по номеру) каждого из узлов ИЛИ (рис. 2).

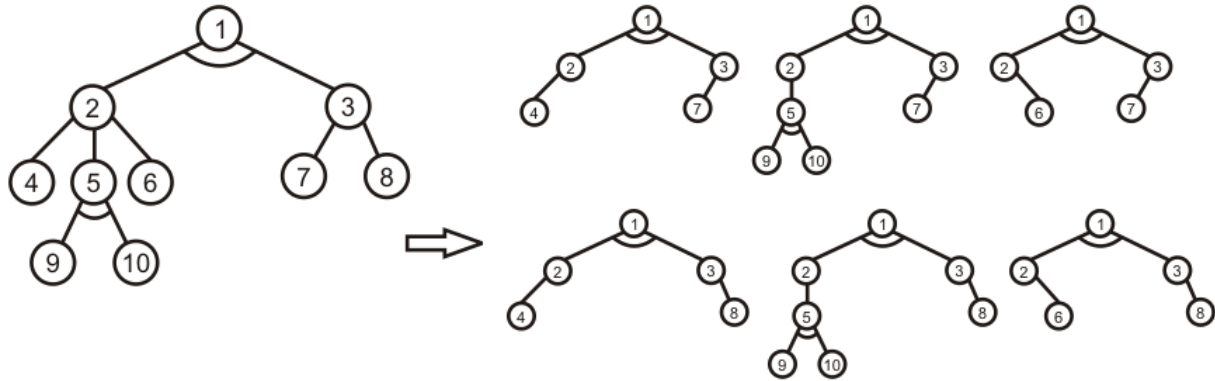


Рис. 2. Дерево И/ИЛИ и все его варианты

Пример построения генератора тестовых заданий на языке GILT. Рассмотрим пример построения генератора задачи по теории вероятности и его описание на языке GILT.

В ящике имеется 10 белых и 15 черных шаров. Вынимается 4 шара. Какова вероятность того, что все вынутые шары будут одного цвета?

Для описания алгоритма генерации данной задачи представим текст как совокупность строковых символов. Данное действие позволит получать различные варианты представления данной задачи за счет изменения заранее определенных групп символов (слов, чисел и т.д.). Весь текст условия задачи разделим на три фрагмента {A, B, C}, тогда узел Q будет И-узлом, содержащим сыновей A, B, C. Каждый из фрагментов разделяется на фиксированные и переменные части. Например, узел B разбивается на три фиксированных фрагмента {T4, T8, T12} и два переменных {V2, V3}, которые имеют по три варианта реализации. Ниже в табл. 1 перечислены значения узлов $\{TN\}_{N=1}^{15}$ [5].

Таблица 1

Значения узлов

T1 – В	T9 – 8
T2 – ящике	T10 – 10
T3 – коробке	T11 – 7
T4 – имеется	T12 – черных шаров
T5 – 10	T13 – вынимается
T6 – 15	T14 – 2 шара. Какова вероятность того, что вынутые шары будут разных цветов.
T7 – 20	T15 – 4 шара. Какова вероятность того, что все вынутые шары будут одинакового цвета
T8 – белых	

Для наглядного отображения всех возможных вариантов условий данной задачи, построим дерево И/ИЛИ (рис. 3).

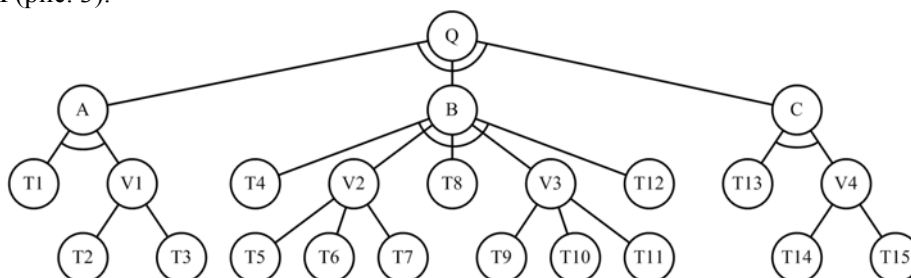


Рис. 3. Схематичное представление задачи в виде дерева И/ИЛИ

Используя правила записи узлов [4]: узлы И – обозначаются в круглых скобках, узлы ИЛИ – в фигурных скобках, именованные узлы содержат название перед скобками, запишем дерево (см. рис. 2) на языке GILT:

```
Q(
  A(T1, V1{T2, T3}),
  B(T4, V2{T5, T6, T7}, T8, V3{T9, T10, T11}, T12),
  C(T13, V4{T14, T15})
)
```

Условие задачи может иметь некоторую другую интерпретацию рассмотренного тестового задания, в котором меняются условия задания в зависимости от варианта одного из узлов:

В ящике имеется 10 белых и 15 черных шаров. Из ящика вынимаются 4 шара. Какова вероятность того, что все вынутые шары будут белого цвета?

При попытке вставки в изменяемый узел ИЛИ слова «белых» с вариантами «белых», «красных», «синих» возможен следующий исход задания:

В ящике имеется 10 красных и 15 черных шаров. Из ящика вынимается 4 шара. Какова вероятность того, что все вынутые шары будут белого цвета?

Как видно, в тексте вопроса осталось слово «белого», что делает задание некорректным. В предотвращении подобной ситуации необходимо использовать условно-именованные узлы, заменив узел T8 на узел ИЛИ и добавив узел T16 в переменный V4 следующим образом:

```
Q(
  A(T1, V1{T2, T3}),
  B(T4, V2{T5, T6, T7}, V5{(T17, #T16("белого")), (T18,
#T16("красного")), (T19, #T16("синего"))}), V3{T9, T10, T11}, T12),
  C(T13, V4{T14, T15, (T20, T16, T21)})
)
```

В табл. 2 приведены значения узлов $\{TN\}_{N=17}^{21}$.

Таблица 2

Значения узлов ИЛИ

T17 – белых и	T20 – 6 шаров. Какова вероятность того, что все вынутые шары будут
T18 – красных и	T21 – цвета
T19 – синих и	

Узел T16 будет инициализироваться в момент генерации того или иного варианта переменного узла V5, но его значение будет использовано, например, для вывода на экран, только в узле V4. Особенность условно-именных узлов, инициализирующихся только в момент выбора того или иного варианта дерева И/ИЛИ, позволит описывать ответы и решения для той или иной задачи.

Разработка интерпретатора GILT. Для реализации программного интерпретатора синтаксической записи дерева И/ИЛИ необходимо учесть ряд требований, предъявляемых к языку GILT:

- учесть рекурсивную природу деревьев;
- реализацию алгоритмов генерации без решения инфраструктурных задач;
- доступ к генераторам случайных чисел;
- наличие математических выражений;
- наличие условного оператора;
- возможность идентификации объектов для их последующего использования.

Функциональная парадигма программирования характеризуется рекурсивными функциями, тем самым предоставляя более удобные возможности обработки древовидных структур. Язык F# предоставляет полный набор инструментов функционального программирования: алгебраические типы данных, функции высшего порядка, средства для композиции функций и неизменяемые структуры данных. Все функциональные возможности F# реализованы поверх общей системы типов .NET Framework [6].

Неотъемлемой частью реализации интерпретатора является лексический и синтаксический анализ введенных пользователем строк, позволяющий переводить скобочную запись деревьев И/ИЛИ в описанный тип данных дерева И/ИЛИ. Среди различных анализаторов кода внимания заслуживает программная библиотека «Yacc», в частности, из-за того, что в своем составе содержит сразу лекси-

ческий и синтаксический анализаторы и позволяет выдавать результат на языке F# (FYacc). Для описания лексического анализатора на языке F# приведена синтаксическая диаграмма дерева И/ИЛИ (рис. 4).

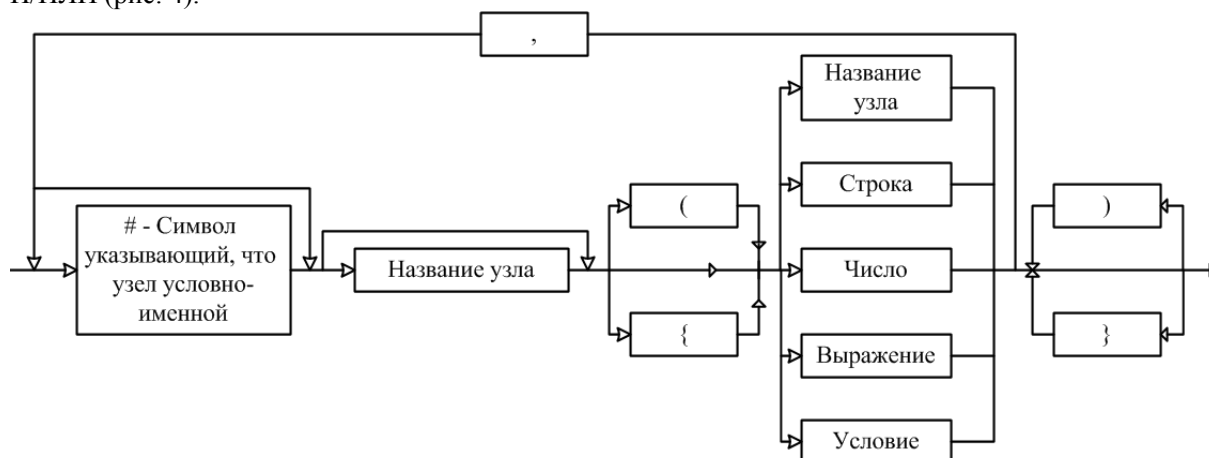


Рис. 4. Синтаксическая диаграмма языка GILT

Синтаксический анализатор на основе символьных конструкций, определенных за счет регулярных выражений, использует рекурсивный спуск при попытке приведения полученной строки к типу Tree, которое в свою очередь будет описано в размеченном объединении:

```
start:
| Prog1          { Tree($1) }
Tree:
| LOR Tree ROR   { Or ([] @ $2) }
| LAND Tree RAND { And ([] @ $2) }
| NAMETREE Tree { SetNameTree($1, $2) }
| STR            { Str $1 }
| INT32          { Int $1 }
| FLOAT         { Float $1 }
```

Размеченные объединения языка F# позволяют описать типы, учитывая рекурсивную природу деревьев И/ИЛИ, и, используя сопоставление с образцом, сразу приступить к применению различных операций над листьями и узлами дерева [6]. Размеченное объединение для дерева И/ИЛИ:

```
type AndOrTree =
//Узел может содержать один или несколько узлов И
| And of AndOrTree list
//Узел может содержать один или несколько узлов ИЛИ
| Or of AndOrTree list
//Листья дерева могут содержать целочисленные значения
| Int of int
//Листья дерева могут содержать строковые значения
| String of string
//Листья дерева могут содержать значения с плавающей точкой
| Float of float
//Узлы могут быть именованные
| SetNameTree of string * AndOrTree
//Узлы могут содержать имя узла, инициализированного ранее
| GetNameTree of string * AndOrTree
// Обозначаем корневой узел
and Tree =
| Tree of AndOrTree
```

К основным характеристикам интерпретатора языка GILT относятся расширяемость и переносимость в различные проекты, разрабатываемые на платформе .NET.

Заключение. Реализация интерпретатора языка GILT даст возможность построения генераторов тестовых задач на основе деревьев И/ИЛИ, что позволит производить идентификацию генератора задания, а также его вариантов, рассчитывать мощность генератора и получать вариант задания

по его номеру. Скобочная нотация записи дерева И/ИЛИ позволит разрабатывать генераторы тестовых заданий на основе визуальных компонентов построения дерева в составе инструментальной системы разработки генераторов тестовых заданий и решить ряд проблем, связанных с отсутствием у преподавателей навыков программирования при разработке генераторов.

Литература

1. Посов И.А. Автоматизация процесса разработки и использования многовариантных учебных заданий: автореф. дис. ... канд. техн. наук. – СПб.: СПбГУ, 2012. – 18 с.
2. Левинская М.А. Автоматизированная генерация заданий по математике для контроля знаний учащихся // *Educational Technology & Society*. – 2002. – Т. 5, вып. 4. – С. 214–221.
3. Кручинин В.В. Генераторы в компьютерных учебных программах. – Томск: Изд-во Том. ун-та, 2003. – 200 с.
4. Титков А.В. Система построения генераторов комбинаторных множеств на основе деревьев И/ИЛИ: автореф. дис. ... канд. техн. наук / А.В. Титков. – Томск: ТУСУР, 2010. – 22 с.
5. Кручинин В.В. Методы генерации тестовых заданий по информатике // *Информатика и образование* (Москва). – 2005. – № 2. – С. 87–93.
6. Смит К. Программирование на F#: учеб. пособие. – М.: Символ-Плюс, 2011. – 448 с.

Зорин Юрий Алексеевич

Аспирант каф. промышленной электроники ТУСУРа

Тел.: 8(3822) 42-30-67

Эл. почта: yura@freebrains.ru

Zorin Yu.A.

The interpreter of programming language for design generators of tests based on AND/OR trees

The language of design generators of tests GILT, based on the methods of generating combinatorial sets of AND/OR trees. The article deals with parsing the language and its implementation in the functional language F#, as well as the example of the designed algorithm for generating the tasks in GILT.

Keywords: generator of tests, functional language, AND/OR tree.

УДК 004.056.53

А.О. Шемяков

Причины повышения уязвимости и снижения стойкости функций безопасности автоматизированных систем вуза

На основе комплексного подхода рассматриваются основные причины повышения уязвимости и снижения стойкости функций безопасности автоматизированных систем высшего учебного заведения. Приводится анализ последствий автоматизации процессов обеспечения безопасности и связанных с ней изменений системы «система защиты – объект защиты».

Ключевые слова: автоматизированная система, защита информации, объект защиты, информационная безопасность.

Следствием применения автоматизированных систем (АС) является взаимная интеграция объекта информатизации и системы защиты. Лежащие в основе этой интеграции современные информационные технологии с каждым годом все больше стирают грани различия между первым и вторым. Это обстоятельство выдвигает новые требования к процессам проектирования и разработки систем управления, в том числе и высших учебных заведений. Суть этих требований состоит в необходимости дополнительного учета топологии и свойств объекта защиты. Существующие естественные тенденции к усложнению объектов защиты требуют соответственно создания более сложных АС.

Анализ ретроспективы развития технологий использования АС и перспектив их дальнейшего совершенствования позволяет выявить устойчивую тенденцию интегрирования компьютерных средств и средств связи в рамках нового класса сложных АС – информационно-телекоммуникационных. Это обстоятельство определило вхождение в оборот таких новых понятий, как *информатизация* [6], *изделия информационных технологий*, *объект информатизации* [1], *инфотелекоммуникационная инфраструктура* и многих других. Вместе с тем результатом автоматизации является не только «технический симбиоз» слабо формализуемых систем «объект защиты – система защиты», но и целый ряд связанных с ними новых проблем – проблем информационной безопасности. Большинство из них возникли после осуществления процесса автоматизации и являются его непосредственным следствием [2].

На рис. 1 показана обобщенная схема совместной эволюции объекта и системы защиты, движущей силой которой является мировой научно-технический прогресс вообще и прогресс современных информационных технологий в частности.

Контрастирующая простота абстрактной схемы «технического симбиоза» системы и объекта защиты 60–70-х годов определяется тем обстоятельством, что на начальном этапе система социальных, экономических, технологических и иных причинно-следственных взаимосвязей в этой сфере деятельности складывалась постепенно. Число такого рода взаимосвязей было невелико и с точки зрения возможности учета их в исследованиях имело незначительный прирост во времени.

Это давало возможность, опираясь на накопленный опыт, легко искать исторические прецеденты, медленно «эволюционно» совершенствовать объект исследований. Актуальность внутренних угроз была незначительной. В настоящее же время многократно возросло число внутренних взаимосвязей. Возникла необходимость учета большого числа причинно-следственных взаимосвязей внешних факторов.

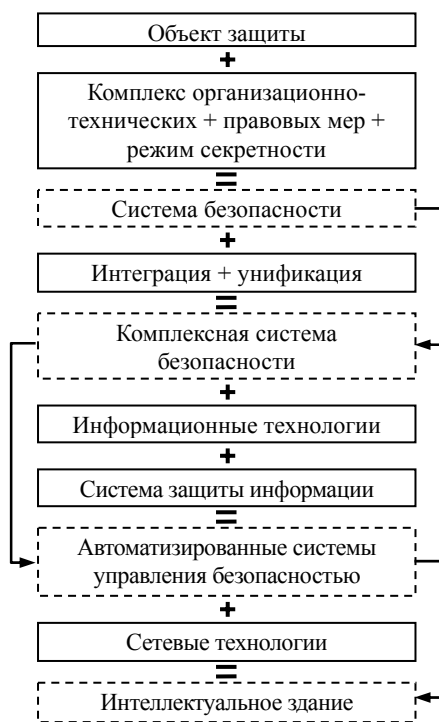


Рис. 1. Обобщенная схема «эволюции» технологий защиты

Значительно повысилась разнородность природы и механизмов структурных составляющих системы и объекта защиты. Произошло усиление их интравертности, приведшее к тому, что развитие их подсистем в большей мере стало осуществляться по своим внутренним закономерностям [2].

Возросло число угроз, явившихся следствием автоматизации и включения систем и объектов защиты в пространство современных информационных технологий. Стала актуальной модель внутреннего нарушителя, являющегося продуктом той же самой научно-технической и культурной парадигмы, что и разработчик средств защиты.

В общем случае последствия «проблемных последствий» автоматизации процессов обеспечения безопасности и связанных с ней изменений системы «система защиты – объект защиты» можно сформулировать в следующем виде.

Во-первых, возрос риск принятия неэффективных решений (вероятность несоответствия реально полученных результатов реализованного решения априорно поставленным целям) по вопросам обеспечения безопасности объектов повышенной потенциальной опасности и критических информационных приложений.

В условиях динамично изменяющегося многообразия внешних и внутренних причинно-следственных взаимосвязей, сопровождающих процессы автоматизации и эволюции рассматриваемого класса систем, не всегда, оказывается, бывает достаточно имеющегося опыта, для того чтобы можно было бы судить, насколько дорого и опасно то или иное принимаемое решение [3].

Во-вторых, возросло число угроз, методы осуществления которых базируются на определенных свойствах АС, обеспечивающих безопасность объекта, которые как бы провоцируют появление средств нападения.

Сегодня нарушитель, получив контроль над объектом повышенной потенциальной опасности с «интеллектуальной» (компьютеризированной) системой управления, может трансформировать его энергетические, материальные и информационные ресурсы в уже им управляемые средства для достижения других, более опасных своих целей.

Это становится возможным за счет возрастающего числа недеklarированных возможностей, которыми «автоматически» наделяются системы, включаемые в пространство современных информационных технологий.

Задействование скрытых функциональных ресурсов защищаемых АС, т.е. способностей систем или их подсистем выполнять «по совместительству» новые не предусмотренные функции, является основой для существования целого класса новых типов угроз – угроз потери контроля за системой и (или) ее несанкционированного применения.

В-третьих, на фоне развития высоких технологий информационной и общей безопасности возросли уязвимости систем, обеспечивающих эти виды безопасности и уязвимость самого объекта защиты.

Статистические данные по росту числа атак в информационных системах (рис. 2, 3) [7] отчасти являются тому подтверждением.

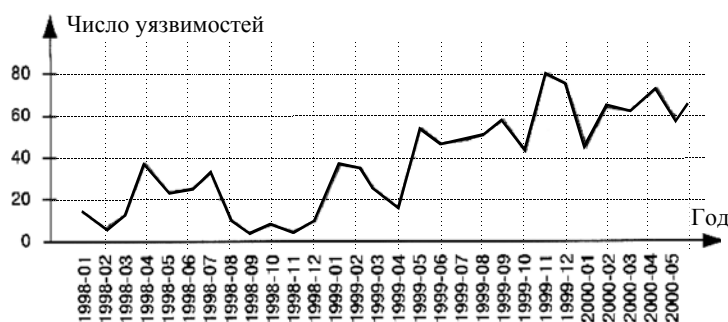


Рис. 2. Современные тенденции изменения числа ежемесячно обнаруживаемых уязвимых мест в программном обеспечении

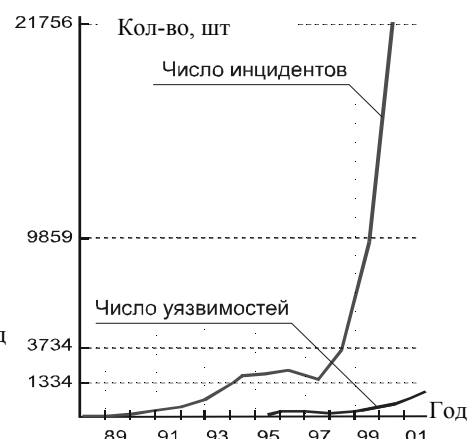


Рис. 3. Статистика роста числа инцидентов и уязвимостей в глобальных сетях передачи данных

Ежегодные отчеты Института компьютерной безопасности – CSI [4:www.gocsi.com] и Координационного центра немедленного реагирования США – CERT [5:www.cert.org] также подтверждают статистику роста числа информационных атак.

Таким образом, автоматизация процесса доступа к информационным ресурсам объекта защиты, в случае нарушения безопасного поведения соответствующей подсистемы защиты, обеспечивает нарушителю возможность воспользоваться этой автоматизацией и «скачать» на магнитный носитель информации в сотни, а то и тысячи раз больше, по сравнению, например, с «традиционными» методами фотокопирования ее с бумажного носителя.

Литература

1. ГОСТ Р 51275-99. Защита информации. Объект информатизации. Факторы, воздействующие на информацию. Общие положения [Электронный ресурс]. – Режим доступа: <http://vsegost.com/Catalog/86/8680.shtml>, свободный (дата обращения: 03.04.2013).
2. Зегжда Д.П. Основы безопасности информационных систем / Д.П. Зегжда, А.М. Ивашко. – М.: Горячая линия-Телеком, 2000. – 452 с.
3. Лазарев И.А. Информация и безопасность. Композиционная технология информационного моделирования сложных объектов принятия решений. – М.: Московский городской центр научно-технической информации, 1997. – 336 с.
4. Отчет Института компьютерной безопасности – CSI [Электронный ресурс]. – Режим доступа: <http://www.gocsi.com>, свободный (дата обращения: 02.04.2013).
5. Отчет Координационного центра немедленного реагирования США – CERT [Электронный ресурс]. – Режим доступа: <http://www.cert.org>, свободный (дата обращения: 03.04.2013).
6. Российская Федерация. Федеральный закон № 24 от 20 февраля 1995 г.: Об информации, информатизации и защите информации. – М., 1995.
7. Сердюк В.А. Криминализация глобальных информационных систем: миф и реальность // Системы безопасности связи и телекоммуникаций. – М.: Гротек, 2000. – № 35. – С. 84–87.

Шемяков Александр Олегович

Ассистент каф. 402 Московского авиационного института
(национального исследовательского университета)
Тел.: 910-423-84-92
Эл. почта: a.shemyakov@gmail.com

Shemyakov A.O.

The reasons of increase of vulnerability and decrease in firmness of functions of safety of the automated systems of higher education institution

On the basis of an integrated approach the main reasons for increase of vulnerability and decrease in firmness of functions of safety of the automated systems of a higher educational institution are considered. The analysis of consequences of automation of processes of safety and the related changes of system «protection system object of protection» is provided.

Keywords: the automated system, information security, object of protection, information security.

УДК 004.75

А.Л. Молчанов, Р.Т. Файзуллин

Изучение свойств эффективности и робастности работы мультиплексной системы

Статья посвящена изучению свойств эффективности и робастности мультиплексных систем. Анализ свойств эффективности и робастности достигается за счет различных и многократных экспериментов по передаче видеофайла.

Ключевые слова: эффективность, робастность, мультиплексная система.

Задача исследования. Сегодня с развитием и глобализацией Internet и используемых технологий передачи данных проблема повышения эффективности функционирования компьютерных сетей встает наиболее остро, затрагивая и подчиняя себе полный спектр различных решений и разработок. На сегодняшний день защищенная передача видеoinформации по вычислительной сети является востребованной задачей. Необходимость в такой защите возникает при передаче конфиденциальной информации, организации видеоконференций, организации платной трансляции видео. Для осуществления защищенной передачи необходимо создать механизм, учитывающий большой объем защищаемых данных (видеоролики высокого качества), однородность передаваемых данных и необходимость потоковой обработки данных в реальном времени. Все чаще для подобных задач используются системы мультиплексной передачи данных. В данной статье будет рассмотрен пример мультиплексной системы передачи данных и проведен анализ ее эффективности и робастности.

Решение задачи. Для проведения экспериментов и изучения свойств эффективности и робастности используем мультиплексную систему. В качестве передаваемой информации используем видеофайл original.avi. Для разбиения видеофайла на части используем ПО VirtualDubMod со встроенным редактором скриптов AviSynth. Для обмена данными между компьютерами используем менеджер XStarter.

Работа данной мультиплексной системы может быть представлена в виде эскиза на рис. 1.

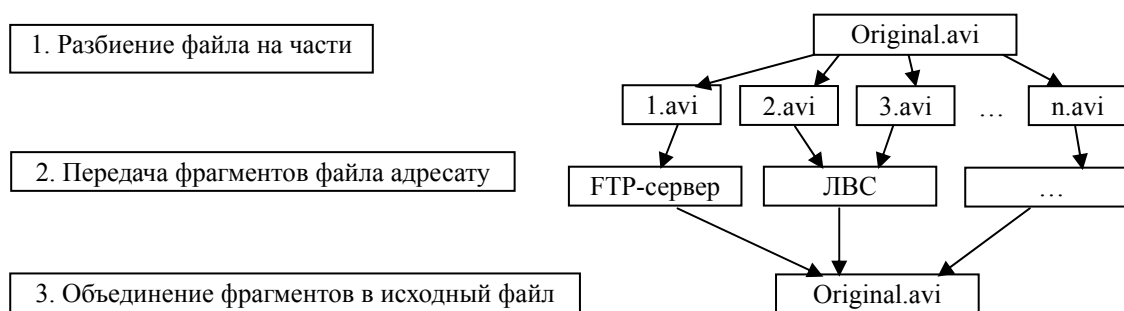


Рис. 1. Схема работы мультиплексной системы

Для первого этапа используем VirtualDubMod со встроенным редактором AviSynth. Пусть n – количество фрагментов разбиения исходного файла N . В результате разбиения avi файла будем иметь n avi-файлов. Редактор AviSynth позволяет разбивать видеофайл покадрово, т.е. наименьшим фрагментом разбиения является 1 кадр. Максимальное количество фрагментов определяется количеством кадров видеофильма.

Пример скрипта, разбивающего видеофайл:

```

A=AviSource("C:\ishod\original.avi")
A1=Trim(A,1,2000)
A1
  
```

Команда AviSource открывает исходный видеофайл. Команда Trim выделяет из файла фрагмент, ограниченный номерами кадров, указанных в скобках.

Использование AviSynth дает возможность применить к разбиению несколько обратимых операций, которые повышают надежность операции передачи. Так, например, команда Reverse дает обратную видеозапись, т.е., применив эту команду на компьютере отправителя, можно формировать фрагменты прямой и обратной записи для передачи, на компьютере получателя восстанавливать исходное изображение.

Для обмена информацией воспользуемся ПО XStarter. Данный менеджер позволяет автоматически забирать фрагменты из заранее определенной папки на компьютере отправителя и отправляет в определенную папку на компьютере получателя, а также запускает программу для объединения фрагментов.

Система XStarter должна быть запущена как на компьютере отправителя, так и на компьютере получателя.

Задача отслеживает изменения в папке Ishot, при обнаружении файла с расширением avi запускается программа VirtualDubMod, в которой исходный файл будет разделен на фрагменты.

Поместим файл original.avi в папку Ishot, задача автоматически перейдет в режим выполнения. Запустим скрипт, который разделит исходный файл. В результате разделения получим в папке Input n -е количество фрагментов файлов. Система XStarter, обнаружив в папке Input файлы с расширением .avi, запустит задачу по передаче фрагментов файлов.

Для передачи фрагментов используем различные способы и протоколы передачи данных.

Рассмотрим передачу фрагментов через FTP-сервер и по локальной сети.

Для проведения экспериментов используется один видеофайл.

Разделим скриптом видеофайл на 5 частей, менеджер XStarter отправит 2 части видеофайла через FTP-сервер, а 3 части – по локальной сети. Запущенная программа XStarter на компьютере получателя, обнаружив в локальной папке avi-файлы, запустит задачу приема файлов с FTP-сервера. Все фрагменты будут собраны в локальной папке на компьютере получателя. Соответственно после окончания выполнения задачи загрузки файлов запустится задача, которая откроет VirtualDubMod для объединения фрагментов файлов в исходный клип.

Таким образом, получаем исходный видеофайл на компьютере получателя. Мультиплексирование проходит на сеансовом уровне модели ISO/OSI.

Рассмотрим графическую функцию выполнения эксперимента, а именно изменение времени выполнения эксперимента при передаче одного и того же файла.

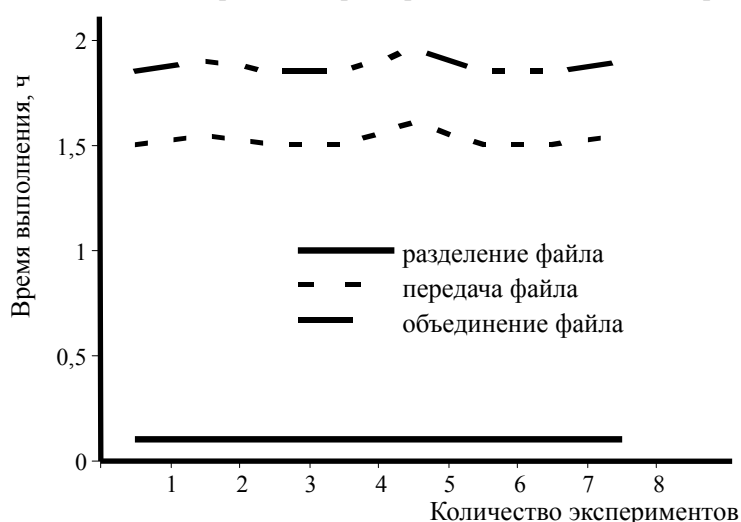


Рис. 1. График изменения времени выполнения экспериментов

Как видно на рис. 1, описанная выше система дает стабильные результаты на этапах разделения и объединения файлов. Так как во всех проведенных экспериментах на компьютер получателя передан файл оригинала, основной функцией для оценки было выбрано время выполнения эксперимента. При выполнении экспериментов не было утеряно ни одного пакета. Во всех экспериментах время разделения файла составило 0,1 ч. Аналогично время объединения пакетов во всех экспериментах стабильно. Небольшие отклонения по времени при передаче частей исходного файла обусловлены посторонними факторами работы сети. Таки образом, при проведении экспериментов вышеописанная система дала стабильные результаты по передаче файла, несмотря на некоторые отклонения при этапе передачи. При выполнении большого количества опытов отклонение от среднего времени передачи файла не превысило 5 мин. Для объективности результата опыт по передаче файла был повторен 35 раз.

Рассмотрим понятие робастности применительно к мультиплексной системе.

Пусть u – проект, мультиплексная система, v – нагрузка, это часть системы, которая не может быть заранее спланирована, предопределена проектом (природа, случай). Проект выбирается из

множества допустимых проектов $u \in U$, нагрузка может принимать любые значения, $v \in E_v$, E_v – пространство значений нагрузки. Рассмотрим случай, когда при создании системы невозможно точно оценить конкретные значения нагрузки, соответственно неизвестно и множество ее возможных значений.

Функционирование системы оценим параметрами a_i и b_j , в формировании значений которых принимает участие как допустимое решение u , так и реализовавшаяся нагрузка v , т.е. $a_i = a_i(u, v)$, $b_j = b_j(u, v)$. Эффективность функционирования системы задается условиями, которые в обобщенном виде можно записать: $a_i(u, v) = 0, b_j(u, v) \leq 0$.

Множество

$$R = \{(u, v) | u \in U, a_i(u, v) = 0, b_j(u, v) \leq 0\}$$

есть отношение, которое задает многозначное отображение $u \rightarrow v$, где каждому проекту u , ставится в соответствие множество эффективно обрабатываемых нагрузок v , называемое множеством робастности проекта u :

$$V(u) = \{v | (u, v) \in R\}.$$

Задача о максимальной робастности есть задача

$$V(u) \rightarrow \max_{u \in U}.$$

Проект (u^*) будет удовлетворять решению данной задачи, если

$$u^* = \arg \max_{u \in U} \xi(V(u)),$$

где $\xi()$ – некоторая функция множества, оценивающая множество робастности $V(u)$ [1].

Основными параметрами оценки эффективности мультимедийной системы могут являться:

- скорость исполнения задания по передаче видеофайла,
- степень надежности программного обеспечения, используемого при передаче,
- уязвимость частей секрета и исходного файла при передаче,
- возможность потери части исходного файла и методы восстановления.

К минусам данной системы можно отнести ограниченность передачи фрагментов возможностями системы XStarter.

При проведении экспериментов система показала стабильные показатели по времени передачи, также потерь и перехват пакетов данных не обнаружено. Возможно такое задание разбиения, при котором потеря даже нескольких частей не повлияет на итоговый результат, при этом отдельные части не дадут представления об оригинале.

В данной системе есть ряд факторов, влияющих на положительный результат, которые не могут быть заранее определены. Так, например, важно наличие необходимых каталогов для автоматического запуска некоторых задач. Условия выполнения задач могут быть изменены пользователями в процессе эксплуатации системы, при этом общая стабильность системы нарушена не будет. При этом могут возникать ситуации неуспешной передачи при неправильной эксплуатации системы XStarter. Возможен вариант, когда получатель не знает заранее количество элементов секрета. В таком случае отправитель может создать скрипт объединения фрагментов и передать его получателю в одном из фрагментов секрета. Для этого возможно использовать методы стеганографии.

Важно также оценить, насколько важна потеря одного из фрагментов при передаче. При достаточно малом размере фрагмента потеря даже нескольких фрагментов не приведет к критическому изменению видеофайла. В противном случае можно предусмотреть фрагменты таким образом, чтобы они накладывались друг на друга, тогда при избирательном объединении фрагмента полученный видеофайл будет очень близок к исходному.

Одним из важных преимуществ данной системы является использование программного обеспечения, находящегося в свободном доступе и не являющегося коммерческим. Это дает возможность построить данную систему в любой сети передачи данных для специалиста, не владеющего основами программирования. Как правило, большинство разработок в этой области базируются на специальном программном обеспечении. Так, в работе П.А. Мельниченко, А.А. Шелупанова [5] в качестве элементов мультимедийной системы выступают различные участники связи, для перехода контрольного пакета данных по произвольным участкам сети потребуется специальное программное обеспечение.

Данная вариация системы не имеет аналогов, тем не менее строится по всем правилам мультиплексной системы. В качестве мультиплексора и демультимплексора используется AviSynth, передающий компонент системы автоматизированный с помощью XStarter.

Не рассматривая в рамках данной статьи уязвимости конкретных способов передачи, остановимся на возможности восстановления исходной информации при перехвате некоторых частей секрета. В случае разбиения исходного файла на достаточно малые части восстановление исходного секрета может быть затруднено. Для усиления степени уязвимости можно использовать обратимые команды AviSynth, такие как воспроизведение записи в обратном направлении, а также некоторые фильтры, которые могут не только убирать лишний «шум» из видеофайла, но и добавлять его.

Что касается степени надежности ПО, используемого в данной системе, то возможна доработка данных систем непосредственно пользователем. Система XStarter на сегодняшний день не располагает возможностями по передаче по всем известным сетям передачи данных и поддерживает ограниченное количество протоколов, фактически позволяя работать с FTP, HTTP, SFTP, электронной почтой, использовать файлообменники [2–4, 6].

Заключение. На основе проведенных исследований возможно сделать следующие выводы:

1. Изученная система передачи данных соответствует мультиплексному способу передачи данных.
2. Изученная система передачи данных может быть эффективна и робастна при соблюдении первоначальных условий приема и передачи данных.
3. Система передачи данных обладает достаточным функционалом для безопасной передачи данных.

Литература

1. Кулагин В.В. Надежность как робастность. Задача о максимальной робастности [Электронный ресурс]. – Режим доступа: http://mmr.gubkin.ru/uploads/submitted_papers/Koulaguin.pdf, свободный (дата обращения: 20.08.2012).
2. Система мультиплексирования разнесенного TCP/IP трафика / В.И. Ефимов, Р.Т. Файзуллин // Вестник Томского государственного университета. – 2005. – № 14. – С. 115–118.
3. Схема разделения секрета на основе метрических характеристик данных для защищенной передачи видеопотоков / А.А. Свенч, Р.Т. Файзуллин // Компьютерная оптика. – 2007. – Т. 31, вып. 3. – С. 70–72.
4. Щерба Е.В. Метод защиты канала передачи видеoinформации на основе мультиплексирования трафика // Вопросы защиты информации. – 2008. – № 1(80). – С. 55–60.
5. Имитационное моделирование системы передачи сообщений, использующей варьирование маршрутов / П.А. Мельниченко, А.А. Шелупанов // Доклады ТУСУРа. – 2008. – Т. 1, № 1. – С. 118–120.
6. Никонов В.И. Методы защиты информации в распределенных компьютерных сетях с помощью алгоритмов маршрутизации // Доклады ТУСУРа. – 2010. – № 1 (21), ч. 2. – С. 219–224.

Молчанов Алексей Леонидович

Магистрант Омского государственного технического университета (ОмГТУ)

Тел.: 913-612-38-62

Эл. почта: rfo31@inbox.ru

Файзуллин Рашид Тагирович

Д-р техн. наук, профессор ОмГТУ

Тел.: (383-2) 67-02-87

Эл. почта: r_fayzullin@omgtu.ru

Molchanov A.L., Faizullin R.T.

Studying of properties of efficiency and robust works of multiplex system

The article is devoted to studying of properties of efficiency and a robust of multiplex systems. The analysis of properties of efficiency and a robust is reached at the expense of various and repeated experiments on video file transfer.

Keywords: efficiency, robust, multiplex system.

УДК 681.3.06:004.89

Ле Хоай, А.Ф. Тузовский

Поиск в семантических электронных библиотеках

Рассматриваются четыре вида поиска на основе использования семантических технологий в электронных библиотеках, анализируются возможные варианты их применения, объясняются решаемые с их помощью задачи и предлагаются новые методы реализации.

Ключевые слова: семантические технологии, семантическая электронная библиотека, семантический поиск, контекстные метаданные, контентные метаданные, поиск по графу.

Поиск информации представляет собой важную функциональность, которую предоставляют почти все информационные системы, и в особенности электронные библиотеки (ЭБ), автоматизирующие работу пользователей с электронными ресурсами, такими как документы, изображения, аудио- и видеофайлы.

В электронных библиотеках поиск информации считается одной из основных задач, решение которой позволяет пользователям находить требуемые им информационные ресурсы. В них предоставляется возможность выполнять полнотекстовый поиск (поиск терминов) для текстовых электронных документов или поиск в метаданных для других типов ресурсов.

Проблемой полнотекстового поиска является отсутствие понимания смысла информационных потребностей пользователей, смысла искомых текстовых документов. Выполнение полнотекстового поиска затрудняется в связи с такими проблемами, как синонимия, полисемия и омонимия. Это повлечет за собой снижение релевантности получаемых результатов поиска. Для их решения требуется обрабатывать документы и формализовать содержащуюся в них семантику с использованием различных моделей представления знаний (таких, как, например, таксономии, тезаурусы и онтологии). В настоящее время уже разработано много методов решения данной задачи, например на основе определения встречаемости простых наборов ключевых слов [1, 2] либо встречаемости ключевых слов из словарей предметных областей [3–5].

Для выполнения эффективной формализации семантики информационных ресурсов организация W3C предложила набор специальных семантических технологий (набор языков и стандартов). Например, RDF (Resource Description Framework) – для описания семантических данных (метаданных); RDFS, OWL (Ontology Web Language) – для описания семантических моделей; SPARQL (Simple Protocol And RDF Query Language) – для описания запросов к таким данным [6]. Целью создания этих стандартов является предоставление средств для явного и формального описания семантики ресурсов в новой концепции Semantic Web.

Семантические электронные библиотеки (СЭБ), т.е. ЭБ, основанные на использовании семантических технологий, позволяют реализовать работу с электронными документами на новом уровне, в особенности при выполнении поиска, учитывающего семантику [7].

В данной статье рассматриваются четыре различных вида поиска ресурсов, такие как простой поиск, поиск по графу, семантический контекстный поиск и семантический контентный поиск, а также описан способ их применения в разрабатываемой авторами семантической электронной библиотеке. Для пояснения предложенных методов вначале необходимо пояснить используемый онтологический подход к описанию ресурсов и способы оценки их семантической близости.

Онтологический подход к созданию метаописания документов. Идея данного подхода заключается в использовании онтологий предметных областей для аннотирования содержания электронных документов [7, 8]. В СЭБ аннотирование электронных документов, содержащих знания из разных предметных областей, выполняется с использованием специально разработанных онтологий этих предметных областей, описанных с использованием таких языков, как RDFS или OWL. Под аннотированием ресурса понимается его семантическое метаописание в виде набора простых высказываний (триплетов), в состав которых могут входить контекстные и контентные семантические метаданные. Под простым высказыванием понимается структура вида $\langle s, p, o \rangle$, где s – субъект высказывания, p – предикат высказывания, а o – это объект высказывания.

Под контекстными метаданными аннотации объекта (документа, пользователя, проекты и т.п.) понимаются утверждения о его связи с другими объектами, понятиями из общих онтологий библиотеки, таких как онтологии пользователей или онтологии ресурсов.

Под контентными метаданными понимаются утверждения о знаниях, содержащихся в самом аннотируемом ресурсе. Наборы триплетов, составляющие контентные метаданные, создаются на основе онтологий, описывающих те предметные области, с которыми связан данный ресурс.

Определение 1. Под онтологической моделью (онтологией) O понимается знаковая система $\langle C, P, I, L, T \rangle$, где C – множество элементов, которые называются понятиями; P – множество элементов, называемых свойствами (двуместными предикатами); I – множество экземпляров понятий; L – множество текстовых меток или значений понятий и свойств; T – частичный порядок на множестве C и P .

Любому свойству $p \in P$ может быть задано значение семантического веса $pv \in [0, 1]$, которое задает смысловую близость между субъектом и объектом утверждения, составленного с использованием данного свойства. При нулевом значении pv -субъект и объект простого высказывания считаются не связанными между собой по семантике, а если значение pv равно единице, то считается, что субъект и объект имеют одинаковую семантику.

На основе онтологической модели можно определить контекстные и контентные метаданные ресурсов.

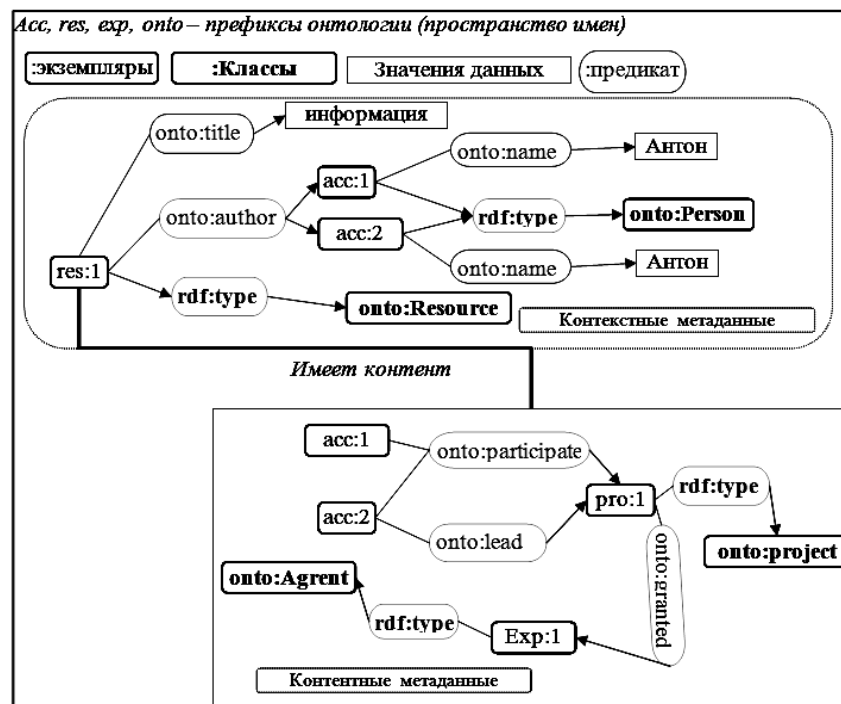


Рис. 1. Пример метаданных

Определение 2. Контекстные метаданные ресурса s – это набор простых утверждений (триплетов) $M_k = \{t_1, t_2, \dots, t_n\}$, описывающих его связи с другими объектами или которые связаны с объектом, но не содержатся в его описании. Каждый триплет t_j имеет вид $t_j = (s, p, o)$, где s – это аннотируемый ресурс (субъект), которому задан идентификатор (URI) ($s \in I$), o – объект утверждения ($o \in C \cup I \cup L$), p – отношение между субъектом и объектом ($p \in P$).

Определение 3. Контентные метаданные ресурса – это набор простых утверждений (кортежей) $M_c = \{t_1, t_2, \dots, t_m\}$, описывающих основные знания, содержащиеся в документе. Каждый кортеж t_i имеет вид $t_i = (s, p, o, v)$, где s – это субъект утверждения ($s \in C \cup I$), o – объект утверждения ($o \in C \cup I$), p – отношение между субъектом и объектом ($p \in P$), v – весовой коэффициент, который оценивает значимость данного утверждения.

Простые утверждения, созданные на основе онтологии O , могут иметь следующий вид: $\langle C, P, C \rangle$, $\langle I, P, I \rangle$, $\langle I, P, C \rangle$, $\langle I, P, L \rangle$, при этом C, P, I представляются в виде уникальных URI-идентификаторов понятий, предикатов и экземпляров.

Семантические метаданные также могут быть графически представлены в виде направленного ациклического графа (directed acyclic graph, DAG) [2], как показано на рис. 1, где вершинами графа являются субъекты и объекты триплетов, а ребрами – их предикаты (отношения).

Семантическая близость. Под семантической близостью между элементами понимается их смысловое сходство. Семантическая близость может определяться между разными компонентами триплетов. При этом в качестве базовой близости можно рассматривать близость между элементами онтологий (классами, предикатами, терминами). Классы и предикаты являются элементами схемы онтологий, а термины – элементами онтологии предметных знаний в виде таксономии или тезауруса, представленных с использованием стандарта SKOS [13] (Simple Knowledge Organization System).

Например, семантическая близость может быть оценена для классов, показанных на рис. 2, таких как *человек*, *сотрудник*, *студент* и *аспирант*, или для терминов – *animal*, *fish*, *mammal* и *puma*.

На основе оценки семантической близости между элементами онтологий может быть определена и близость между метаданными ресурсов ЭБ.

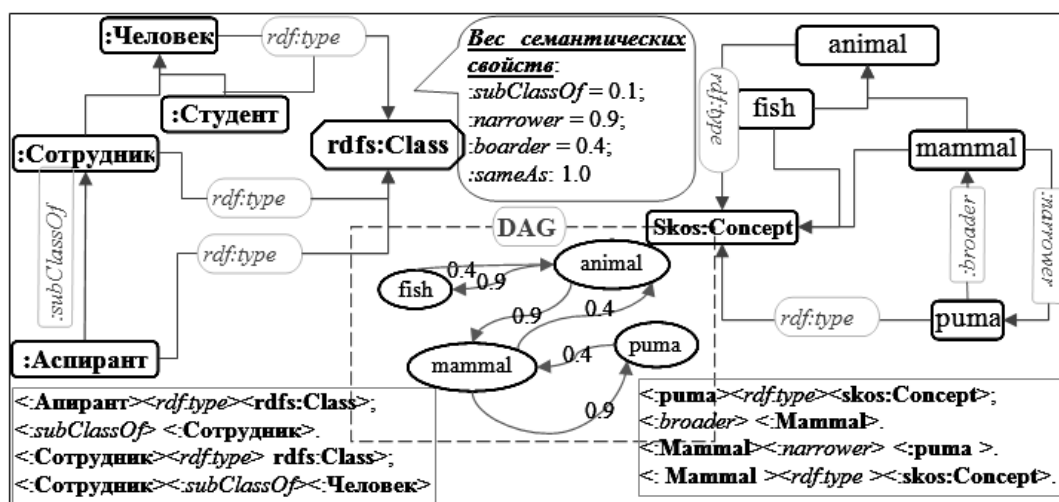


Рис. 2. Примеры оценки семантической близости

Следует отметить, что с учетом направления семантических отношений (от субъекта к объекту) *семантическая близость не является симметричной*. Допустим, необходимо вычислить близость между терминами *<mammal>* и *<puma>* с использованием отношения (*<mammal>* *<:narrower>* *<puma>*), как показано на рис. 2 и обозначено $Sim(mammal, puma)$. Если под $Sim(mammal, puma)$ понимается результат запроса на поиск всех *<mammal>*, то ответом могут быть *<mammal>* и *<puma>*, а если искать $Sim(puma, mammal)$, то первым ответом не может быть *mammal*. Это означает, что $Sim(mammal, puma)$ не равно $Sim(puma, mammal)$. На рис. 2 показан пример того, что $Sim(mammal, puma)$ равно значению семантического веса предиката «:narrower» (0,9), а $Sim(mammal, puma)$ – «:broader» (0,4).

Предикаты (свойства) онтологий будем называть семантическими, если им будут заданы весовые коэффициенты $pv, \in (0,1]$. Например: предикаты *<:broader>*, *<:subClassOf>* являются семантическими (т.к. им заданы веса 0,9 и 0,4 соответственно), а предикат *<:rdf:type>* не является семантическим, т.к. ему вес не задается. Задание веса предикатов возлагается на создателей онтологий в соответствии с их пониманием онтологии и потребностями решаемых задач эмпирическим способом [3; 9. С. 85–87; 11]. При этом не следует путать весовые коэффициенты предикатов с весовыми коэффициентами триплетов в *определении 3*.

Для вычисления семантических близостей необходимо использовать GO-граф. Он строится в соответствии со следующими правилами:

- Используются только триплеты с семантическими предикатами ($pv \neq 0$).
- Вершинами графа являются субъекты и объекты триплетов, а ребра графа от субъекта к объекту триплета имеют вес, равный значению семантического веса (pv) предиката данного триплета.
- Симметричное семантическое отношение добавляет в граф два ребра с равными весами, например: *<owl:sameAs>* добавляет два ребра со значениями $pv = 1,0$.

Под путем $PATH(a, b)$ между двумя вершинами a и b графа GO понимается набором ребер (предикатов), ведущих от вершины a до вершины b с учетом их направленности.

Пусть a и b – это элементы некоторой онтологии O . Тогда значение семантической близости $Sim(a, b)$ между этими вершинами вычисляется следующим образом:

$$Sim(a, b) = \max_{i=1 \rightarrow k} (Sim_{PATH_i}(a, b)), \quad (1)$$

где k – число возможных путей графа GO от вершины a до вершины b . Значение семантической близости между a и b по направлению пути i $sim_{PATH_i}(a, b)$ определяется по следующей формуле:

$$Sim_{PATH_i}(a, b) = \prod_{j=1}^{h_i} pv_{i,j}, \quad (2)$$

где h_i – число возможных семантических отношений между a и b на пути i ; $pv_{i,j}$ – значение веса ребра на основе j -го семантического предиката на пути i . Из (1) и (2) можно получить окончательную формулу для определения семантической близости между вершинами a и b :

$$Sim(a, b) = \max_{i=1 \rightarrow k} (Sim_{PATH_i}(a, b)) = \max_{i=1 \rightarrow k} \left(\prod_{j=1}^{h_i} pv_{i,j} \right). \quad (3)$$

$Sim(a, b)$ удовлетворяет следующим свойствам: $Sim(a, b) \in [0, 1]$; равно 0 при отсутствии пути от a к b ; $Sim(a, a) = 1$.

В качестве примера можно рассмотреть вычисление семантической близости между терминами $\langle puma \rangle$ $\langle animal \rangle$ с использованием графа DAG , показанного на рис. 2. Между ними имеются две дуги ($\langle puma \rangle$ $\langle :broader - 0,4 \rangle$ $\langle mammal \rangle$) и ($\langle mammal \rangle$ $\langle :broader - 0,4 \rangle$ $\langle animal \rangle$). С использованием этих дуг можно построить два пути $Sim(puma, animal) = 0,4 \times 0,4 = 0,16$ и $Sim(animal, puma) = 0,9 \times 0,9 = 0,81$. Таким образом, окончательная оценка семантической близости равна максимальной из них – 0,81.

На основе оценки семантической близости между базовыми элементами можно уже определять семантическую близость между триплетами, экземплярами и между наборами триплетов. Такие оценки будут рассматриваться, использоваться при описании разных видов семантических поисков. Другие методы определения различных видов семантической близости описаны в [9].

Виды поиска в семантической ЭБ. В СЭБ все информационные объекты (ИО – электронные ресурсы, пользователи, категории и др.) описываются метаданными, представленными с помощью *RDF-триплетов* на основе использования элементов некоторых онтологий. Онтологии с их экземплярами формируют базу знаний (*KB – Knowledge Base*) СЭБ. В *KB* могут выполняться запросы, описанные с помощью языка *SPARQL* [12].

Простой поиск. Поиск информационных объектов на основе использования их лексических меток.

Пусть каждый ИО s описывается в базе знаний *KB* метаданными M_L , состоящими из m триплетов, которые определяют его интерпретацию на естественных языках, и обозначается на основе онтологий O :

$$M_L(s) = (p_1(s, l_1) \vee p_2(s, l_2) \vee \dots \vee p_m(s, l_m)),$$

где каждое утверждение $p_j(s, l_j)$ (соответствующее *RDF-триплету* (s, p_j, l_j)) состоит из субъекта $s \in C \cup I$ отношения $p_j \in R$, для которого определяются метаданные и значения $l_j \in L$.

Постановка задачи: Требуется найти все ИО в *KB* с заданными текстовыми метками $l \in L$.

Решение: Результатом данного запроса должен быть любой ИО с его набором M_L , в котором содержится хотя бы одно значение l_j из лексических меток, удовлетворяющее следующее условие:

$$\text{Подстрока}(l_j) = l \quad (4)$$

Решение данной задачи заключается в формировании в *KB SPARQL*-запроса, который содержит только один известный параметр l .

Допустим, в *KB* на рис. 1 для интерпретации документа используется предикат $\langle :title \rangle$, а автора – $\langle :name \rangle$. Кроме этих предикатов, могут использоваться и другие.

Решение: Данная задача может быть решена следующим образом (пояснение выполнено на основе рис. 1).

1. Создание нового предиката, для которого все свойства ($\langle :title \rangle$, $\langle :name \rangle$ и др.), выполняющие текстовое описание, являются подсвойствами нового созданного свойства (например: $\langle :hasName \rangle$) с помощью отношения (*rdfs:subProperty*);


```

/* здесь комментарий*/
/* <:hasName> – новое созданное отношение*/
/*Добавленные триплеты: */
<:title, rdfs:subProperty, :hasName>
<:name, rdfs:subProperty, :hasName>
..... /* другие добавленные триплеты при необходимости*/

```

2. Составление SPARQL-запроса на основе триплета с созданным свойством для поиска в KB с логическим выводом;

```

/* <?X> - искомый параметр*/
/*Запрос на языке SPARQL: */
SELECT * FROM
?IO :hasName ?L, /*?IO – сам искомый объект, ?L – его название*/
?IO rdf:type ?T /* ?T – его тип*/
... /* включить другие отношения */
WHERE ?L LIKE “*/” /* условие запроса для фильтрации*/

```

3. Получение набора ИО с их названиями и типами или др.

Семантика каждого информационного объекта в KB описывается явно, т.е. он имеет тип и отношения с другими объектами. В связи с этим в задаваемый SPARQL-запрос можно включать разные значимые отношения (например, тип объекта и т.п.), которые помогают пользователю уточнить требуемый ответ.

Поиск по графу. Данный вид поиска похож на простой поиск, который использует запрос SPARQL для решения задачи. Все отличие будет заключаться в том, что в запросе **идентификатор (URI)** рассматриваемого объекта будет **известен**.

Пусть для рассматриваемого объекта s в KB имеется конечный набор из h триплетов его контекстных метаданных и обозначается как: $M_k(s) = \{t_i = \langle s, p_i, o_i \rangle \vee i = 1 \div h\}$, где $s \in I$, $p_j \in R$ и $o_j \in C \cup I \cup L$.

Постановка задачи: Для некоторого объекта s , имеющего набор отношений $p_j(s, o_j)$, требуется найти все объекты o_j , которые связаны с s отношением p_k и обладают заданными свойствами p_n . Поиск по графу может повторяться для задания $s = o_j$.

Например: нужно **найти всех друзей Бориса, живущих в городе Томске и обучающихся в ТПУ**. Здесь *Борис* – это объект s , *Друзья* – это объекты o на основании отношения p_k , а *Живут и Обучаются* – свойства (отношения) p_n .

Решение задачи: Для выбранного пользователем отношения p_k и для заданных отношений (свойств) p_n формируется соответствующий им запрос в KB на основе языка SPARQL. При выборе пользователя отправляется его соответствующий запрос в KB и полученные результаты показываются пользователям.

Например, на рис. 1 для объекта проекта $\langle :pro1 \rangle$ система может предложить для поиска отношение $\langle onto:participate \rangle$ – *участвующие* и свойства $\langle onto:profession \rangle$ для выполнения такого запроса, как найти всех участников $\langle onto:participate \rangle$ проекта $\langle :pro1 \rangle$, имеющих профессию *Программист*. Имеется следующий шаблон запроса на языке SPARQL:

```

/*Запрос на языке SPARQL: */
SELECT * FROM
S :pk ?O, /*?O – сам искомый объект, :pk – выбранное
отношение для S*/
?O :p1 ?Oi, /* ?T – его тип*/
..., /* включить другие отношения */
?O :pn ?Oi+1
WHERE /* условие запроса для фильтрация*/

```

Простой поиск и поиск по графу не учитывают семантические близости, в связи с этим они не обеспечивают ранжирование результатов. Ранжирование может быть осуществлено при контекстном и контентном поиске.

Контекстный поиск. Контекстный поиск объектов выполняется путем выполнения поискового запроса к контекстным метаданным. Примером такого поиска может служить поиск проектов, документов и других информационных объектов, которые включают заданный набор отношений.

Постановка задачи: требуется найти объекты (проект, документ, и др.) с заданным шаблонным запросом $Q_{context} = \{t_i = \langle s_q, p_i, o_i \rangle \vee i \in [1, k]\}$ и проранжировать полученные результаты по убыванию семантической близости.

Решение: Пусть в KB имеется конечный набор M_{PR} возможных объектов, которые удовлетворяют заданному запросу $Q_{context}$, тогда результатом будет любой объект $s_r \in M_R \subseteq M_{PR}$ с набором триплетов его контекста $R_{context} = \{t_j = \langle s_r, p_j, o_j \rangle \vee j \in [1, h]\}$, если значение семантической близости $Sim_{context}$ между $Q_{context}$ и $R_{context}$ будет удовлетворять следующему условию:

$$Sim_{context}(Q_{context}, R_{context}) > \varepsilon, \quad (5)$$

где ε – заданное пороговое значение. В контекстном поиске s_q и s_r не участвуют в формуле (5), так как s_q – искомая неизвестная переменная, а используются только компоненты p и o их триплетов $Q_{context}$ и $R_{context}$. Для вычисления выражения (5) на основе семантической близости между экземплярами может быть использована следующая формула:

$$Sim_{context}(Q_{context}, R_{context}) = \frac{\sum_{i=1}^k \max_{j=1 \rightarrow h} (Sim(t_i, t_j))}{k} > \varepsilon, \quad (6)$$

где $Sim(t_i, t_j)$ определяется следующим образом:

$$Sim(t_i, t_j) = Sim(p_i, p_j) \times Sim(o_i, o_j), \quad (7)$$

где $Sim(p_i, p_j)$ и $Sim(o_i, o_j)$ – семантические близости между предикатами и объектами триплета t_i и t_j , соответственно, которые вычисляются по формуле (3), $Sim_{context}$ – значение семантической близости между экземплярами (контекстами).

Результаты поиска упорядочиваются по убыванию значений семантических близостей, вычисленных по формуле (6).

Контентный поиск. Контентный поиск применяется для объектов, проаннотированных наборами триплетов контентных метаданных.

Постановка задачи: Найти информационные объекты, соответствующие шаблонному запросу $Q_{content} = \{t_i = \langle s_i, p_i, o_i \rangle \vee i \in [1, k]\}$, и проранжировать полученные результаты по убыванию семантической близости.

Решение: Пусть в KB имеется конечный набор N_{PR} возможных объектов, которые удовлетворяют заданному запросу $Q_{content}$ и тогда результатом будет любой объект $s_r \in N_R \subseteq N_{PR}$ с набором триплетов его контента $R_{content} = \{t_j = \langle s_j, p_j, o_j \rangle \vee j \in [1, h]\}$ при условии того, что значение семантической близости $Sim_{content}$ между $Q_{content}$ и $R_{content}$ удовлетворяет следующему условию:

$$Sim_{content}(Q_{content}, R_{content}) > \varepsilon, \quad (8)$$

где ε – заданное пороговое значение. Разложение левой части формулы (8) похоже на формулу (6) и отличается только тем, что в контентном поиске в вычислении семантической близости участвуют все компоненты триплетов из контентных метаданных. Для вычисления семантической близости между триплетами t_i, t_j может быть использована следующая формула:

$$Sim(t_i, t_j) = \begin{cases} |k| \frac{Sim(s_1, s_2) + Sim(o_1, o_2)}{2}, & \forall k > 0, \\ |k| \frac{Sim(s_1, o_2) + Sim(o_1, s_2)}{2}, & \forall k \leq 0, \end{cases} \quad (9)$$

где $k = Sim(p_1, p_2)$ и $Sim(t_1, t_2) \in [0, 1]$, $Sim(t_1, t_2) = 0$ при $k = 0$. При $k < 0$ два предиката p_1 и p_2 имеют отношение $\langle p_1, owl:inverseOf, p_2 \rangle$, при наличии триплета $\langle s, p_1, o \rangle$ подразумевается триплет $\langle o, p_2, s \rangle$. Семантическая близость между компонентами триплета t_i и t_j вычисляется по формуле (3), а весовые коэффициенты всех утверждений контентных метаданных одинаковы ($v = 1$).

Результаты поиска ранжируются по убыванию семантических близостей, вычисленных по формуле (8).

На практике наборы M_{PR} и N_{PR} формируются путем фильтрации объектов данного типа с помощью языка SPARQL [10] с целью ограничения количества объектов, для которых вычисляется семантическая близость.

Реализация видов поиска в СЭБ. Все рассмотренные виды поиска являются полезными в семантических ЭБ для повышения точности поиска и использования возможностей семантических технологий.

Примеры реализации интерфейса простого вида поиска, позволяющего пользователям составлять набор триплетов для поиска по контенту и контексту, показан на рис. 3 и 4.

Все показанные примеры интерфейсов реализованы в разрабатываемой авторами семантической электронной библиотеке *SemDL*. Данная библиотека предоставляет возможность составлять триплеты запроса для контентных и контекстных метаданных базы знаний системы.

Для контекстного поиска вначале необходимо выбрать тип объекта, а затем можно составлять другие триплеты о контексте данного объекта. Возможные свойства (предикаты) будут определяться автоматически для выбранного типа объекта и предлагаться пользователю.

Результаты поиска ранжируются по убыванию значения семантической близости между набором триплетов искомого объекта и набором триплетов релевантных объектов.

Проводилась оценка точности работы предлагаемых видов контентного и контекстного поиска с использованием экспертов [3, 9]. В результате были получены достаточно высокие оценки точности и полноты, примерно 85%. В некоторых случаях точность достигает 100% при учете *строгой несимметричности* семантических близостей между базовыми элементами.

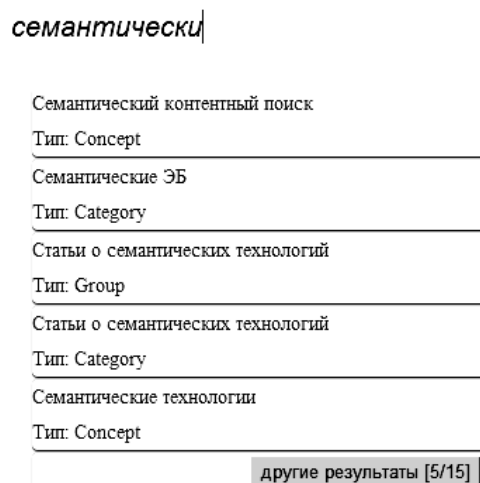


Рис. 3. Пример интерфейса простого поиска

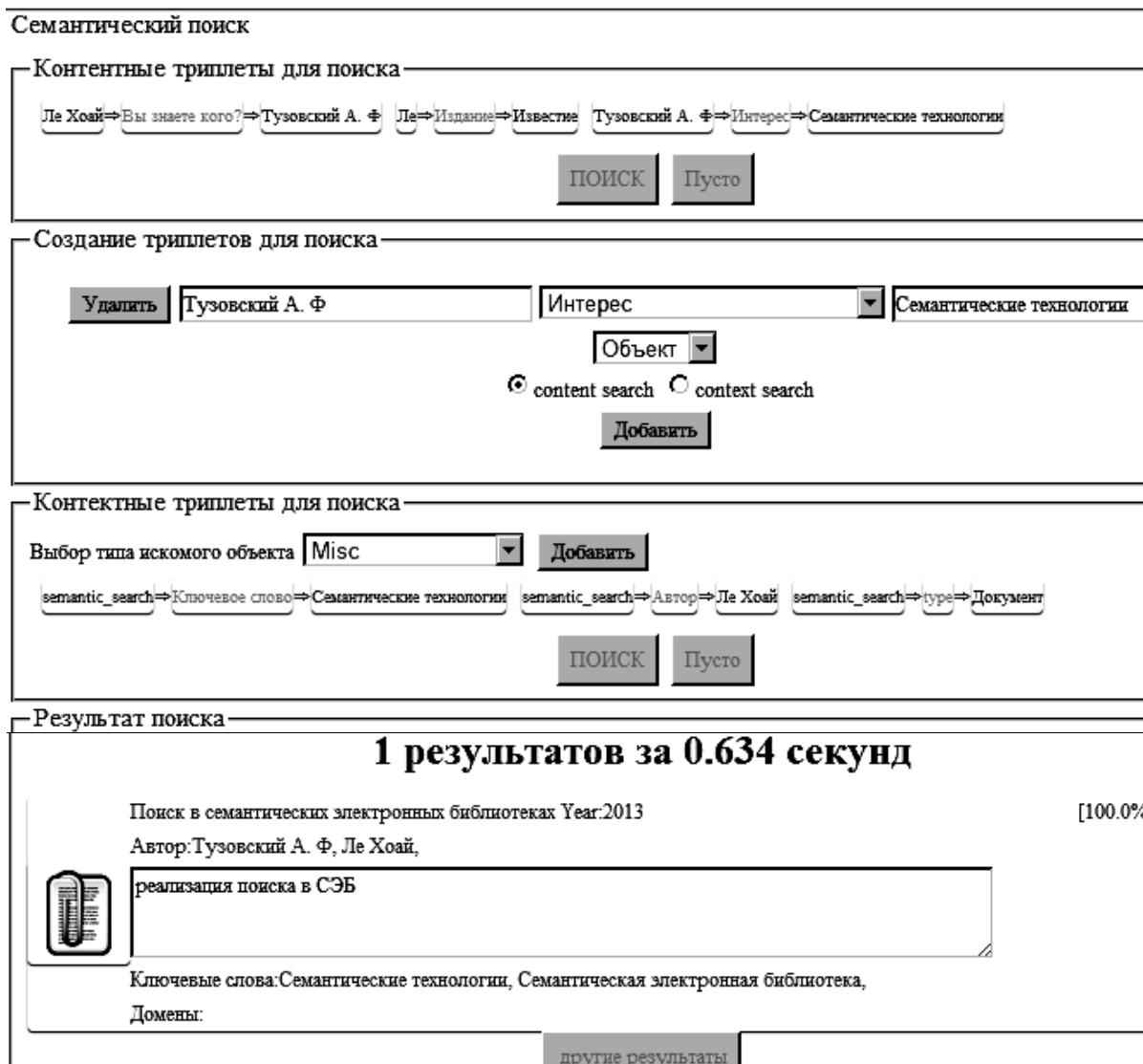


Рис. 4. Пример интерфейса для выполнения контекстного и контентного поиска и результата поиска по контексту

Заключение. Поиски по контенту и контексту позволяют ранжировать результаты поиска на основе их значений семантической близости. Контентный поиск применяется для поиска по содержанию объектов, контекстный поиск используется для поиска всех объектов системы на основе их контекстных метаданных. В настоящее время широкое применение получили простой поиск и поиск по графу, по сравнению с остальными видами, так как они не требуют вычислений семантических близостей. Однако использование контекстного и контентного поиска позволяет выполнять более сложные запросы и получать более точные результаты.

Литература

1. Нгуен Б.Н. Классификация текстов на основе оценки семантической близости терминов / Б.Н. Нгуен, А.Ф. Тузовский // Изв. Том. политехн. ун-та. – 2012. – Т. 320, № 5. – С. 43–48.
2. Taylor William P. A comparative study on ontology generation and text clustering using VSM, LSI, and document ontology models. – Clemson University, 2007. – 58 p.
3. Saša Nešić. Semantic Document Architecture for Desktop Data Integration and Management: Doctoral Dissertation. – Italia, 2010. – 228 p.
4. Hotho A. Wordnet improves Text Document Clustering / A. Hotho, S. Staab, and G. Stumme // Proceedings of the SIGIR 2003 Semantic Web Workshop. – 2003. – P. 541–544.
5. Using WordNet Similarity and Antonymy Relations to Aid Document Retrieval [Электронный ресурс]. – Режим доступа: <http://www-users.cs.york.ac.uk/kazakov/papers/desimone-kazakov-crc.pdf> (дата обращения: 01.03.2013).
6. Hendler A. J. Handbook of Semantic Web Technologies. – Springer, 2011. – 1077 p.
7. Ле Х.Х. Разработка электронных библиотек на основе семантических технологий // Научно-технический вестник Поволжья (Казань). – 2012. – № 3. – С. 138–145.
8. Тузовский А.Ф. Формирование семантических метаданных для объектов системы управления знаниями // Изв. Том. политехн. ун-та. – 2007. – Т. 310, № 3. – С. 108–112.
9. Нгуен Б.Н. Модели и методы поиска информационных ресурсов с использованием семантических технологий: дис. ... канд. техн. наук. – Томск, 2012. – 198 с.
10. Ле Х. Формирование рекомендаций в семантических электронных библиотеках / Хоай Ле, А.Ф. Тузовский // Проблемы информатики. – 2012. – № 3. – С. 113–119.
11. Gong Z. Multi-term Web Query Expansion Using WordNet / Z. Gong, C.W. Cheang, U.H. Leong // DEXA'06 Proceedings of the 17th Database and Expert Systems Applications Conference. – Springer-Verlag Berlin, Heidelberg, 2006. – P. 379–388.
12. SPARQL 1.1 overview [Электронный ресурс]. – Режим доступа: <http://www.w3.org/TR/2012/WD-sparql11-overview-20120501/>, свободный (дата обращения: 01.03.2013).
13. SKOS Simple Knowledge Organization System RDF Schema [Электронный ресурс]. – Режим доступа: <http://www.w3.org/TR/2008/WD-skos-reference-20080829/skos.html> (дата обращения: 01.03.2013).

Ле Хоай

Аспирант каф. оптимизации систем управления НИ ТПУ
Тел.: +7-913-108-01-44
Эл. почта: lehotomsk@yahoo.com

Тузовский Анатолий Федорович

Д-р техн. наук, профессор каф. оптимизации систем управления НИ ТПУ
Тел.: +7 (382-2) 42-14-85
Эл. почта: tuzovskyaf@tpu.ru

Hoai Le, Tuzovsky A.F.

Search in semantic digital libraries

In this paper four types of search in digital libraries on the basis of using semantic technologies are considered, possible options of their application are analyzed, tasks solved with their help are explained and new methods of implementation are proposed.

Keywords: semantic technology, semantic digital library, semantic search, contextual metadata, content metadata, graph search.

УДК 510.644

А.А. Ефремов

Новые операции над нечеткими числами и интервалами

Предложены новые операции над нечеткими числами и интервалами, сохраняющие тип и функциональную форму функций принадлежности и позволяющие образовывать составные термы лингвистических переменных, соответствующие применению лингвистических модификаторов «очень» и «более или менее». Получены выражения, позволяющие для отдельных типов функций принадлежности найти характеристические точки нечетких результатов введенных операций без вычисления интегралов.

Ключевые слова: нечеткое множество, нечеткое число, лингвистические модификаторы, полиномиальные функции принадлежности, кусочно-непрерывные функции.

В работе [1] Заде вводит операции концентрирования и растяжения нечеткого множества \tilde{A} как средство порождения составных термов соответствующей лингвистической переменной с помощью модификаторов «очень» и «более или менее», соответственно. Если $\mu_A(x)$ – функция принадлежности (ФП) нечеткого множества \tilde{A} , то операция концентрирования определяется как $CON(\tilde{A}) = (\mu_A(x))^2$, а растяжения – $DIL(\tilde{A}) = \sqrt{\mu_A(x)}$.

Критики в работах [2, 3] указывают на то, что подобные операции не изменяют носитель и ядро нечеткого множества, т.к. полученные ФП принимают значения 0 и 1 в тех же точках, что и исходная, и предлагают производить сдвиг ФП после операций концентрирования и растяжения.

В настоящей работе вводятся новые операции над нечеткими множествами, заданными на множестве действительных чисел \mathfrak{R} (нечеткие числа и интервалы), позволяющие модифицировать нечеткую величину, не изменяя функциональную форму ФП. Также, показывается простота вычисления результатов введенных операций для некоторых типов ФП.

Операции сжатия и размытия нечетких величин. Пусть \tilde{A} – нечеткое множество, заданное на универсальном множестве U , с функцией принадлежности $\mu_A(x)$, и C_A – результат дефаззификации данного множества по методу центра максимума. Иными словами, C_A – центр ядра нечеткого множества \tilde{A} . Также, пусть S_A – площадь под функцией принадлежности $S_A = \int_{\text{supp } \tilde{A}} \mu_A(x) dx$, где $\text{supp } \tilde{A}$ – носитель нечеткого множества \tilde{A} .

Определение 1. Результатом операции *сжатия* множества \tilde{A} со степенью $r \geq 1$ является нечеткое множество \tilde{B} с функцией принадлежности $\mu_B(x)$ такого же вида, что и $\mu_A(x)$, с центром ядра $C_B = C_A$, площадь под функцией принадлежности которого $S_B = \frac{S_A}{r}$, где $S_B = \int_{\text{supp } \tilde{B}} \mu_B(x) dx$.

Операцию сжатия со степенью r нечеткого множества \tilde{A} будем обозначать $\tilde{B} = \text{sharp}_r \tilde{A}$.

Определение 2. Результатом операции *размытия* множества \tilde{A} со степенью $r \geq 1$ является нечеткое множество \tilde{B} с функцией принадлежности $\mu_B(x)$ такого же вида, что и $\mu_A(x)$, с центром ядра $C_B = C_A$, площадь под функцией принадлежности которого $S_B = r \cdot S_A$, где $S_B = \int_{\text{supp } \tilde{B}} \mu_B(x) dx$.

Будем обозначать операцию размытия со степенью r $\tilde{B} = \text{blurr}_r \tilde{A}$.

Очевидно, введенные операции при $r > 0$ являются коммутативными, так что

$$\text{blurr}_r \tilde{A} = \text{sharp}_{(r^{-1})} \tilde{A}; \quad \text{sharp}_r \tilde{A} = \text{blurr}_{(r^{-1})} \tilde{A}.$$

Исходя из введенных определений, можно получить следующие выражения:

$$\begin{aligned} \text{blurr}_1 \tilde{A} = \text{sharp}_1 \tilde{A} = \tilde{A}; \quad \text{blurr}_0 \tilde{A} = \text{sharp}_\infty \tilde{A} = C_A; \\ \text{blurr}_\infty \tilde{A} = \text{sharp}_0 \tilde{A} = U. \end{aligned}$$

В дальнейшем будем говорить о нечетких множествах, заданных на множестве действительных чисел \mathfrak{R} , т.е. о нечетких числах и интервалах.

Пример с трапецидальными ФП. Рассмотрим нечеткий интервал $\tilde{A} = \langle L, K_L, K_R, R \rangle$ с ФП

$$\mu_A(x) = \begin{cases} \frac{x-L}{K_L-L}, & x \in [L, K_L); \\ 1, & x \in [K_L, K_R]; \\ \frac{x-R}{K_R-R}, & x \in (K_R, R]; \\ 0, & x \notin [L, R]; \end{cases} \quad (1)$$

где $[L, R]$ — носитель нечеткого множества, $[K_L, K_R]$ — его ядро, и $\frac{K_L + K_R}{2} = C_A$ — центр ядра.

Функция принадлежности (1) нечеткого множества \tilde{A} изображена на рис. 1, а, б сплошными линиями и при $x \in [L, R]$ представляет собой трапецию с высотой $h=1$ и основаниями $a=R-L$ и $b=K_R-K_L$, площадь которой определяется по известной формуле [4]:

$$S_A = \frac{a+b}{2} \cdot h. \quad (2)$$

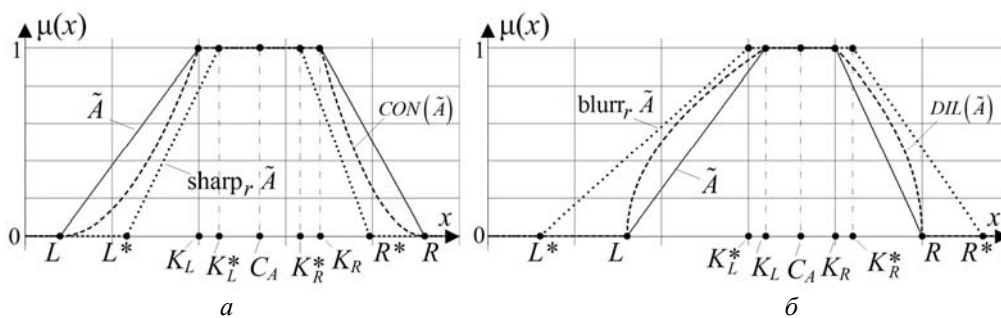


Рис. 1. Сравнение ФП результатов операций над \tilde{A} : а — концентрирование (CON) и сжатие ($sharp$); б — растяжение (DIL) и размытие ($blurr$)

ФП, полученные в результате использования операций $CON(\tilde{A})$ и $DIL(\tilde{A})$, изображены на рис. 1, а, б штриховыми линиями. Из рис. 1 видно, что левые и правые части ФП исказились, в то время как носитель и ядро новых нечетких интервалов совпадают с исходными.

Нечеткий интервал $\tilde{B} = \text{sharp}_r \tilde{A}$ представляет собой нечеткое множество $\tilde{B} = \langle L^*, K_L^*, K_R^*, R^* \rangle$, ФП которого также является трапецией с высотой $h=1$, площадь которой меньше площади исходной трапеции в r раз. Согласно (2) получаем

$$\frac{(R-L) + (K_R - K_L)}{2} \cdot h = r \cdot \frac{(R^* - L^*) + (K_R^* - K_L^*)}{2} \cdot h. \quad (3)$$

Так как центры меньших оснований трапеций должны совпадать, то, используя (3), получаем:

$$\begin{aligned} L^* = C_A - \frac{C_A - L}{r}; \quad R^* = C_A - \frac{C_A - R}{r}; \\ K_L^* = C_A - \frac{C_A - K_L}{r}; \quad K_R^* = C_A - \frac{C_A - K_R}{r}. \end{aligned} \quad (4)$$

Следуя аналогичным рассуждениям, можно получить зависимости для характерных точек нечеткого интервала $\tilde{B} = \text{blurr}_r \tilde{A}$:

$$\begin{aligned} L^* = C_A - (C_A - L) \cdot r; \quad R^* = C_A - (C_A - R) \cdot r; \\ K_L^* = C_A - (C_A - K_L) \cdot r; \quad K_R^* = C_A - (C_A - K_R) \cdot r. \end{aligned} \quad (5)$$

Примеры ФП нечетких интервалов, полученных в результате применения операций сжатия и размытия, приведены на рис. 1 (пунктирные линии).

Соотношения (4) и (5) можно использовать и в том случае, когда нечеткое множество \tilde{A} представляет собой треугольное нечеткое число $\tilde{A} = \langle L, C, R \rangle$, принимая во внимание, что $C = C_A = K_L = K_R$.

Пример с кусочно-непрерывными полиномиальными ФП. В работе [5] введен класс нечетких чисел (L-R)-типа с кусочно-непрерывными ФП такими, что функции $f_L(x)$ и $f_R(x)$ левой и правой частей ФП являются полиномами второго порядка с дополнительными условиями о равенстве нулю производных функций $f_L(x)$ и $f_R(x)$ в характеристических точках нечеткой величины. Тогда для нечеткого интервала $\tilde{A} = \langle L, K_L, K_R, R \rangle$ общий вид ФП можно записать как

$$\mu_A(x) = f_L(x) \cdot H(x-L) \cdot H(K_L-x) + H(x-K_L) \cdot H(K_R-x) + f_R(x) \cdot H(x-K_R) \cdot H(R-x), \quad (6)$$

где $H(x)$ - единичная функция Хевисайда, а $f_L(x)$ и $f_R(x)$ удовлетворяют условию

$$\begin{cases} f_L(L) = 0; \\ f_L(K_L) = 1; \\ f_R(K_R) = 1; \\ f_R(R) = 0. \end{cases} \quad (7)$$

Форму ФП, определяемой выражением (6), можно изменять, налагая на функции $f_L(x)$ и $f_R(x)$ одно из четырех дополнительных условий:

$$\begin{cases} f'_L(L) = 0; \\ f'_R(R) = 0. \end{cases} \quad (8)$$

$$\begin{cases} f'_L(K_L) = 0; \\ f'_R(K_R) = 0. \end{cases} \quad (9)$$

$$\begin{cases} f'_L(L) = 0; \\ f'_R(K_R) = 0. \end{cases} \quad (10)$$

$$\begin{cases} f'_L(K_L) = 0; \\ f'_R(R) = 0. \end{cases} \quad (11)$$

Тип ФП, определяемой выражением (6), будем обозначать $2P$ с добавлением греческих букв $-\alpha, -\beta, -\gamma, -\delta$ при соответствии ФП одному из условий (8)–(11).

Внешний вид ФП типов $2P\alpha - 2P\delta$ представлен на рис. 2, $a - \varepsilon$ соответственно.

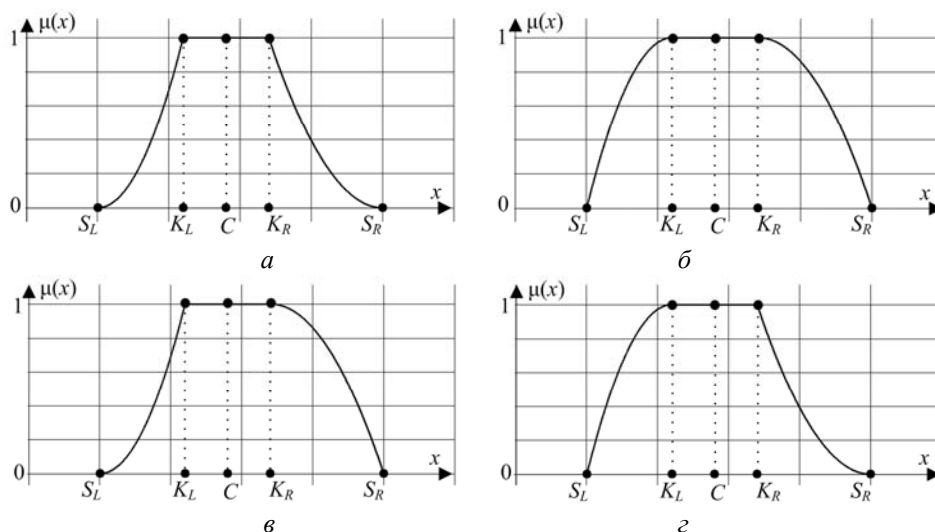


Рис. 2. Примеры кусочно-непрерывных полиномиальных ФП: a – тип $2P\alpha$; $б$ – тип $2P\beta$; $в$ – тип $2P\gamma$; $г$ – тип $2P\delta$

Коэффициенты функций $f_L(x)$ и $f_R(x)$, а, следовательно, и площадь под кривой ФП, однозначно выражаются через характерные точки нечеткого интервала путем решения системы уравнений, составленной из условия (7) и одного из условий (8)-(11), соответствующего типу ФП.

Пример. Рассмотрим операции сжатия и размытия нечеткого интервала $\tilde{A} = \langle L, K_L, K_R, R \rangle$ с ФП типа $2P\alpha$ (рис. 2, а). Составим систему уравнений, полученную из условий (7) и (8):

$$\begin{cases} a_0 + a_1L + a_2L^2 = 0; \\ a_0 + a_1K_L + a_2K_L^2 = 1; \\ b_0 + b_1K_R + b_2K_R^2 = 1; \\ b_0 + b_1R + b_2R^2 = 0; \\ a_1 + 2a_2L = 0; \\ b_1 + 2b_2R = 0; \end{cases} \quad (12)$$

где a_i и b_i — коэффициенты полиномов $f_L(x) = \sum_{i=0}^2 a_i x^i$ и $f_R(x) = \sum_{i=0}^2 b_i x^i$. Решением системы (12)

будут коэффициенты функций $f_L(x)$ и $f_R(x)$, выраженные через характерные точки нечеткого интервала:

$$\begin{aligned} a_0 &= \frac{L^2}{(L - K_L)^2}; & a_1 &= \frac{-2L}{(L - K_L)^2}; & a_2 &= \frac{1}{(L - K_L)^2}; \\ b_0 &= \frac{R^2}{(R - K_R)^2}; & b_1 &= \frac{-2R}{(R - K_R)^2}; & b_2 &= \frac{1}{(R - K_R)^2}. \end{aligned} \quad (13)$$

Тогда, площадь под кривой $\mu_A(x)$ также выражается через данные характерные точки:

$$S_A = \int_{\text{supp } \tilde{A}} \mu_A(x) dx = \int_L^{K_L} f_L(x) dx + \int_{K_L}^{K_R} dx + \int_{K_R}^R f_R(x) dx = \frac{K_L - L}{3} + (K_R - K_L) + \frac{R - K_R}{3}.$$

Согласно Определению 1, ФП нечеткого интервала $\tilde{B}_1 = \text{sharp}_r \tilde{A} = \langle L^*, K_L^*, K_R^*, R^* \rangle$ также будет являться кусочно-непрерывная функция типа $2P\alpha$, площадь под которой будет меньше S_A в r раз, причем центры ядер нечетких множеств \tilde{A} и \tilde{B}_1 должны совпадать. Исходя из этого, легко можно показать, что характерные точки нечеткого интервала \tilde{B}_1 можно также определить, используя выражения (4). Аналогично, можно определить характерные точки нечеткого интервала $\tilde{B}_2 = \text{blur}_r \tilde{A}$ с использованием выражений (5).

Примеры ФП типа $2P\alpha$ нечетких интервалов, полученных применением операций сжатия и размытия, изображены штриховыми линиями на рис. 3.

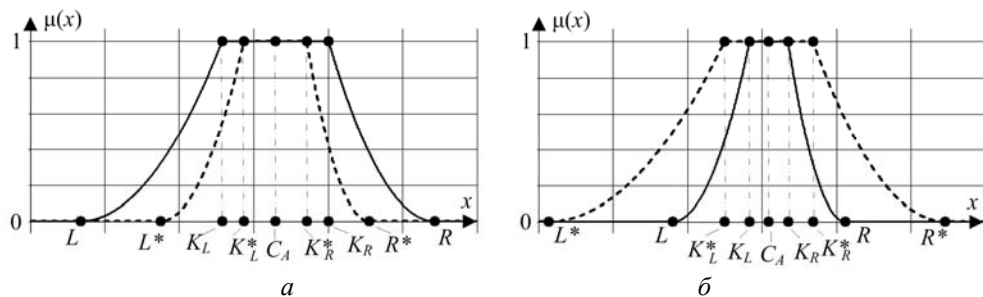


Рис. 3. Результат вычисления операций для ФП (L-R)-типа: а – сжатие; б – размытие

Легко доказать, что выражения (4) и (5) применимы для всех четырех введенных выше типов функций принадлежности. Также, при $C_A = K_L = K_R$ эти выражения позволяют получить харак-

терные точки нечетких чисел с кусочно-непрерывными ФП, левые и правые части которых являются полиномами второго порядка, удовлетворяющие условию (7) и одному из условий (8)–(11).

Заключение. Представленные в настоящей работе операции сжатия и размытия нечетких множеств позволяют образовывать составные термы лингвистических переменных без усложнения выражений для функций принадлежности. Простота вычисления результатов введенных операций для отдельных распространенных типов ФП позволяет предложить их в качестве альтернативы традиционным операциям концентрирования и растяжения.

Литература

1. Заде Л.А. Понятие лингвистической переменной и его применение к принятию приближенных решений. - М.: Мир, 1976. – 167 с.
2. Lackoff G. Hedges: A study in meaning criteria and the logic of fuzzy concepts // Journal of Philosophical Logic. – 1973. – Vol. 2, Iss. 4. – pp. 458-508.
3. Huynh V.N., Ho T.B., Nakamori Y. A parametric representation of linguistic hedges in Zadeh's fuzzy logic // International Journal of Approximate Reasoning. - 2002. – Vol. 30, Iss. 3. – P. 203–223.
4. Корн Г. А. Справочник по математике для научных работников и инженеров: определения, теоремы, формулы / Г.А. Корн, Т.М. Корн. - 6-е изд., стер. - СПб.: Лань, 2003. – 832 с.
5. Ефремов А.А. О применении кусочно-непрерывных функций к заданию функций принадлежности нечетких чисел (L-R)-типа / А.А. Ефремов, А.М. Кориков // Вестник науки Сибири. – 2011. - №. 1(1) [Электронный ресурс] - Режим доступа: <http://sjs.tpu.ru/journal/article/view/70/117>, свободный (дата обращения: 20.03.2012).

Ефремов Александр Александрович

Ассистент каф. автоматизации и компьютерных систем НИ ТПУ

Тел.: 8 (383-2) 41-89-07

Эл. почта: Yefremov@aics.ru

Yefremov A.A.

New operations on fuzzy numbers and intervals

New operations on fuzzy numbers and intervals which keep the type and the shape of membership functions intact and allow producing complex terms of linguistic variables corresponding linguistic hedges «very» and «more or less» are suggested in this paper. We derive mathematical expressions which allow determining characteristic points of fuzzy results of introduced operations for particular types of membership functions without integral evaluation.

Keywords: fuzzy set, fuzzy number, linguistic hedges, polynomial membership functions, piecewise continuous functions.

УДК 004.031

Ю.П. Ехлаков, Ю.Б. Гриценко, О.И. Жуковский

Автоматизация технологий формирования и мониторинга электронного генерального плана инженерной инфраструктуры

Рассмотрены ряд общих особенностей инженерных систем, технология использования генеральных планов на крупных промышленных предприятиях, эксплуатирующих инженерные сети, приводятся этапы создания и ведения электронного генерального плана, предлагается архитектура информационной системы управления электронным генеральным планом, в рамках которой представлены информационные подсистемы для управления инженерной инфраструктурой крупных промышленных предприятий, приведен их функциональный состав.

Ключевые слова: электронный генеральный план, инженерная инфраструктура, геоинформационные системы, хранилище данных, интеллектуальный анализ, электронный документооборот.

Эффективность автоматизации большинства технологических процессов производства во многом зависит от устойчивости функционирования его инженерной инфраструктуры. Организации, занятые в процессе эксплуатации инженерной инфраструктуры, зачастую сталкиваются с рядом довольно сложных проблем, связанных с недостоверностью и задержкой получения информации об объектах подотчетных им инженерных сетей.

Несмотря на значительные различия в типе транспортируемого продукта и технологиях, используемых эксплуатационными организациями, современные инженерные сети характеризуются рядом общих специфических особенностей [1]:

- обширным территориальным распределением и огромным числом элементов, формирующих систему;
- непрерывным развитием в пространстве и во времени;
- сложной иерархической структурой управляемой и управляющих систем и непосредственным наличием субъекта в контуре управления;
- непрерывностью во времени процессов транспортировки и распределения целевого продукта;
- высокой степенью централизации управления с одновременной децентрализацией оперативного управления технологическими процессами транспортировки и распределения целевого продукта;
- инерционностью процессов транспортировки продукта и необходимостью создания его оперативных запасов в резервуарах и хранилищах.

Процесс эффективного управления инженерными сетями тесно связан с понятием электронного генерального плана. Генеральный план промышленного предприятия – одна из важнейших частей проекта промышленного предприятия, содержащая комплексное решение вопросов планировки и благоустройства территории, размещения зданий, сооружений, транспортных коммуникаций, инженерных сетей, организации систем хозяйственного и бытового обслуживания, а также расположения предприятия в промышленном районе (узле).

Постановка задачи. Основой технических решений по генеральному плану промышленного предприятия является технология основного и вспомогательного производства, состав основных и вспомогательных цехов, а также их взаимная связь в процессе производства. Технологическая связь цехов осуществляется с помощью транспортных коммуникаций и оказывает влияние на решение по взаимному размещению цехов.

На решения, принимаемые на основе генерального плана, влияет ряд факторов: размещение предприятия в составе промышленного узла; близость и взаимное расположение его по отношению к селитебному району; необходимость создания архитектурно-выразительного облика предприятия; вертикальная планировка с учетом рельефа местности; ориентация площадки относительно сторон света; преобладающее направление ветра и ряд других факторов.

Под понятием «генеральный план промышленного предприятия» будем понимать комплексное технологическое, транспортное и архитектурно-строительное решение промышленного предпри-

ятия. В составе проекта генеральный план разрабатывается как чертеж (или комплект чертежей), на котором в определенном масштабе показано взаимное расположение всех зданий и сооружений, транспортных и энергетических коммуникаций в увязке с рельефом местности и планировочными решениями на площадке с учетом требований технологии и транспорта.

Генеральный план должен также предусматривать возможность дальнейшего развития предприятия и обеспечивать такую производственную структуру, при которой могут быть достигнуты наивысшие результаты производства при наименьших затратах; создавать условия для максимального удовлетворения интересов всех работников предприятия.

По назначению и содержанию различают проектный, строительный и исполнительный генпланы:

– генплан (проектный) разрабатывается на всех стадиях проектирования, но с различной степенью детализации. Он необходим для строительства предприятия, как основной документ, определяющий, где и какие здания и сооружения необходимо строить;

– строительный генплан разрабатывается на основе проектного генплана и служит для решения вопросов размещения временных сооружений, необходимых для нужд строительства предприятия, например: складов стройматериалов, площадок для монтажа технологического оборудования, строящихся цехов и др.;

– исполнительный генплан ведется на предприятиях. На нем фиксируется фактическое положение всех построенных на предприятиях зданий и сооружений и отмечаются возможные отступления (как правило, незначительные) от проектного генерального плана, связанные с технологией строительства, неточностью ведения строительных работ и непредвиденными обстоятельствами. Исполнительный генеральный план по своему содержанию является постоянно действующим топографическим планом предприятия, который непрерывно дополняется по мере его строительства.

По охвату территории различают общеплощадочный и объектный генеральный план:

– общеплощадочный генплан разрабатывается на всю территорию промышленной площадки предприятия. Как правило, это делается либо при строительстве, либо при полной реконструкции предприятия;

– объектный генплан составляется на участок территории, новые здания и промышленные сооружения.

На практике используются также комбинированные варианты типов генеральных планов, например: «проектный общеплощадочный генеральный план» или «исполнительный объектный генеральный план» и т.п.

Существует также понятие дежурного генерального плана (ДГП), под ним понимают актуальный, постоянно обновляемый и действующий на предприятии документ. ДГП по своему содержанию является топографическим планом предприятия, создаваемым на основе исполнительного общеплощадочного генерального плана, на который наносятся проектные и временные (строительные) объекты и коммуникации с чертежей проектных и строительных генеральных планов (как общеплощадочных, так и объектных).

Жизненный цикл объекта автоматизации – генерального плана промышленного предприятия – состоит из следующих основных этапов: создание (проектирование), функционирование (ведение, сопровождение и использование), развитие (изменение).

Между этапами жизненного цикла генерального плана предприятия и самого предприятия прослеживается тесная взаимная связь. На первом этапе проектирования предприятия создается генеральный план предприятия, на следующем этапе функционирования предприятия генеральный план используется для принятия решений на основе его данных, а на этапе развития (реконструкции) предприятия в генеральный план вносятся изменения.

Этап функционирования генерального плана по времени – это самый объемный этап. Основные задачи, которые должны решаться на этом этапе, можно объективно поделить на две группы: задачи сопровождения (ведения), задачи использования.

Сопровождением генерального плана предприятия, как правило, занимается конструкторское бюро генерального плана. На крупных промышленных предприятиях оно входит в состав общезаводского проектно-конструкторского отдела.

Основными функциями конструкторского бюро на этапе функционирования жизненного цикла генерального плана предприятия являются: внесение изменений в дежурный генплан, согласование новых проектов, формирование выкопировок из дежурного генплана по заказу.

Как правило, созданием генерального плана предприятия занимается сторонняя проектная организация (проектировщик). Результатом деятельности проектировщика является проект предприятия, неотъемлемой частью которого являются проектный и строительный общеплощадочные генеральные планы. После окончания строительства предприятия геодезистами (топографами) выполняется исполнительная съемка территории предприятия, в результате чего появляется исполнительный общеплощадочный генеральный план. Далее вся накопленная документация, в том числе генеральный план, передается в конструкторское бюро генерального плана для создания ДГП.

На этапе функционирования генерального плана при реконструкции предприятия выполняются работы по созданию проектных и строительных объектных генеральных планов, обычно эту работу выполняют проектные бюро предприятия. После выполнения проекта разработанные генеральные планы передаются в конструкторское бюро генерального плана предприятия, где наносятся на дежурный генеральный план.

Далее начинается реализация проекта, после окончания строительных работ по заказу подразделения топографами создается исполнительный объектный генеральный план, объекты которого наносятся на генеральный план.

Перед проектированием инженерная служба или проектное бюро предприятия заказывает у конструкторского бюро выкопировку требуемой территории из ДГП. Согласование с генпланом необходимо при выполнении любого проекта на территории предприятия.

Таким образом, генеральный план состоит из дежурного генплана, находящегося в ведении конструкторского бюро генплана, и архива генплана, хранящегося в проектно-конструкторском отделе.

Состояние дежурного генерального плана на большинстве отечественных предприятий является критическим. В лучшем случае на предприятии имеется исполнительный общеплощадочный генеральный план и ДГП, который по ряду объективных причин не содержит достоверной информации. Основной и самой важной причиной является нехватка квалифицированной рабочей силы.

Для решения всех существующих проблем, связанных с процессом ведения и использования генерального плана, требуется внедрение в этот процесс информационных технологий. Создание электронной версии генерального плана позволит сократить время формирования топосъемок, согласования проектов, поиска необходимой информации на генплане и др.

Основной результат. Электронный генеральный план (ЭГП) по структуре данных состоит из двух основных составляющих: графические (пространственные) данные, атрибутивные (семантические, описательные) данные. При этом каждый графический объект генплана может иметь атрибутивное описание.

Источником для формирования набора пространственных данных ЭГП является ДГП и архив генплана. Набор атрибутивных данных формируется на базе описательных данных, которыми располагают инженерные службы предприятия. Графические данные в обязательном порядке связаны с атрибутивной информацией.

Переводом генерального плана в электронный вид должны заниматься проектно-конструкторские отделы предприятия, службы главного архитектора и вновь создаваемые службы генерального плана с применением ПС, позволяющих формировать слои ЭГП в форматах файлов shp, dxf, dwg, mif, которые поддерживаются популярными инструментальными ГИС, такими как Autodesk Map, Arcinfo, MapInfo, Геокад.

Технология создания и внедрения ЭГП должна содержать следующие основные этапы:

- перевод имеющихся «твердых копий» (планшетов ДГП, документов архива генплана) в электронную форму;
- создание классификатора слоев графической части ЭГП и векторизация генплана;
- проектирование единого, централизованного хранилища данных ЭГП;
- выбор ГИС-решения, которое будет являться основой системы (ГИС-платформой);
- разграничение прав доступа к данным ЭГП на основе утверждаемого регламента;
- организация удобного доступа к ЭГП проектировщиков, диспетчеров, администрации предприятия, а также всех заинтересованных специалистов.

В начале процесса создания ЭГП вся имеющаяся информация должна быть переведена в удобную для работы форму, сканируются все документы генплана и в структурированном виде размещаются на файловом сервере с определенными правами доступа к ним.

Следующим этапом на основе нормативных актов (ГОСТ, СНиП, СН и т.п.) и принятых на предприятии положений создается классификатор слоев ЭГП. Для удобства использования и анализа слои группируются в группы по назначению (виду инженерной сети). Из всего объема графических данных, полученных на первом этапе, отбираются актуальные данные, на основе которых создается пространственная база данных (ПБД) электронного генплана.

Затем необходимо на базе корпоративной вычислительной сети предприятия создать централизованное хранилище данных (Data Warehouse), в котором будет аккумулироваться вся пространственная и атрибутивная информация.

Для обеспечения автоматизации процессов работы с электронным генеральным планом инженерной инфраструктуры предприятия необходимо внедрить ряд программных подсистем, обеспечивающих необходимый функционал (рис. 1).

Программные подсистемы непосредственно взаимодействуют с хранилищем данных – базой данных электронного генплана (БД ЭГП) [2]. Эти данные включают пространственную (графическое представление объекта на карте) и атрибутивную (описательную) информацию об объектах инженерной инфраструктуры, зданиях и сооружениях предприятия. Пространственная информация представляется в хранилище как пространственная схема данных, организованная в реляционной СУБД, а также файлы, содержащие пространственную информацию (растровые, векторные). Помимо этих двух схем данных, в хранилище содержатся метаданные, дополняющие структуру основных данных. Система хранения данных имеет следующие возможности:

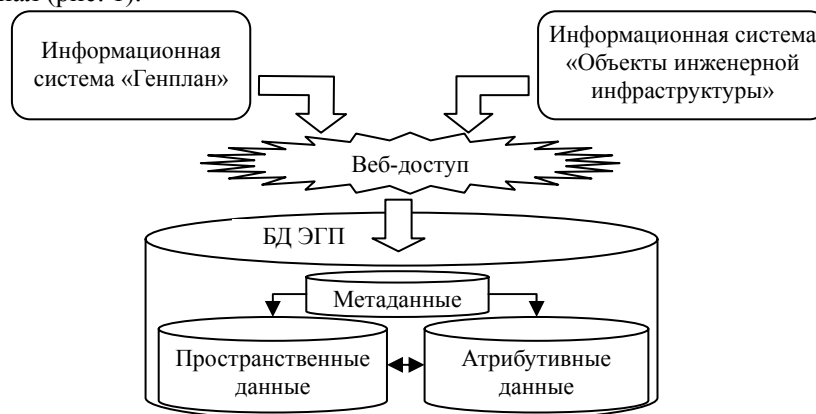


Рис. 1. Архитектура информационной системы управления электронным генеральным планом

– хранение пространственной информации в пространственной схеме СУБД;

- журналирование действий пользователя по работе с объектами;
- ведение архива пространственных данных;
- резервное копирование данных;
- многопользовательский доступ к данным;
- гибкая модель данных, позволяющая детально описать каждый объект инженерной инфраструктуры;
- безопасность данных, хранимых в едином хранилище, с определенным регламентом доступа.

Информационная система «Генплан» предполагает ведение пространственной информации БД ЭГП. С этой системой работают сотрудники ПКО предприятия. Система содержит инструменты, предоставляющие следующие возможности:

- работа с пространственной схемой данных: создание, изменение и удаление объектов. Создаваемые в этой системе объекты – это графические объекты с некоторым минимумом описательной информации, которая включает: наименование объекта, источник данных, на основе которого введен объект (топосъемка);
- импорт данных и работа с локальным вариантом части ЭГП, экспорт данных из локального варианта в единое хранилище ЭГП;
- поиск объектов по их описательной информации;
- работа с описательной информацией объектов (копирование, редактирование, прикрепление, открепление);
- введение в хранилище данных информации с оцифрованных носителей;
- классификация объектов инженерной инфраструктуры по слоям;
- автоматический контроль типов вводимых объектов в слоях в соответствии с принятым классификатором (тип графического примитива);

- получение метаданных: данных об ограничениях, стилях и цветах слоев ЭГП;
- сбор и обработка статистики об объектах ЭГП;
- оперативное взаимодействие бюро генплана с инженерными службами предприятия (обмен пространственной информацией).

Информационная система «Объекты инженерной инфраструктуры» предполагает ведение атрибутивной информации БД ЭГП. Ее использование ведут сотрудники инженерных служб предприятия. В рамках этой системы авторы приняли участие в реализации ряда подсистем:

- подсистема Web-ГИС-сервер, предназначенная для обеспечения Web-доступа к средствам хранения и анализа данных электронного генерального плана;
- подсистема Web-ГИС-клиент, предназначенная для обеспечения графического интерфейса конечного пользователя электронного генерального плана (ЭГП) в среде Интернет/Интранет;
- подсистема информационной безопасности, предназначенная для обеспечения информационной защищенности пространственных и атрибутивных данных ЭГП на основе авторизованного доступа к пространственным данным и функциональной составляющей программного комплекса;
- подсистема интеллектуального анализа данных, предназначенная для обеспечения извлечения новых знаний из информации хранилища пространственно-временных данных и прогнозирования динамики процессов жизненного цикла инженерной инфраструктуры предприятия;
- подсистема организации документооборота электронного генерального плана, предназначенная для обеспечения организационно-распорядительного механизма развития электронного генерального плана на всех этапах его жизненного цикла.

Подсистема веб-ГИС-сервер. Подсистема веб-ГИС-сервер обеспечивает функциональные возможности, реализуемые компонентами, входящими в его состав:

- а) формирование данных в виде xml-файла, необходимых для публикации средствами Web-ГИС-клиента;
- б) размещение сформированного для публикации файла на Web-сервере;
- в) поддержка многослойного представления электронного генплана. Возможность создания, настройки и удаления как отдельного слоя, так и организации групп слоев по определенной проблемной направленности. Каждая группа слоев и слой должны иметь уникальное наименование. Слои должны объединять в своем составе однотипные геометрические объекты (точки, линии, полигоны);
- г) выполнение пространственно-атрибутивных запросов пользователей, формируемых с помощью Web-ГИС-клиента;
- д) выполнение функций оперативного геоинформационного анализа, в состав которых входят геометрический анализ (измерение расстояний, дистанций, площадей); тематическая классификация; построение буферных зон. Геометрический анализ зависит от типа геометрии объектов (полигон, линия, точка). Для полигональных объектов рассчитываются периметр и площадь, для линейных – длина. Также для всех типов объектов отображаются координаты точек, из которых образуются объекты;
- е) выделение пространственных объектов на основе атрибутивного запроса.

Подсистема веб-ГИС-клиент. Подсистема веб-ГИС-клиент обеспечивает функциональные возможности, реализуемые компонентами, входящими в его состав:

- а) доступ к графическим и атрибутивным данным электронного генплана. Выбор объектов карты может осуществляться следующими способами: выбор объектов левой клавишей мыши, выбор объектов полигоном, выбор объектов окружностью;
- б) доступ пользователя к функциям геоинформационной системы, поддерживаемым Web-ГИС-сервером;
- в) формирование запросов на выборку пространственных данных на основе значений атрибутивных их описаний;
- г) интерактивное измерение расстояний, дистанций и площадей на карте;
- д) отображение результатов тематического анализа;
- е) интерактивное редактирование атрибутивного описания объекта электронного генплана;
- ж) динамическое масштабирование графического представления карты.

Подсистема информационной безопасности [3]. Подсистема информационной безопасности обеспечивает функциональные возможности, реализуемые компонентами, входящими в его состав:

а) ведение регламента доступа к пространственным данным и функциональной составляющей системы Web-ГИС;

б) организация ролевого регламентированного доступа к данным генерального плана согласно указанным администратором объектам, группам объектов, функциям и территориям;

в) аудит действий пользователей в среде ЭПП. Основными контролируруемыми действиями должны являться создание пространственных объектов, просмотр сведений об объектах, редактирование пространственно-атрибутивных данных, удаление данных;

г) назначение пользователям и их группам права на элементы пользовательского интерфейса веб-ГИС-клиента: ГИС-слои, информационные отчеты, ГИС-функции;

д) формирование отчетов о доступе пользователей к объектам хранилища пространственных данных и действиям над этими объектами за указанный период времени.

Подсистема интеллектуального анализа данных. Подсистема интеллектуального анализа данных обеспечивает функциональные возможности, реализуемые компонентами, входящими в его состав:

а) классификация объектов электронного генерального плана инженерной инфраструктуры предприятия на основе технологии нечеткого моделирования;

б) аппроксимация пространственно-временных данных, описывающих состояние элементов инженерной инфраструктуры, на основе нечеткого моделирования.

Подсистема организации документооборота электронного генерального плана. Подсистема организации документооборота электронного генерального плана обеспечивает функциональные возможности, реализуемые компонентами, входящими в его состав:

а) ведение первичной регистрации всех типов документов, находящихся в общем потоке документооборота ЭПП;

б) организация автоматизированного контроля исполнения документов: постановка документов на контроль; формирование напоминаний и писем-предупреждений о необходимости исполнения в срок соответствующих документов; продление срока исполнения и снятие документа с контроля;

в) ведение технологии электронного взаимодействия между подразделениями организации (пользователями ЭПП) посредством работы с документами в электронном виде (просмотр документов, просмотр резолюций руководителя, добавление резолюции и т.д.).

Заключение. Уточненные методы решения задачи создания и ведения электронного генплана, структуры входных и выходных его данных, общее описание алгоритма поддержки его жизненного цикла разработаны при финансовой поддержке Министерства образования и науки Российской Федерации в рамках мероприятия 2.4 федеральной целевой программы «Исследования и разработки по приоритетным направлениям развития научно-технологического комплекса России на 2007–2013 годы», государственного контракта № 07.524.11.4013.

Литература

1. Гриценко Ю.Б. Геоинформационные технологии мониторинга инженерных сетей / Ю.Б. Гриценко, Ю.П. Ехлаков, О.И. Жуковский. – Томск: ТУСУР, 2010. – 148 с.

2. Жуковский О.И. Архитектура корпоративной WEB-ориентированной ГИС / О.И. Жуковский, Н.Б. Рыбалов // Доклады ТУСУРа. – 2008. – № 2 (18), ч. 2. – С. 46–57.

3. Ехлаков Ю.П. Методы анализа защищенности образовательных ресурсов / Ю.П. Ехлаков, Р.В. Мещеряков // Доклады ТУСУРа. – 2007. – № 2 (16). – С. 18–24.

Ехлаков Юрий Поликарпович

Д-р техн. наук, профессор, проректор ИиУ ТУСУРа

Тел.: (382-2) 53-24-20

Эл. почта: ure@muma.tusur.ru

Гриценко Юрий Борисович

Канд. техн. наук, доцент кафедры АОИ ТУСУРа

Тел.: (382-2) 90-01-80

Эл. почта: ubg@muma.tusur.ru

Жуковский Олег Игоревич

Канд. техн. наук, доцент кафедры АОИ ТУСУРа

Тел.: (382-2) 90-01-80

Эл. почта: ol@muma.tusur.ru

Ehlaikov Yu.P., Gritsehko Yu.B., Zhukovskiy O.I.

Automation technologies for shaping and monitoring of electronic master plan engineering of infrastructure

The paper considers a number of common features of engineering systems, the technology, the usage of master plans for large industrial enterprises which use engineering networks. There are described the steps for creating and maintaining an electronic master plan, we propose the architecture of an electronic master plan of information system management, which presents information subsystem for managing large engineering infrastructures of industrial enterprises, as well as their functional composition is described.

Keywords: electronic master plan, engineering infrastructure, geographic information systems, data mining, electronic document management.

УДК 681.5.017

А.А. Шилин, В.Г. Букреев, Е.И. Гладышева

Моделирование нелинейных систем на FBD-блоках с ограниченным базисом

Предложены процедура моделирования и организация алгоритмов управления в реальном времени сложными объектами в программной среде Functional Block Diagram (FBD). Рассмотрен иллюстративный пример решения задачи управления теплообменной системой, представленной нелинейной моделью в пространстве состояний. Приведены алгоритм преобразования Simulink-модели теплообменной системы в структуру FBD и сравнительная оценка результатов моделирования.

Ключевые слова: моделирование нелинейных систем, функциональные блоки, реализация на ПЛК.

Значительная часть исследований в области теории автоматического управления содержит прикладную часть работы в виде результатов экспериментов и моделирования, представленных в виде программ на языке MatLab или структур, организованных в среде Simulink. Это определяется тем, что основная работа, связанная с анализом свойств замкнутых систем управления, достаточно эффективно выполняется на персональном компьютере (ПК). Однако среда программирования контроллеров не совместима с программным кодом, полученным в среде MatLab, и требует решения проблемы реализации этого алгоритма непосредственно в контроллере. Цель работы: показать возможность моделирования нелинейных объектов в реальном времени. Назначение статьи: предложить технологию имитационного моделирования и тестирования систем управления в среде FBD для программирования контроллера.

Краткий обзор методов реализации алгоритмов в контроллерах. Решение задач синтеза и реализация систем управления требуют выполнения предварительного моделирования с последующей коррекцией алгоритмов на технологическом объекте для исключения ошибок, которые могут стать причиной повреждения дорогостоящего оборудования. Мощный математический пакет программного обеспечения Simulink, который входит в состав MatLab, интуитивно понятен для специалистов, работающих непосредственно с оборудованием. Так, в статье [1] хорошо представлен подход к реализации научных разработок до стадии виртуального лабораторного стенда на базе пакет Simulink в рамках возможностей ПК.

Известны также и инженерные, удобные для практического использования методы настройки и подходы к изучению работы регулятора, где проверка результатов моделирования поддерживается программной средой Simulink [2]. Оптимальные настройки регулятора можно получить, используя известные, реализованные в MatLab такие методы, как: методы Циглера–Николса, Шеделя и др., входящие в пакет NCD [3]. Однако размещение в контроллере системы управления исходного кода на языке MatLab не представляется возможным, алгоритм работы регулятора необходимо реализовать на языке программно-логического контроллера (ПЛК). Следует отметить, что таких языков достаточно много: низкоуровневые языки Ассемблер, Си, Tehno-IL; языки высокого уровня Си++, Питон и др. Кроме того, некоторые ПЛК позволяют выполнять скрипты, построенные на основе диаграммы из функциональных блоков в среде визуального конструирования: Functional Block Diagram (FBD). Набор блоков FBD призван соответствовать международному стандарту МЭК 61131, но, как правило, во многих ПЛК реализован лишь их минимальный набор. Методология построения систем управления на FBD-логике аналогична построению структур в среде Simulink, отличие заключается в достаточно больших возможностях Simulink, но которые невозможно полностью реализовать на ограниченной базе блоков среды FBD.

Предполагается, что специалисты в области синтеза систем управления в большей мере являются математиками, нежели программистами, и многие из них испытывают трудности в реализации своих научных разработок на реальных объектах управления. Тем не менее созданную модель Simulink можно использовать для автоматического создания структурированного текста, соответствующего международному стандарту МЭК 61131 и предназначенного для программирования ПЛК

при развертывании. Для этого существует программный пакет расширения Simulink PLC Coder™ [4], который формирует программу, код или скрипт, соответствующий разработанной математической модели в системе Simulink. Далее, если возможности FBD позволяют, код легко импортируется в среду моделирования, используемую для работы с целевым ПЛК. Это действительно удобный инструмент для решения прикладных научно-технических задач, который позволяет, не вдаваясь в подробности программирования контроллера, быстро реализовать исследуемую систему на действующем объекте управления. Однако существуют некоторые препятствия: пакеты MatLab и Simulink PLC Coder™ имеют значительную стоимость, также и список контроллеров, поддерживаемых пакетом Simulink PLC Coder™, небольшой и состоит обычно из дорогих ПЛК, чаще всего избыточных по вычислительным ресурсам при построении многоуровневых систем управления. Следует отметить, что кроме Simulink PLC Coder™ имеются другие пакеты, такие как «SoftLogic система МастерPLC», поддерживающие в меру эффективные и экономичные ПЛК, но они не имеют сопряжения с программным обеспечением MatLab.

Постановка задачи и процедура создания имитационной модели в среде FBD ПЛК. Для нелинейных систем с элементами запаздывания требуется апробировать технологию имитационного моделирования и тестирования в среде FBD ПЛК управляющего регулятора. Предполагается, что количество FBD-блоков ограничено вычислительными возможностями контроллера.

Рассмотрим нелинейный объект управления – теплообменную систему, состоящую из теплообменного аппарата, во внутреннем контуре которого используется электромеханический клапан, обеспечивающий стабилизацию температуры теплоносителя в контуре на основании измерительной информации датчика. Внешний контур теплообменного аппарата образован распределенной трубопроводной сетью, определяющей запаздывающую реакцию теплоносителя в этом контуре на протекающий теплоноситель во внутреннем контуре [5].

Рассматривая задачу стабилизации температуры теплоносителя во внутреннем контуре теплообменного аппарата, динамику процессов можно представить в пространстве таких переменных состояния, как нормированное значение положения штока h клапана относительно максимального значения, температуры $T_{\text{пд}}^{\circ}$ теплоносителя во внутреннем контуре, температуры $T_{\text{из}}^{\circ}$ корпуса изолятора датчика и температуры $T_{\text{д}}^{\circ}$ термодатчика. На основании практического опыта и априорных знаний о характере протекающих процессов математическая модель нелинейной теплообменной системы представляется в виде следующей системы дифференциальных уравнений:

$$\begin{cases} \frac{dh}{dt} = \frac{u(t)}{T_{\text{пр}}}, \\ \frac{dT_{\text{пд}}^{\circ}}{dt} = \frac{\left((T_1^{\circ} - T_{\text{обр}}^{\circ}) \cdot (1 - e^{-k_h \cdot h}) \cdot k_{\text{max}} \right) + T_{\text{обр}}^{\circ} - T_{\text{пд}}^{\circ}}{T_{\text{см}}}, \\ \frac{dT_{\text{из}}^{\circ}}{dt} = \frac{(T_{\text{пд}}^{\circ} - T_{\text{из}}^{\circ})}{T_{\text{из}}}, \\ \frac{dT_{\text{д}}^{\circ}}{dt} = \frac{(T_{\text{из}}^{\circ} - T_{\text{д}}^{\circ})}{T_{\text{д}}}, \end{cases} \quad (1)$$

где $u(t)$ – релейный сигнал управления, принимающий дискретные значения $[-1, 0, 1]$; $T_{\text{пр}}, T_{\text{см}}, T_{\text{из}}, T_{\text{д}}$ – соответственно постоянные времени электропривода клапана, процесса смешения теплоносителей, изолятора датчика и датчика; T_1° и $T_{\text{обр}}^{\circ}$ – соответственно температура теплоносителя от источника и температура обратного теплоносителя; k_h и k_{max} – параметры, характеризующие нелинейность процесса смешения теплоносителя во внутреннем контуре теплообменного аппарата, определяются для конкретных режимов работы и, в частности, могут быть приняты $k_h = 0,9$ и $k_{\text{max}} = 1,1$.

С учетом фиксированных значений k_h и k_{max} линеаризованная математическая модель в точках равновесного состояния нелинейной теплообменной системы может быть представлена в виде передаточной функции

$$W_{ob}(s) = \frac{k_0(h_0, T_1^\circ, T_{обр}^\circ)}{T_{пр} \cdot s \cdot (1 + T_{см} \cdot s) \cdot (1 + T_d \cdot s) \cdot (1 + T_{из} \cdot s)},$$

$$k_0(h_0, T_1^\circ, T_{обр}^\circ) = (T_1^\circ - T_{обр}^\circ) \cdot \left(1 - e^{k_h \cdot h_0}\right) \cdot k_{max} / h_0, \quad (2)$$

где k_0 – коэффициент передачи разомкнутой системы в точке равновесия, соответствующей положению клапана $h=h_0$.

Задача синтеза ПИ-регулятора. Для теплообменной системы, представленной в виде передаточной функции (2), достаточно использовать корректирующее звено, реализующее ПИ-закон регулирования

$$W_{pi}(s) = \frac{k_1 \cdot (1 + T_{рег} \cdot s)}{(1 + k_2 \cdot T_{рег} \cdot s)}, \quad (3)$$

где k_1 – коэффициент усиления регулятора. Параметр регулирования k_2 положительный, меньше единицы, как правило, выбирается в пределах $k_2 = (0,1 \dots 0,3)$ и имеет содержательный смысл относительного коэффициента усреднения при дифференцировании. Выбор значения этого коэффициента определяется качеством измерения регулируемой величины. Так, при отсутствии аддитивных шумов и эффекта квантования усреднять оценку значения производной не следует. В этом случае $k_2 = 0$ и выражение (3) будет соответствовать ПД-регулятору. Параметр $T_{рег}$ – постоянная времени регулятора. Если рассмотреть последовательное соединение (2) и (3), то пропорциональный k_p и интегрирующий k_i коэффициенты ПИ-регулятора определяются следующим равенством:

$$k_p = k_1 \cdot T_{рег} / T_{пр}; k_i = k_1 / T_{пр}. \quad (4)$$

Задача синтеза данного типа регулятора предполагает определение параметров $k_1, T_{рег}, k_2$ регулирования с учетом условий устойчивости замкнутой системы с обратной связью при возможных изменениях значений параметра k_0 . Предполагая, что все параметры, кроме k_0 , стационарные, можно отметить, что минимальный запас устойчивости и условия оптимального управления следует искать при максимальном значении коэффициента $k_0 = k_{0max}$. Очевидно, что уменьшение коэффициента k_0 приведет к затягиванию переходного процесса и увеличению запаса устойчивости, что является допустимым. Существует достаточно большое количество методов выбора оптимальных настроек ПИ-регулятора, использующих различные критерии оптимальности. В нашем случае, не отдавая предпочтений конкретным методам синтеза оптимальных регуляторов, для определения параметров закона регулирования будем использовать численный подход. Наиболее распространенным критерием оптимальности является квадратичный функционал качества

$$J(k_1, T_{рег}) = \int_{t=t_0}^{\infty} \Delta x(t)^2 \cdot (t - t_0) \cdot dt, \quad (5)$$

где $\Delta x(t)$ – значение рассогласования (ошибка регулирования), определяется равенством $\Delta x(t) = T_{задан}^\circ(t) - T_d^\circ(t)$. При вычислении функционала (5) рассматривается замкнутая система с объектом (2) и регулятором (3), на сигнал управления наложено ограничение $u(t) \leq 1$ с максимально возможным значением рассогласования $\Delta x_{max}(t) = 10^\circ\text{C}$. Результатом решения вариационной задачи относительно коэффициентов ПИ-регулятора является поверхность (рис. 1), на которой однозначно определяется минимум, соответствующий оптимальным параметрам $k_1, T_{рег}$.

Имитационное моделирование теплообменной системы. Работоспособность теплообменной системы и эффективность регулятора можно проверить на имитационной SimuLink модели или на ее аналоге в программной среде Xcos пакета SciLab [6]. Структура системы управления в среде Xcos пакета SciLab представлена на рис. 2.

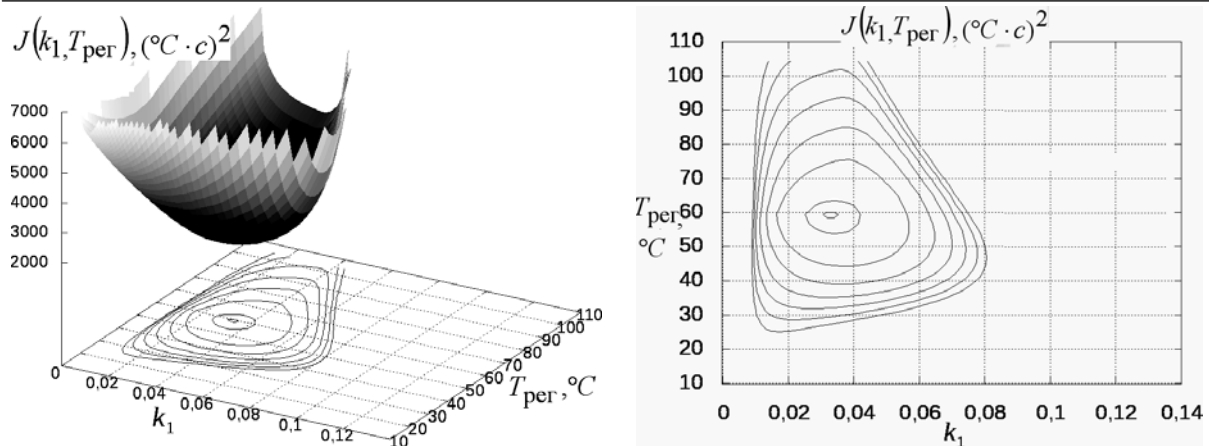


Рис. 1. Поверхность значений функционала в пространстве параметров ПИ-регулятора

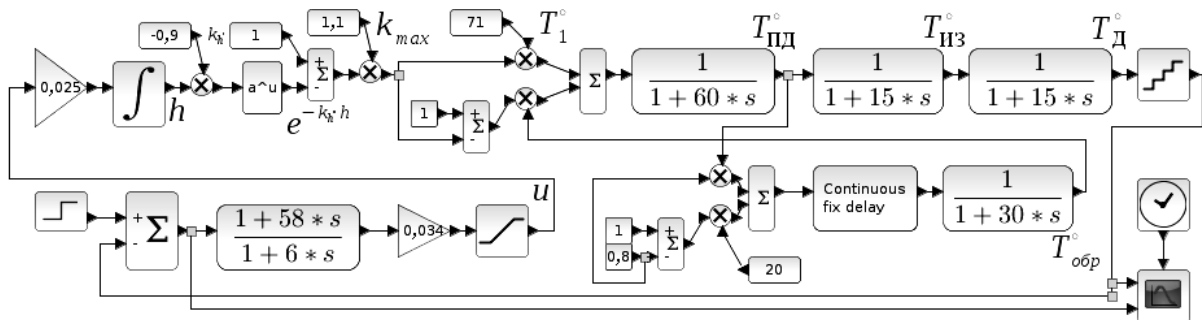


Рис. 2. Структура системы управления в среде Xcos (SimuLink)

В приведенной структурной схеме учитывается такое свойство теплообменной системы, как зависимость температуры теплоносителя во внешнем контуре от температуры теплоносителя внутреннего контура. Для реализации модели (1) в среде Xcos элемент задержки аппроксимируется цепочкой из последовательно включенных апериодических звеньев первого порядка, и на этапе моделирования можно выполнить необходимые численные эксперименты, подтверждающие работоспособность и эффективность алгоритмов управления теплообменной системой в реальном времени. Существенным затруднением при создании адекватной имитационной модели нелинейной системы в среде FBD используемого ПЛК [7] является ограниченность набора FBD-блоков, что значительно усложняет решение задач предварительного тестирования и перенастройки параметров регулятора. Одним из вариантов решения этой проблемы является предварительное преобразование априорной нелинейной модели системы в линейную модель путем корректной замены координат.

В нашем случае с целью исключения экспоненциальной нелинейности в первом уравнении системы (1) используем замену переменной $K_{CM}(t) = (1 - e^{-k_h \cdot h}) \cdot k_{max}$. В результате исходная математическая модель теплообменной системы записывается в следующем виде:

$$\begin{cases} \frac{dK_{CM}(t)}{dt} = (k_{max} - K_{CM}(t)) \cdot \frac{k_h}{T_{пр}} \cdot u(t), \\ \frac{dT_{пд}^{\circ}(t)}{dt} = \frac{\left((T_1^{\circ} - T_{обр}^{\circ}(t)) \cdot K_{CM}(t) \right) + T_{обр}^{\circ} - T_{пд}^{\circ}(t)}{T_{CM}}, \\ \frac{dT_{из}^{\circ}(t)}{dt} = \frac{\left(T_{пд}^{\circ}(t) - T_{из}^{\circ}(t) \right)}{T_{из}}, \\ \frac{dT_{д}^{\circ}(t)}{dt} = \frac{\left(T_{из}^{\circ}(t) - T_{д}^{\circ}(t) \right)}{T_{д}}. \end{cases} \quad (6)$$

Используя данную математическую модель теплообменной системы и алгоритм работы регулятора, можно создать имитационную модель непосредственно на FBD ПЛК (рис. 3).

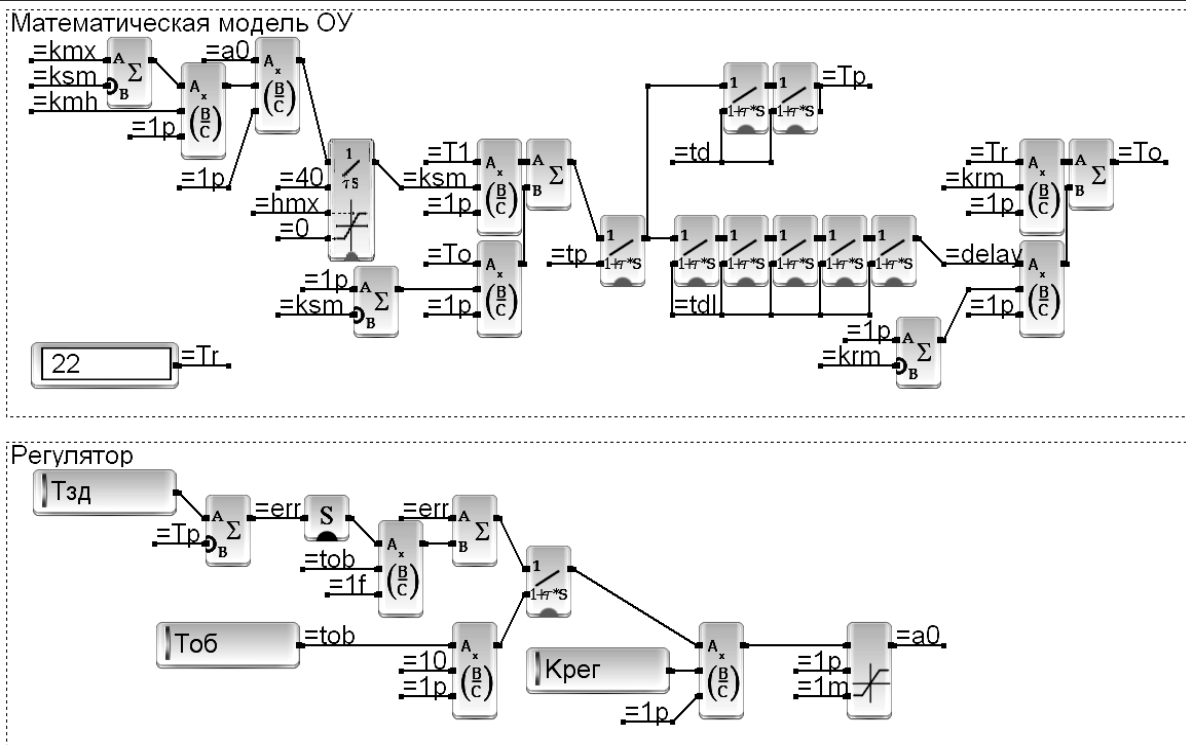


Рис. 3. Реализация теплообменной системы и ПИ-регулятора на FBD блоках ПЛК

На рисунке приняты обозначения регистров соответствуют переменным и параметрам для (3) и (6), диктуемые компилятором FBD и скрипта программы производителя: $=T_1 \Rightarrow T_1^\circ$; $=Tp \Rightarrow T_{пд}^\circ$; $=ksm \Rightarrow K_{см}$. Подобным образом в структуру введены константы $=Tr \Rightarrow T_{ком}^\circ = 20^\circ\text{C}$; $=kmh \Rightarrow k_h = 0,9$; $=ktx \Rightarrow k_{max} = 1,1$; $=p \Rightarrow 100$.

Некоторые результаты моделирования. Результаты имитационного моделирования и тестирования экспериментальной системы, представленные в виде переходных процессов, соответствуют положению клапана в окрестности открытого (рис. 4) и закрытого (рис. 5) состояния и отражают работоспособность имитационной модели теплообменной системы и ПИ-регулятора.

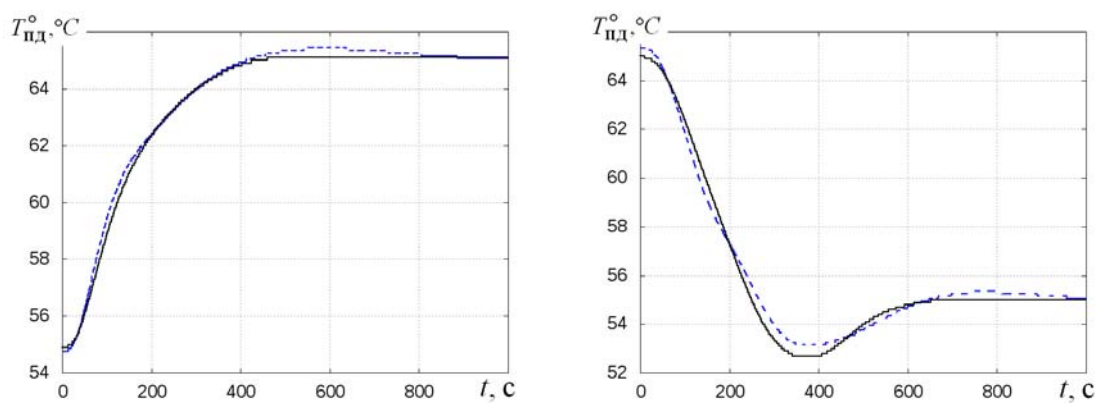


Рис. 4. Переходные процессы при открытом положении клапана

Различие переходных процессов обусловлено нелинейностью процесса смешения ОУ, представленного как в (1) так и в (6), причем незначительное отличие переходных процессов разными методами моделирования свидетельствует об адекватности обеих моделей ОУ. Отличие переходных процессов в одинаковых условиях объясняется тем, что в Simulink-модели (пунктирная линия) транспортное запаздывание реализовано в виде буфера оперативной памяти, а в FBD-модели (сплошная линия) – аппроксимацией инерционными звеньями.

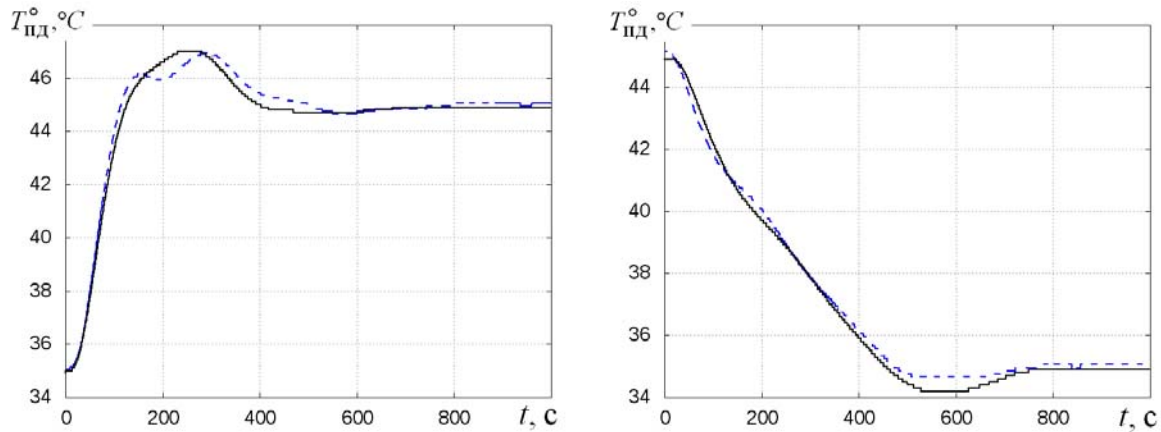


Рис. 5. Переходные процессы при закрытом положении клапана

Выводы

1. Опыт имитационного моделирования в среде FBD ПЛК показывает, что функциональные возможности FBD незначительно уступают возможностям пакета Simulink.
2. Скорость моделирования на FBD-блоках значительно выше, примерно в 8–10 раз, по сравнению со скоростью вычисления на Simulink и Xcos, потому что FBD-блоках реализован быстродействующий и оптимизированный под микроконтроллеры код подпрограмм.
3. Отсутствует проблема реализации алгоритма на действующем ПЛК, так как в результате моделирования в среде FBD полученный код можно в режиме реального времени импортировать из среды моделирования непосредственно в ПЛК для выполнения эксперимента на действующем объекте управления.
4. С точки зрения обучения специалистов настройке параметров регулирования представляет интерес тестирование работы регулятора совместно с моделью объекта управления в реальном времени непосредственно на ПЛК.

Литература

1. Герман-Галкин С.Г. Моделирование устройств силовой электроники. Создание виртуальных лабораторий в среде Matlab-Simulink // Силовая электроника. – 2008. – №2. – С. 144–150.
2. Настройка ПИД-регулятора. Aauto-works [Электронный ресурс]. – Режим доступа: <http://autoworks.com.ua/teoreticheskie-svedeniya/nastrojka-pid-regulyatora>, свободный (дата обращения: 06.02.2013).
3. Сидорова А.А. Исследование настройки ПИД-регулятора в Simulink-Matlab / А.А. Сидорова, А.М. Мальшенко // Технологии Microsoft в теории и практике программирования: тр. 7-й Всерос. науч.-практ. конф. – Томск: ТПУ, 2011. – С. 13–15.
4. Кодати П. Разработка и реализация на базе ПЛК комплексных стратегий управления/ П. Кодати, Т. Эрркинен, А. Туревский // Электронные компоненты: сетевой журн. 2012. – №3 [Электронный ресурс]. – Режим доступа: <http://sl-matlab.ru/news/PLC%20Modeling%20Simulink.pdf> свободный (дата обращения: 06.02.2013).
5. Шилин А.А. Нелинейная математическая модель теплотребления с учетом характеристик элементов теплового узла / А.А. Шилин, В.Г. Букреев // Научный вестник НГТУ. – 2012, №2(47). – С.107–114
6. SciLab. Open source software for numerical computation [Электронный ресурс]. – Режим доступа: <http://www.scilab.org/download/5.4.0> свободный (дата обращения: 06.02.13).
7. Автоматизированный комплекс исследования автоматических регуляторов [Электронный ресурс]. – Режим доступа: <http://npowest.ru/7-programmnoe-obespechenie.html> свободный (дата обращения: 06.02.13).

Шилин Александр Анатольевич

Канд. техн. наук, доцент каф. электропривода и электрооборудования НИ ТПУ

Тел.: 8 (382-2) 56-31-47

Эл. почта: shilin@tpu.ru

Букреев Виктор Григорьевич

Д-р техн. наук, профессор каф. электропривода и электрооборудования НИТПУ

Тел.: 8 (382-2) 56-30-45

Эл. почта: bukreev@tpu.ru

Гладышева Екатерина Игоревна

Студентка каф. промышленной электроники ТУСУРа

Тел.: 8-952-886-60-39

Эл. почта: yegl@yandex.ru

Shilin A.A., Bukreev V.G., Gladysheva E.I.

Simulation of nonlinear systems in FBD blocks with limited basis

The technology of modeling and organization of the control algorithms in real-time complex object in the software environment Functional Block Diagram (FBD) are offered. An illustrative example of the solution of the heat-exchange control system, represented by nonlinear state-space model is considered. The algorithm that transforms Simulink-model heat exchanger system in the structure of the FBD and comparative evaluation of the simulation results are presented.

Keywords: modeling of nonlinear systems, function blocks, implementation of the PLC.

УДК 004.94

В.В. Грачев, В.А. Силич, М.П. Силич

Методология проектирования ситуационных центров принятия решений

Рассматривается методология проектирования ситуационных центров в различных предметных областях. Описана типовая структура и функции ситуационного центра принятия решений.

Ключевые слова: ситуационный центр, технология проектирования, моделирование.

Назначение ситуационных центров. Экономический и научно-технический прогресс стран и регионов, укрупнение и усложнение объектов хозяйствования, экологические и техногенные проблемы породили и выдвинули в число весьма актуальных проблематику создания и развития ситуационных центров. Такие центры создаются на базе достижений современных информационно-коммуникационных технологий и их главным предназначением является сбор, обработка, отображение и анализ информации, поступающей с объектов мониторинга и управления, прогнозирование и моделирование кризисных ситуаций, выработка и оптимизация решений по их предотвращению и эффективной ликвидации.

Развитие современных информационных технологий стимулировало развитие систем поддержки принятия решений, в частности ситуационных центров. Современные ситуационные центры (СЦ) представляют (или должны представлять) собой совокупность новейших программно-технических средств, математических методов, сетевых и системотехнических решений для автоматизации процессов сбора, отображения, моделирования и анализа ситуаций и оперативного управления [1–7].

Ситуационные центры позволяют решать такие задачи, как:

- обеспечение информацией лиц, принимающих решения (ЛПР);
- предоставление доступа руководителям к территориально разнесенным программно-техническим компонентам подразделений;
- обеспечение согласованности и надежности функционирования программно-технических компонент системы;
- предоставление доступа к информации организаций, участвующих в принятии решения;
- сокращение временных и финансовых затрат, связанных с программно-технической несогласованностью информационно-телекоммуникационных систем, дублированием и несогласованностью данных, их противоречивостью, затруднениями с доступом, выборкой и передачей информации;
- формирование единого информационного пространства для ЛПР всех структурных подразделений.

Основные возможности, предоставляемые ситуационными центрами:

- обоснование принимаемых решений путем их оценки и моделирования на основе возможностей современных информационных технологий;
- повышение качества предварительного анализа информации и вырабатываемого решения путем использования современных информационных технологий;
- обеспечение интеграции систем связи, систем обработки данных и визуализации информации, систем принятия решений.

В настоящей статье предлагается методология проектирования ситуационных центров.

Этапы проектирования СЦ. В качестве основы взята технология реинжиниринга бизнес-процессов, видоизмененная с учетом объекта проектирования. Методология включает следующие 4 основных этапа.

Этап 1. Формирование директивы на выполнение проекта.

Директива на создание ситуационного центра служит основанием для начала работ. Директива должна быть составлена в терминах высокого уровня и выражать ожидания от реализации проекта по созданию СЦ. Прежде всего необходимо определить тип проектируемого центра, его основное назначение.

К основным типам ситуационных центров относят [2, 4]:

- СЦ анализа и управления кризисными ситуациями;
- СЦ виртуальной реальности;
- СЦ мониторинга и принятия стратегических решений.

Этап 2. Разработка концепции ситуационного центра.

На этом этапе формируется образ будущего центра, определяются основные цели и задачи. Спецификацию целей предлагается осуществлять на основе анализа целевых установок стейкхолдеров (заинтересованных лиц) – заказчиков, будущих участников процессов принятия решения, лиц, выступающих источниками информации и реализующих решения и т.д. По результатам анализа требований стейкхолдеров формируется система целей, отражающая предназначение СЦ, и разрабатываются основные стратегические решения относительно структуры центра. Строятся прототипы – сценарии работы будущего СЦ. Необходимо также определить список факторов успеха и риска неудач выполнения проекта. Помимо идентификации факторов риска и оценки их значимости и вероятности желательно предложить меры по их устранению или снижению степени влияния.

Этап 3. Обратный инжиниринг.

На данном этапе анализируется опыт создания и функционирования центров, подобных разрабатываемому. Это поможет снизить трудоемкость разработки, избежать типовых ошибок. В качестве образца необходимо выбирать СЦ, показывающие наилучшие результаты.

Этап 4. Прямой инжиниринг.

Это наиболее трудоемкий этап, на котором осуществляется проектирование центра в соответствии со стратегией, выдвинутой на этапе разработки концепции, с опорой на результаты этапа обратного инжиниринга. В самом общем виде технология проектирования включает следующие подэтапы:

1. Выявление типовых проблем и ситуаций, для решения которых создается ситуационный центр. Все выявленные типовые проблемы должны быть сгруппированы. Для каждой создается спецификация – структурированное описание, позволяющее ее идентифицировать.
2. Выбор или разработка моделей и методов для идентификации, анализа и решения выявленных типовых проблем.
3. Разработка обеспечивающих комплексов:
 - информационного – системы сбора, хранения, мониторинга, визуализации информации;
 - программно-технического (разработка архитектуры СЦ);
 - организационного – структуры управления и взаимодействия между участниками.
4. Разработка регламента технологии ситуационного управления для различных вариантов функционирования СЦ.
5. Создание ситуационного центра в соответствии с разработанным проектом, тестирование, оценка эффективности.

Дальнейшая детализация процесса разработки СЦ осуществляется с учетом типа и конкретных особенностей создаваемого центра.

Рассмотрим методологию выделения и описания типовых проблем (*этап 4.1*) для кризисного ситуационного центра, основными функциями которого являются:

- оценка текущей ситуации (тенденций в развитии ситуации), выявление проблемной ситуации (неблагоприятного развития ситуации);
- выбор решения для устранения проблемной ситуации, оценка последствий.

Выявление типовых проблем, решаемых кризисным ситуационным центром. Необходимо выявить множество типовых потенциально опасных ситуаций в заданной предметной области. Формирование данного множества может осуществляться как путем группировки случившихся в прошлом нежелательных ситуаций, так и с помощью декомпозиции рассматриваемой предметной области. В первом случае сначала группа экспертов формирует список составляющих потенциальную угрозу ситуаций, а затем схожие ситуации группируются. Во втором случае осуществляется последовательное разбиение предметной области на подсистемы в соответствии с различными признаками (основаниями декомпозиции) и выделение ситуаций для каждой подсистемы. Таким образом, формируется иерархия, в которой ситуации сгруппированы по различным признакам, например по аспектам (производственному, социальному, экономическому, экологическому), по территориальному признаку, по типам объектов мониторинга, по видам деятельности и т.д. Использование стандартных оснований декомпозиции позволяет обеспечить полноту иерархии.

Для каждой из выделенных ситуаций создается описание. Типовые ситуации могут быть представлены в виде прецедентов – структурированных описаний, представляющих собой набор слотов (атрибутов). Ключевыми слотами такого описания являются: индикаторы; оценки степени опасности ситуации (классы опасности); меры реагирования на ситуацию.

Индикаторы – это специальные атрибуты, свидетельствующие о наступлении или возможности наступления ситуации. Индикаторами могут выступать не только конкретные значения некоторого показателя, но и индексы изменения показателя по отношению к некоторому базовому периоду, характеристики тренда (тенденция изменения), частотные характеристики (например, частота попадания в заданные области значений), а также прогнозируемые значения показателя. Значения индикаторов могут непосредственно измеряться, вводиться экспертами (в том числе в виде балльных оценок, лингвистических значений) либо определяться на основе исходных данных с помощью выбранного метода. При этом могут быть определены нечеткие значения индикатора, т.е. значения, характеризующиеся факторами уверенности. Это может быть обусловлено как нечеткими исходными данными (например, в ходе доопределения ситуации данные определяются с некоторой долей уверенности), так и использованием систем нечеткого вывода. Кроме того, один и тот же индикатор может определяться на основе различных источников, которые могут как подкреплять друг друга (значение фактора уверенности при этом увеличивается), так и противоречить друг другу (значение фактора уверенности снижается).

Для того чтобы по индикаторам можно было оценить возможность возникновения ситуации, они должны фиксировать наличие предпосылок. Одним из способов выявления подобных индикаторов является построение дерева причин. Данный метод состоит в последовательном выстраивании цепочек причинно-следственных связей от ситуаций к коренным причинам их возникновения. Затем предлагаются показатели, по которым можно судить о наличии той или иной причины. Изучение статистики изменения данного показателя в прошлом и соотнесение со статистикой появления исследуемой ситуации может помочь подтвердить или опровергнуть гипотезу о связи между динамикой показателя и наличием ситуации.

Индикаторы используются для определения уровня опасности ситуации. Для каждой отслеживаемой типовой ситуации могут быть выделены уровни (классы) опасности, например: нормальная ситуация, вызывающая определенные опасения, опасная, кризисная, катастрофическая. Классы опасности, как и индикаторы, могут быть представлены в виде нечетких значений лингвистической переменной. Это позволяет задавать «размытые» границы классов опасности. При этом формируются функции принадлежности на базовом множестве значений индикатора. Классификация на основе нескольких индикаторов может осуществляться различными способами. Один из них – объединение нечетких множеств, полученных на основе различных индикаторов. Как правило, используется либо формула тах-объединения, либо формула выпуклой комбинации (в случае, если заданы веса индикаторов). Второй способ – вывод на модели функциональных зависимостей, связывающей показатель опасности с индикаторами. В частности, зависимость уровня опасности от значений индикаторов может задаваться в виде системы нечетких правил-продукций «ЕСЛИ ... И ... ТО ...» или в виде нечеткой нейронной сети. Объединение нейронных сетей и нечеткой логики более предпочтительно, поскольку дает возможность обучения модели путем подстройки параметров функций принадлежности. Помимо автоматического определения класса проблемной ситуации, может быть предусмотрена возможность и оценки ее человеком.

Для каждой типовой ситуации и для каждого уровня опасности (за исключением нормального уровня) заранее определяются возможные способы реагирования. Если степень принадлежности текущей ситуации к одной из типовых ситуаций достаточно высока (близка к 1), то используется типовое решение, связанное с данной стандартной ситуацией (при этом оно может быть конкретизировано – настроено на конкретные данные, характеризующие текущую ситуацию). Если же степень принадлежности текущей ситуации к одной или нескольким стандартным ситуациям – на среднем уровне (например, от 0,5 до 1), то либо окончательный выбор делает человек, либо система осуществляет оценку последствий реализации решений и выбор. Если же ситуация нестандартна, т.е. степень принадлежности к любой из стандартных ситуаций низка (например, менее 0,5), то осуществляется генерация решений в автоматическом или автоматизированном режиме, а затем также может быть проведена оценка последствий и выбор. Для оценки последствий реализации тех или иных решений и выбора оптимальных решений используются заранее разработанные модели, в частности, имитационные модели, когнитивные, функциональные сети, нейросети и др.

В целях структурирования информации, описывающей ситуацию, предлагается использовать объектно-ориентированную методологию [8, 9]. На рис. 1 представлен фрагмент диаграммы классов объектно-ориентированной модели оценки ситуаций.

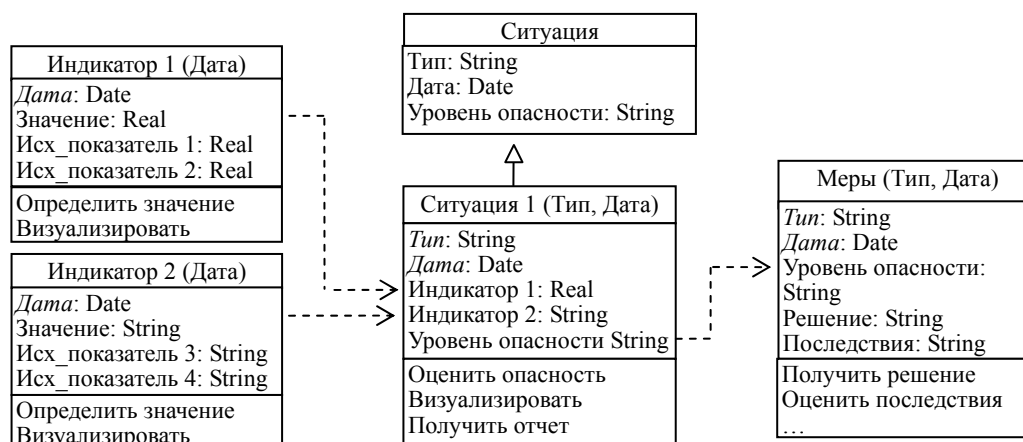


Рис. 1. Диаграмма классов модели оценки ситуаций

Класс описания ситуации включает ключевые атрибуты, содержащие тип ситуации и дату оценки, а также атрибуты для хранения значений индикаторов и уровня опасности. К методам относятся процедура оценки опасности ситуации, метод визуализации результатов оценки и ряд служебных процедур для отправки сообщений, формирования отчетов и т.д. Информацию, относящуюся к определенному индикатору, лучше хранить в отдельном классе, т.к. один и тот же индикатор может использоваться для оценки различных ситуаций, и к тому же даты вычисления значений индикаторов и оценки ситуаций могут не совпадать. Класс описания индикатора кроме ключевого атрибута с датой включает атрибуты с его текущим значением и значениями показателей, являющихся исходными данными, методы для получения исходных данных, вычисления индикатора, визуализации результатов расчета и ряд других процедур. Для выбора мер реагирования на ситуацию используется отдельный класс, содержащий процедуру получения решения с учетом типа и уровня опасности ситуации, метод оценки последствий и ряд служебных процедур.

Запуск процедуры оценки текущей ситуации на степень принадлежности тому или иному уровню опасности заданной типовой ситуации осуществляется либо по запросу пользователя, либо при наступлении даты оценки (в соответствии с заранее составленным графиком), либо при превышении значений индикаторов некоторых пороговых величин. При этом формируется экземпляр класса ситуации, идентифицируемый значениями атрибутов с типом ситуации и датой оценки, запрашиваются текущие (прошлые, прогнозные) значения индикаторов, которые вычисляются методами соответствующих классов, выполняются процедуры определения уровня опасности и визуализации. Если уровень опасности не является нормальным, инициируются процедуры поиска типового решения, его настройки и оценки последствий.

Заключение. В данной статье предложена системная последовательность проектирования, которая может быть использована для проектирования ситуационных центров принятия решений в различных предметных областях. Более подробно рассмотрен этап выделения и описания типовых ситуаций для кризисного ситуационного центра. Дальнейшие исследования связаны с выбором или разработкой моделей и методов для идентификации, анализа и решения выявленных типовых проблем, разработкой обеспечивающих комплексов (информационного – системы сбора, хранения, мониторинга, визуализации информации; программно-технического); организационно-правового обеспечения и др., разработкой регламента технологии ситуационного управления для различных вариантов функционирования ситуационного центра.

Литература

1. Поспелов Д.А. Ситуационное управление: теория и практика. – М.: Наука, 1986. – 288 с.
2. Бекренев В. Ситуационные центры и социально-экономическое моделирование // Управление персоналом. – 2000. – №12. – С. 61–63.

3. Филиппович А.Ю. Интеграция систем ситуационного, имитационного и экспертного моделирования. – М.: ООО Эликс+, 2003. – 300 с.
4. Филиппович А.Ю. Ситуационные центры: определения, структура и классификация [Электронный ресурс]. – Режим доступа: http://it-claim.ru/Library/Articles/publications_Philippovich_Andrew/sit_centres/sit_centres.htm (дата обращения: 13.12.2012).
5. Райков А.Н. Ситуационная комната для поддержки корпоративных решений // Открытые системы. – 1999. – № 7–8. – С. 56–66.
6. Мещеряков Р.В., Росошек С.К., Шелупанов А.А., Сонькин М.А. Криптографические протоколы в системах с ограниченными ресурсами // Вычислительные технологии. – 2007. – №12.1 – С. 51–61.
7. Клыков Ю.И. Ситуационное управление большими системами. – М.: Энергия, 1974. – 213 с.
8. Силич М.П. Технология разработки целевых программ на основе объектно-ориентированного подхода. – Томск: Том. гос. ун-т систем управления и радиоэлектроники, 2007. – 207 с.
9. Силич В.А., Силич М.П. Системная технология, использующая объектно-ориентированный подход // Известия Томского политехнического университета. – 2009. – Т. 314, № 5. – С. 155–160.

Грачёв Владимир Викторович

Канд. военных наук, нач. отделения ФГУП НИИ «Квант»

Тел.: 8-913-825-78-62

Эл. почта: vas@tpu.ru

Силич Виктор Алексеевич

Д-р техн. наук, профессор кафедры оптимизации систем управления НИ ТПУ

Тел.: 8-913-825-78-62

Эл. почта: vas@tpu.ru

Силич Мария Петровна

Д-р техн. наук, профессор кафедры автоматизации обработки информации ТУСУРа

Тел.: 8-913-825-78-62

Эл. почта: vas@tpu.ru

Grachev V.V., Silich V.A., Silich M.P.

Methodology of design of decision-making situation centers

The article presents a methodology of design of situation centers in various subject areas. The functions and a typical structure of a decision-making situation center are described.

Keywords: situation center, design technology, modeling.

УДК 65.011.56

И.В. Лазарев

Математические модели планирования технологических регламентов реконструкции и ремонта инженерных сетей

Рассмотрена математическая модель планирования технологических регламентов реконструкции и ремонта инженерных сетей. На примере задачи по реконструкции и ремонту участков сети холодного водоснабжения предприятия приведен пример расчета сетевого графика регламентных мероприятий с использованием правила предпочтения «кратчайшая операция».

Ключевые слова: технологический регламент, календарный план, сетевой график.

Основные понятия и определения регламента. Успешное достижение организацией целевых результатов зависит от способности руководства спланировать деятельность подчиненных и обеспечить её оперативное исполнение по заданным правилам и с гарантированным качеством. Как правило, эти вопросы рассматриваются при регламентации основных и вспомогательных бизнес-процессов деятельности организации.

В литературе есть несколько определений термина «регламент». Толковый словарь Ожегова дает два определения [1]: правила, регулирующие порядок какой-нибудь деятельности; время, отведенное на собрании для речи, выступления. Толковый словарь Даля дает следующее определение термина «регламент» [2]: устав, порядок или правила какой-либо службы, разъясненные на письме. В государственном стандарте ГОСТ Р 522294–2004, регламент определен как «совокупность правил, устанавливающих порядок проведения работ или осуществления деятельности» [3].

Любая регламентная деятельность осуществляется определенным субъектом. В качестве субъектов деятельности, участвующих в реализации всех этапов жизненного цикла регламента, могут выступать организация в целом, отдельное подразделение, конкретный сотрудник. Обязательное введение при описании регламентов параметров времени и исполнителей позволяет обеспечить синхронизацию их деятельности как в пределах одного подразделения, так и при их взаимодействии. Исходя из вышеизложенного, можно выделить следующие основные свойства регламента деятельности:

- ориентация деятельности на достижение определенных целевых результатов;
- задание правил и порядка осуществления и последующего регулирования деятельности;
- определение сроков (времени) осуществления этой деятельности;
- распределение функций между исполнителями и задание правил взаимодействия между ними в процессе реализации деятельности;
- документальное оформление правил и порядка осуществления деятельности.

Тогда под регламентом будем понимать набор правил, определяющих содержание и порядок реализации конкретных функций подразделений и отдельных должностных лиц по достижении заданных целевых результатов деятельности, их документальное оформление при разработке и последующем исполнении. Для дальнейшего изложения материала введем понятие регламентного мероприятия как действия по реализации конкретной функции определенным исполнителем в определенный период времени.

С учетом вышеизложенного бизнес-процесс управления регламентной деятельностью организации можно представить в виде следующих этапов:

- построение нормативной модели регламента в виде совокупности взаимосвязанных процессов (регламентных мероприятий);
- разработка календарного плана реализации регламента;
- организация контроля, анализ и регулирования исполнения регламента.

Для построения нормативной модели регламента в виде совокупности взаимосвязанных регламентных мероприятий предлагается использовать методы системной декомпозиции [4].

Процедура декомпозиции предполагает итерационное разбиение целого на части (системы на подсистемы, цели на подцели, задачи на подзадачи и т.д.). Для того чтобы обеспечить системность

декомпозиции, однозначность разбиения целого на части и обосновать элементы результирующего множества составных частей, при декомпозиции будем использовать формальную модель декомпозиции типа «жизненного цикла», которая обеспечивает выделение строго упорядоченной совокупности элементов, описывающих эволюционное изменение состояния системы.

Важной характеристикой процедуры декомпозиции является «глубина» разбиения целого на части. Понятие глубины декомпозиции в данном случае связано с понятием принципа элементарности. Использование этого принципа предполагает продолжение декомпозиции до получения результата, не требующего дальнейшего разложения, т.е. «программного» мероприятия, простого и понятного исполнителю. Такие регламентные мероприятия будем называть элементарными.

Зачастую для управления регламентными мероприятиями создаются временные органы (различные комиссии, оргкомитеты, рабочие группы), выполняющие набор вполне конкретных управленческих функций, содержание и организационно-технологическая схема реализации которых периодически повторяются. Отсутствие постоянных органов управления такого рода регламентными мероприятиями и следующая из этого несогласованность действий руководства являются основной причиной трудностей, возникающих на этапе подготовки решений по проведению регламентных мероприятий. Эти трудности имеют субъективный и объективный характер. Причиной трудностей субъективного характера являются периодические смены состава органов управления такими мероприятиями, что, как правило, приводит к необходимости самообучения со всеми вытекающими отсюда издержками [5].

К числу трудностей объективного характера следует отнести масштабность мероприятий, принципиальную сложность охвата всего комплекса работ на этапе планирования, а также сложность внесения корректировок на этапе оперативного управления. Кроме того, сложность проблемы усугубляется еще и периодическими изменениями условий проведения регламентных мероприятий (сроки, ресурсы, дополнительные работы и т.д.).

Вместе с тем таким мероприятиям присущи общие характерные признаки:

- периодичность (повторяемость) мероприятий;
- технологическая и ресурсная взаимосвязь составляющих мероприятия работ;
- относительная стабильность состава и структуры работ;
- межотраслевой характер и отсутствие постоянных органов управления;
- многовариантность стратегий реализации регламентных мероприятий.

Все представленное выше позволяет сделать вывод об объективной потребности и необходимости использования методов оптимизации при управлении регламентными мероприятиями деятельности организации. В данной статье рассматривается одна из перечисленных выше задач управления регламентной деятельностью организации – формирование календарного плана реализации технологических регламентов на примере деятельности по реконструкции и ремонту инженерных сетей.

Постановка задачи. В общем случае будем считать, что совокупность регламентных мероприятий связана некоторой технологией, т.е. для каждого мероприятия известно множество предшествующих мероприятий, после выполнения которых может быть начато выполнение данного. Формально такая технология может быть представлена в виде направленного графа $G=(X,U)$, где X – упорядоченное множество регламентных мероприятий (вершины графа); U – множество взаимосвязей между мероприятиями (дуги графа).

Взаимосвязи мероприятий могут иметь *технологическую* и *ресурсную* компоненты. Технологическая компонента определяет функциональную взаимосвязь мероприятий и определяется условием: *<последующее мероприятие начинается после полного завершения всех предшествующих мероприятий>*. При этом мероприятия в рамках одного регламента могут быть заданы как в виде линейной последовательности работ, так и в виде сетевого графика. Ресурсная компонента определяет факт наличия взаимосвязи между различными мероприятиями *при условии использования ими однотипных ресурсов*.

Задача формирования календарного плана регламентной деятельности может быть поставлена и решена в двух вариантах: как задача согласования работ и как задача оптимального распределения ресурсов при заданном времени выполнения работ. В данной работе внимание уделено постановке и решению задачи формирования календарного плана как задачи согласования работ.

Математическая модель календарного планирования исполнения регламентных мероприятий как задача распределения ресурсов при заданном времени выполнения работ. Пусть в заданном периоде времени в рамках реализации регламентной деятельности по реконструкции и ремонту инженерных сетей предприятие планирует провести работы на $I = \{1, 2, \dots, i, \dots, n\}$ объектах. Регламент работы по каждому из объектов инженерных сетей состоит из совокупности регламентных мероприятий – $P_i = \{1, 2, \dots, j, \dots, m\}$ и может быть представлен в виде строго заданной линейной последовательности. Прерывание исполнения регламентных мероприятий не допускается. Для проведения работ по реконструкции и ремонту инженерных сетей необходимо использовать $R = \{1, 2, \dots, r, \dots, k\}$ ресурсов. Каждый из ресурсов в определенный момент может использоваться только для исполнения одного мероприятия. Задана нормативная потребность в r -м ресурсе для выполнения j -го мероприятия на i -м объекте – p_{ij} .

Требуется определить множество времен начала и окончания регламентных мероприятий – $Q = \{t^H(x_i), t^K(x_i) | x_i \in P\}$ и потребности в ресурсном обеспечении при минимизации времени исполнения регламента $T = \max_{x_i \in X} t^K(x_i) \rightarrow \min$.

По аналогии с [5] представим графически процесс исполнения технологического регламента деятельности в виде точек на плоскости. Каждое из регламентных мероприятий как точка на плоскости, может быть описано следующим кортежем: $\langle i, j, t(i, j), r \rangle$. Каждая две точки могут иметь временную либо ресурсную зависимость.

Если $i_1 = i_2$, $j_1 \neq j_2$ и $r_1 \neq r_2$, то точки (i_1, j_1, r_1) и (i_2, j_2, r_2) связаны временной зависимостью в том смысле, что время начала выполнения одной, зависит от времени окончания другой. Графически эти две точки следует соединить связью типа «дуга», направленной от одной точки к другой.

Если $i_1 \neq i_2$, $j_1 \neq j_2$ и $r_1 = r_2$, то точки (i_1, j_1, r_1) и (i_2, j_2, r_2) связаны ресурсной зависимостью, и данные мероприятия не могут выполняться одновременно, однако очередность их выполнения заранее не задана. Такие две точки следует соединить связью типа «ребро».

В результате получаем смешанный граф $G = (X, \bar{U}, U)$, где X – множество операций (вершин); \bar{U} – множество дуг; U – множество ребер. Процедура определения календарного плана состоит в формировании ориентированного графа $G' = (X', U')$ путем замены ребер графа G на направленные дуги и расчета методом критического пути времени начала и окончания всех регламентных мероприятий графа G' .

Рассмотрим технологию построения смешанного графа на примере описания регламентов по реконструкции и ремонту участков сети холодного водоснабжения предприятия, когда мероприятия заданы в виде линейной последовательности работ.

Требуется произвести плановый ремонт сети холодного водоснабжения на четырех объектах $I = \{i_1, i_2, i_3, i_4\}$: i_1 – подземный участок трубопровода; i_2 – насосная станция; i_3 – наружный участок трубопровода; i_4 – распределительный узел.

Для выполнения ремонтных работ привлекаются четыре специализированные бригады – $R = \{r_1, r_2, r_3, r_4\}$: r_1 – бригада по обслуживанию экскаватора; r_2 – инженерная бригада; r_3 – бригада сварщиков; r_4 – бригада по благоустройству.

Распределение регламентных мероприятий по объектам ремонта и реконструкции, требуемые ресурсы и нормативная потребность в них (в человеко-часах) представлены в таблице.

С учетом описанных выше правил построения смешанного графа технологические регламенты по реконструкции и ремонту участков сети холодного водоснабжения можно представить в виде структуры (рис. 1).

Основная идея алгоритма *согласования работ* (регламентных мероприятий) сводится к получению ориентированного графа по определенным правилам из исходного смешанного графа путем замены ребер на дуги.

Алгоритм расчета календарного плана можно представить в виде следующей процедуры:

1. Выберем множество вершин графа G , в которые не входит ни одна дуга. Назовем это множеством множеством «ожидаемых» мероприятий.

2. Выберем по определенному правилу из множества «ожидаемых» мероприятий одно из них.
3. Заменяем все ребра, соединяющие данное мероприятие с другими мероприятиями графа G , на дуги. Удалим данное мероприятие из множества ожидаемых мероприятий.
4. Если множество ожидаемых мероприятий пусто, переходим к шагу 5, иначе к шагу 2.
5. Методом определения критического пути рассчитываем времена начала и окончания мероприятий на полученном графе G' .

Распределение нормативной потребности в ресурсах в рамках регламентов по реконструкции и ремонту участков сети холодного водоснабжения предприятия

Регламентное мероприятие		Объекты регламентной деятельности и их нормативная потребность в ресурсах (чел./ч)			
		i_1	i_2	i_3	i_4
p_{11}	Подготовка участка	$(r_4; 15)$			
p_{12}	Вскрытие грунта	$(r_1; 20)$			
p_{13}	Проведение ремонтно-восстановительных работ	$(r_3; 45)$			
p_{14}	Проведение испытаний	$(r_2; 30)$			
p_{15}	Восстановление грунта	$(r_1; 15)$			
p_{21}	Проведение ремонтно-восстановительных работ		$(r; 30)$		
p_{22}	Проведение испытаний		$(r_2; 30)$		
p_{31}	Подготовка участка			$(r_4; 27)$	
p_{32}	Наладка оборудования			$(r_2; 4)$	
p_{33}	Проведение ремонтно-восстановительных работ			$(r_3; 30)$	
p_{34}	Проведение испытаний			$(r_2; 30)$	
p_{41}	Проведение ремонтно-восстановительных работ				$(r_3; 36)$
p_{42}	Проведение испытаний				$(r_2; 20)$

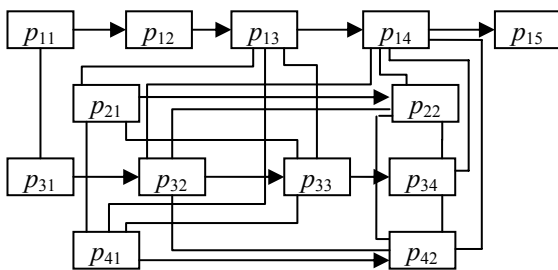


Рис. 1. Смешанный граф технологического регламента по реконструкции и ремонту участков сети холодного водоснабжения предприятия

Очевидно, что замена связи типа «ребро», на связь типа «дуга» может быть проведена на основе различных **правил**, каждое из которых может привести к различному результирующему графу G' . В теории расписаний такие правила получили название **функции предпочтения, или приоритетов**. В качестве правил предпочтения могут быть использованы следующие классические в теории расписаний функции предпочтения:

- Правило «первым пришел – первым обслуживается» – FIFO.
- Правило «последним пришел – первым обслуживается» – LIFO.
- Правило «кратчайшей операции» – SIO.
- Правило «первым уйдет – первым обслуживается» – FOFO.
- Правило «наибольшего оставшегося времени обслуживания» – LRT.

Помимо обозначенных выше правил, при разработке календарных планов технологических регламентов для объектов инженерной инфраструктуры может быть применено правило **«первым требуется – первым обслуживается»**. Данное правило предполагает приоритет исполнения мероприятий на тех объектах, директивный срок запуска в эксплуатацию которых минимальный.

На рис. 2 представлен вариант бесконтурного графа, порожденного исходным смешанным графом в результате замены всех его ребер дугами, на основе использования правила предпочтения SIO. Каждой дуге приписана длительность выполнения, каждой вершине – календарное время начала выполнения соответствующего мероприятия. Расчет произведен с учетом производительности труда каждой из бригад. Данный граф дополнен вершиной R – временем окончания всех регламент-

ных мероприятий. В соответствии с данным календарным планом работы в рамках технологического регламента по реконструкции и ремонту участков сети холодного водоснабжения предприятия можно завершить в течение 88 ч.

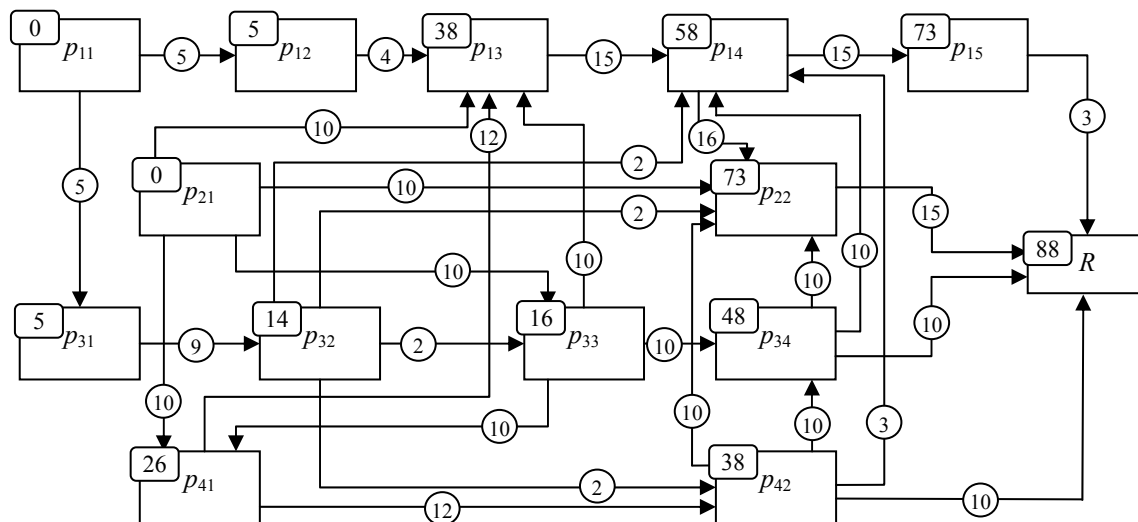


Рис. 2. Вариант ориентированного графа календарного плана технологического регламента

В классической литературе по теории расписаний [6–8] отмечается, что вряд ли удастся вывести какое-либо универсальное правило предпочтения для всех *возникающих* ситуаций, поэтому для получения наилучшего (оптимального) календарного плана рекомендуется:

- исследовать для конкретного объекта моделирования некоторое множество функций предпочтения, не противоречащих здравому смыслу;
- выявлять и ограничивать классы задач, для которых конкретные правила приводят к хорошим результатам;
- комбинировать использование различных правил с помощью случайного либо направленного поиска.

Заключение. Таким образом, задача разработки календарного плана технологического регламента методом сетевого планирования сводится к перебору вариантов ориентированных графов, полученных в результате применения различных правил предпочтения на этапе выбора мероприятия из множества ожидаемых, и выбору такого графа, который позволяет получить наименьшее время выполнения всех регламентных мероприятий. Данная задача может быть решена посредством разработки программного генератора допустимых расписаний.

Выполнение данной работы проводилось при финансовой поддержке Министерства образования и науки Российской Федерации в рамках мероприятия 2.4 федеральной целевой программы «Исследования и разработки по приоритетным направлениям развития научно-технологического комплекса России на 2007–2013 годы», проект «Разработка Web-ориентированных геоинформационных технологий формирования и мониторинга электронного генерального плана инженерной инфраструктуры», государственный контракт № 07.524.11.4013 от 03 ноября 2011 г.

Литература

1. Ожегов С.И. Толковый словарь русского языка: 80 000 слов и фразеологических выражений / С.И. Ожегов, Н.Ю. Шведова. – М.: Азбуковник, 1999. – 944 с.
2. Даль В.И. Толковый словарь русского языка / В.И. Даль. – М.: Эксмо, 2010. – 928 с.
3. ГОСТ Р 52294–2004 Информационная технология. Управление организацией. Электронный регламент административной и служебной деятельности. Основные положения. – М.: ИПК Изд-во стандартов, 2005. – 28 с.
4. Перегудов Ф.И. Основы системного анализа: учеб. пособие для студентов вузов / Ф.И. Перегудов, В.Ф. Тарасенко. – Томск: Изд-во НТЛ, 1997. – 397 с.
5. Ехлаков Ю.П. Теоретические основы автоматизированного управления: учебник / Ю.П. Ехлаков. – Томск: Том. гос. ун-т систем управления и радиоэлектроники, 2001. – 337 с.

6. Португал В.М. Модели планирования на предприятии / В.М. Португал, А. И. Семенов. – М.: Наука, 1978. – 272 с.
 7. Танаев В.С. Введение в теорию расписаний / В.С. Танаев, В.В. Шкурба. – М.: Наука, 1975. – 256 с.
 8. Конвей Р.В. Теория расписаний / Р.В. Конвей, В.Л. Максвелл, Л.В. Миллер. – М.: Наука, 1975. – 360 с.
-

Лазарев Иван Васильевич

Аспирант каф. АОИ ТУСУРа

Тел.: +7-913-812-46-79

Эл. почта: yalazarev@gmail.com

Lazarev I.V.

Mathematical models of the process of planning in public utility infrastructure

A mathematical model of the process of planning of reconstruction and repair of utility infrastructure is described. An example of calculation of the activity network is given for the problem of reconstruction and repair of cold water supply network.

Keywords: production schedules, schedule, network schedule.

УДК 004.94:551.345

В.Ю. Полищук

Программный комплекс имитационного моделирования динамики полей термокарстовых озер в зонах вечной мерзлоты

Описан комплекс программ имитационного моделирования полей термокарстовых озер, в состав которого входят следующие основные компоненты: подсистема моделирования, подсистема отображения результатов моделирования и база данных. Проведено прогнозирование динамики полей термокарстовых озер на территории многолетней мерзлоты Западной Сибири с учетом повышения среднегодовой температуры воздуха до 2030 г.

Ключевые слова: имитационное моделирование, программный комплекс, прогнозирование, многолетняя мерзлота.

Потепление климата, ставшее одной из наиболее значимых глобальных проблем современности, приводит на северных территориях к росту аварийности на трубопроводах и других сооружениях нефтегазового комплекса. Снижение прочности многолетнемерзлых пород, вызываемое ускорением термокарстовых процессов под воздействием потепления, сопровождается ростом экономических и экологических ущербов на предприятиях отечественного нефтегазового комплекса, так как большинство газовых месторождений и значительная часть месторождений нефти Западной Сибири располагаются в зоне вечной мерзлоты. Разработка мероприятий по снижению ущербов нефтегазодобывающих предприятий требует получения прогнозных оценок динамики термокарстовых процессов на территории вечной мерзлоты в условиях глобального потепления климата. В связи с этим возникает необходимость в разработке методов и средств математического моделирования динамики термокарстовых процессов, позволяющих путем проведения компьютерных экспериментов с моделью полей термокарстовых озер прогнозировать изменения их состояния с учетом климатических изменений.

Вопросы моделирования термокарстовых озер рассматривались в работах Веста и Плуга, Дюге, Субина и Райли, Хостетлера, и др., обзор которых приведен в [1], где показано, что использование аналитических моделей позволяет достаточно эффективно исследовать различные процессы в одиночных термокарстовых озерах. Однако такие модели не пригодны для исследования пространственно-временных изменений полей термокарстовых озер на территории многолетней мерзлоты. С точки зрения моделирования полей термокарстовых озер интерес вызывают разрабатываемые А.С. Викторовым [2, 3] методы математической морфологии ландшафта, ориентированные на использование аналитических моделей для моделирования динамики территорий. Методы и модели математической морфологии ландшафта могут использоваться для оценки геоэкологических рисков при развитии термокарстовых процессов на территориях в зоне мерзлоты. Однако эти методы также не предназначены для исследования пространственно-временных изменений полей термокарстовых озер в условиях изменяющегося климата.

Поэтому перспективным рассматривается использование имитационных моделей [4], позволяющих исследовать динамику полей термокарстовых озер в условиях современного глобального потепления. В последнее время в рамках идеологии имитационного моделирования сформировалось одно из новых направлений компьютерного моделирования – геоимитационное моделирование, понимаемое как имитационное моделирование сложных объектов с пространственной структурой и реализуемое с использованием методов и средств геоинформатики [5–7]. В [8] предложена имитационная модель поля термокарстовых озер в виде совокупности случайных окружностей, свойства которой учитывают основные характеристики реальных полей термокарстовых озер, определенные по экспериментальным данным дистанционных исследований. Однако вопросы программной реализации имитационного моделирования полей термокарстовых озер в настоящее время рассмотрены недостаточно, что и определило цель настоящей работы.

Реализация математической модели полей термокарстовых озер предполагает создание генератора псевдослучайных числовых последовательностей – троек: первые два псевдослучайных числа распределены по закону равномерной плотности, а третье – по показательному закону. В общем случае совместная плотность вероятностей представляется в виде

$$f(x, y, s), \quad (1)$$

где x и y – случайные координаты центров озер; s – площадь озер.

Разработана структурная схема программного комплекса имитационного моделирования (ПКИМ), которая представлена на рис. 1.

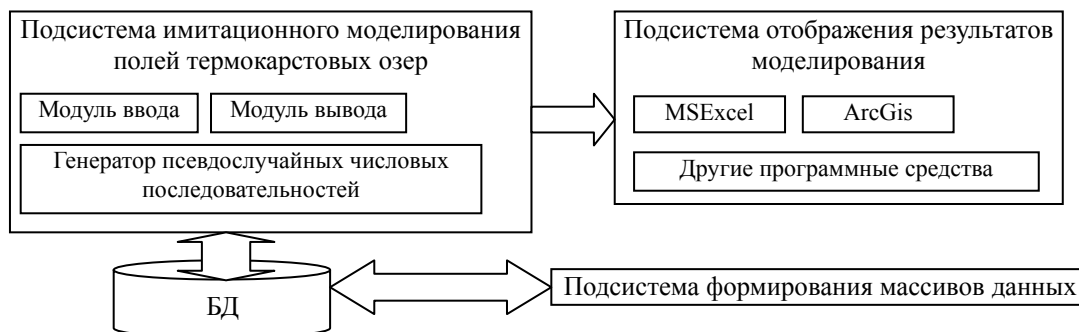


Рис. 1. Обобщенная схема программного комплекса имитационного моделирования полей термокарстовых озер

В состав разработанного программного комплекса входят следующие основные компоненты:

- 1) подсистема моделирования полей термокарстовых озер;
- 2) подсистема отображения результатов моделирования;
- 3) база данных (БД);
- 4) подсистема формирования массивов данных.

Ниже приведено описание компонентов ПКИМ. *Подсистема имитационного моделирования* полей термокарстовых озер, разработанная автором, представляет собой совокупность программных модулей, обеспечивающих ввод параметров модели, формирование псевдослучайных числовых последовательностей и вывод результатов моделирования. В состав подсистемы имитационного моделирования полей термокарстовых озер входят следующие основные блоки:

- *модуль ввода данных* (предназначен для обеспечения генератора псевдослучайных числовых последовательностей (ГПЧП) значениями параметров модели);
- *генератор псевдослучайных числовых последовательностей* (является основной составной частью подсистемы имитационного моделирования и предназначен для формирования случайных числовых последовательностей при реализации алгоритмов численного моделирования полей термокарстовых озер). Достаточно подробно алгоритм работы генератора описан в [9];
- *модуль вывода данных* (предназначен для преобразования результатов моделирования в один из следующих форматов: Microsoft Excel (*.xls), векторный формат (*.shp), растровый формат (*.jpeg)).

Подсистема отображения результатов моделирования позволяет отобразить выходные данные либо на цифровой карте средствами геоинформационной системы ArcGIS, либо в виде электронной таблицы и графика в MSExcel. *База данных* в составе ПКИМ представляет собой хранилище пространственной и атрибутивной информации об объектах исследования, полученной в ходе натурного эксперимента. Описание структуры и возможностей БД приводится в [9]. *Подсистема формирования массивов данных* позволяет извлекать из базы данных необходимую информацию об объекте исследования и создавать массивы данных для проведения модельных экспериментов.

Подсистема имитационного моделирования термокарстовых процессов, включенная в программный комплекс, зарегистрирована Роспатентом (свидетельство №2011614293 от 31 мая 2011 г.). Разработанная база данных также зарегистрирована Роспатентом (свидетельство № 2010620330 от 17 июня 2010 г.).

На рис. 2 приведена схема уровневое представление работы системы имитационного моделирования. На первом уровне выбирается совокупность параметров. На уровне алгоритмов численного моделирования в соответствии с параметрами, указанными на первом уровне, выбирается алго-

ритм численного моделирования. Третий уровень представляет результат работы алгоритма либо в виде массива данных, либо в виде графического изображения. Полученный массив данных обрабатывается средствами MSExcel, а сформированное изображение поля термокарстовых озер отображается на цифровой карте в геоинформационной системе ArcGIS.

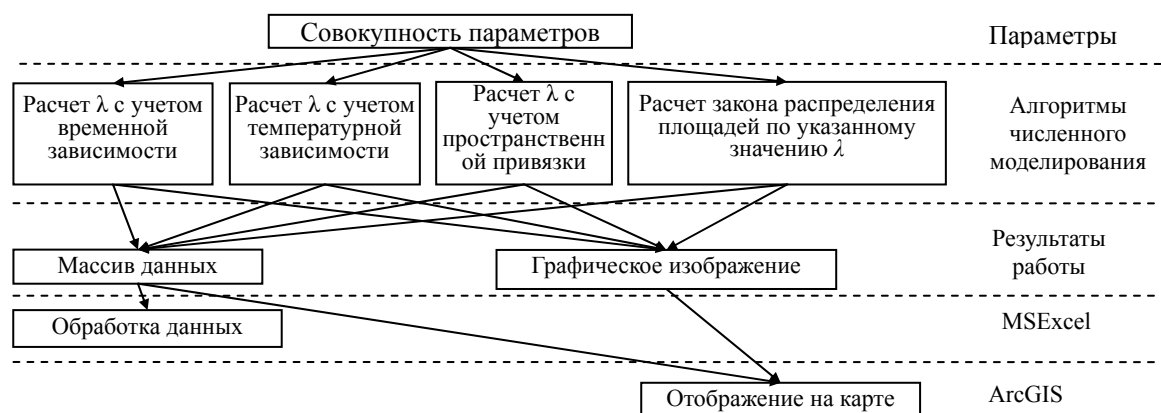


Рис. 2. Схема уровневой представления работы подсистемы имитационного моделирования

Для демонстрации работы программного комплекса проведено прогнозирование временных рядов средних значений площадей термокарстовых озер на исследованной территории многолетней мерзлоты Западной Сибири с учетом потепления климата на период до 2030 г., для чего сформированы сценарии модельных экспериментов для прогнозирования динамики полей термокарстовых озер. График прогнозного временного хода приведен на рис. 3.

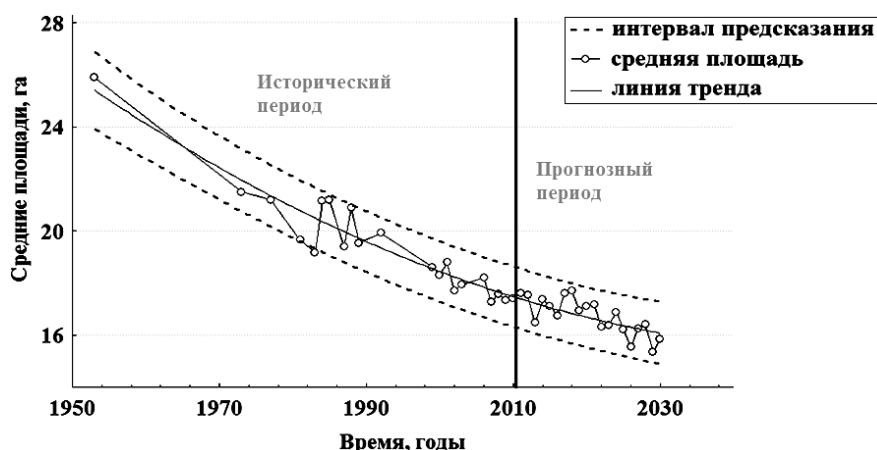


Рис. 3. Прогнозный временной ход средних площадей термокарстовых озер

Как видно из графика на рис. 3, в условиях продолжающегося потепления климата следует ожидать к 2030 г., что сокращение площадей термокарстовых озер на территории мерзлоты Западной Сибири в среднем может составить от 8 до 12% по сравнению с 2010 г.

Литература

1. Modeling thermokarst lakes dynamics and carbon flux [Электронный ресурс]. – Режим доступа: http://www.docstoc.com/docs/35351317/Methods_of_thermokarst_lakes_modeling, свободный (дата обращения: 02.09.2012).
2. Викторов А.С. Модели математической морфологии ландшафта в решении задач геоэкологии // Сергеевские чтения. Моделирование при решении геоэкологических задач: Матер. годичной сессии Научного совета РАН по проблемам геоэкологии, инженерной геологии и гидрогеологии (23–4 марта 2009 г.). – М.: ГЕОС, 2009. – Вып. 11. – С. 154–159.
3. Викторов А.С. Основные проблемы математической морфологии ландшафта. – М.: Наука, 2006. – 252 с.

4. Моисеев Н.Н. Системный анализ динамических процессов биосферы. Концептуальная модель биосферы / Н.Н. Моисеев, Ю.М. Свирежев // Вестник АН СССР. – 1979. – № 2. – С. 47–54.
5. Основы геоинформатики: в 2 кн.: учеб. пособие для вузов [Текст] / Е.Г. Капралов, А.В. Кошкарёв, В.С. Тикунов и др. ; под ред. В.С. Тикунова. – М. : Изд. центр «Академия», 2004. – 352 с.
6. Zhao Y. Urban dynamics analysis using spatial metrics geosimulation / Y. Zhao, Y. Murayama // Spatial analysis and modeling in geographical transformation process. – Dordrecht-Heidelberg-New York-London: Springer, The GeoJournal Library (Eds. Y. Murayama, R. Thapa), 2011. – Vol. 100. – P. 153–168.
7. Полищук Ю.М. Геоимитационное моделирование зон атмосферного загрязнения в результате сжигания газа на нефтяных месторождениях / Ю.М. Полищук, О.С. Токарева // Информационные системы и технологии. – 2010. – № 2. – С. 39–46.
8. Полищук Ю.М. Моделирование пространственно-временной динамики термокарста в зонах многолетней мерзлоты / Ю.М. Полищук, В.Ю. Полищук // Информационные системы и технологии. – 2011. – № 3. – С. 25–31.
9. Полищук В.Ю. Динамика полей термокарстовых озёр. Моделирование и прогноз. – Saarbrücken: LAP LAMBERT Academic Publishing GmbH & Co, 2012. – 126 с.

Полищук Владимир Юрьевич

Канд. техн. наук, мл. науч. сотрудник

Института мониторинга климатических и экологических систем СО РАН

Тел.: 8-953-914-15-38

Эл. почта: liquid_metal@mail.ru

Polishchuk V.Yu.

The software system of dynamics simulation of thermokarst lake fields in permafrost zones

The software complex for simulation of thermokarst lake fields is described. It consists of the following main components: simulation subsystem, display subsystem of simulation results and database. Forecast of dynamics of thermokarst lake fields was carried out in West-Siberian permafrost taking into account the increase of the average annual air temperature till 2030 year.

Keywords: simulation, software system, forecast, permafrost.

УДК 004.428.4

Ю.Б. Гриценко, О.И. Жуковский, И.В. Лазарев, М.М. Милихин, П.В. Сенченко

Интерфейс взаимодействия геоинформационной технологии ведения электронного генерального плана со сторонними программными системами

Рассмотрены возможности интеграции программного обеспечения информационной технологии ведения электронного генерального плана инженерной инфраструктуры со смежными информационными системами. Описаны методы кроссдоменного взаимодействия. Рассмотрен вариант использования интерфейса API для интеграции приложений.

Ключевые слова: геоинформационные системы, электронный генеральный план, интеграция информационных систем, API, документооборот.

Геоинформационная система ведения электронного генерального плана промышленного предприятия (ГИС ЭГП) представляет собой единую корпоративную информационную систему обработки геоданных пространственных и атрибутивных описаний объектов предприятия. Являясь распределенным Web-ориентированным программным продуктом, данная система нуждается в прикладном интерфейсе (API), служащем для ее интеграции с другими программными системами, используемыми на предприятии. Примером приложения, которому необходим такой интерфейс, может служить система электронного документооборота (СЭД) предприятия, для которой в результате интеграции с ГИС становятся возможными формирование и учет специфических документов, необходимых для работы с генеральным планом.

Другим способом взаимодействия между системами является обмен данными посредством файлов установленного формата. Потребность в таком способе возникает в том случае, когда стороннему программному продукту требуются только данные, которыми оперирует ГИС (без необходимости интеграции пользовательского интерфейса приложений), например, предоставление атрибутивных описаний объектов генерального плана, необходимых для выполнения функций анализа и восстановления данных. В настоящей статье рассмотрены основные этапы реализации интерфейса взаимодействия ГИС ЭГП со сторонними программными продуктами.

Программный интерфейс взаимодействия web-ориентированной геоинформационной системы со сторонними программными продуктами. Одним из требований к ГИС ЭГП является масштабируемость системы, в соответствии с которым ГИС состоит из отдельных независимых друг от друга программных модулей, отвечающих за выполнение определенных функциональных возможностей системы. Первоочередной задачей, возникающей при разработке внешнего API системы, является разработка модели взаимодействия, которая обеспечила бы возможность вызова упрощенной версии ГИС ЭГП в качестве виджета (элемента графического пользовательского интерфейса) стороннего приложения.

Трудность интеграции Web-ориентированных систем, клиентской частью которых выступает Web-браузер, определяется необходимостью организации кроссдоменного взаимодействия между такими системами.

Разрабатываемая ГИС ЭГП предполагает возможность установки и эксплуатации в сетях, управление которыми осуществляется на основе службы каталогов Active Directory корпорации Microsoft. Active Directory имеет иерархическую структуру, состоящую из объектов (ресурсов, служб и учетных записей пользователей и групп пользователей). Совокупность таких объектов формирует лес – верхний уровень организации инфраструктуры сети промышленного предприятия. Лес содержит одно или несколько деревьев, связанных отношениями доверия. Деревья, в свою очередь, включают в себя один или несколько доменов, которые идентифицируются своими пространствами имен DNS (Domain Name Service).

В соответствии с данной структурой сети предприятия различные Web-приложения могут функционировать в различных доменах. При этом возможности взаимодействия между такими системами определяются правилами ограничения домена (Same Origin Policy). Такие правила реализо-

ваны во всех современных Web-браузерах и не имеют определенного стандарта [1]. Основной целью их использования является ограничение доступа к функциям управления содержимым web-страниц и Web-приложений внутри домена со стороны недоверенных сайтов.

Таким образом, выбор модели взаимодействия определяется поддержкой правил ограничения домена всех современных Web-браузеров и обеспечением достаточного уровня безопасности такого взаимодействия.

Методы кроссдоменного взаимодействия. Существует несколько основных способов реализации кроссдоменного взаимодействия:

1. Использование технологии JSONP (Java Script Object Notation with Padding). В отличие от формата JSON, служащего для передачи данных в виде пар «ключ: значение» (объекта вида {ключ: значение, ключ2: значение2, ...}), данная технология предполагает возможность выполнения GET-запросов на основе использования HTML-тега (элемента разметки страницы HyperText Markup Language) <script>, выполняющего загрузку приложения на языке JavaScript и позволяющего передать в качестве параметра запроса имя функции, которую необходимо выполнить для обработки данных [2]. Одним из основных недостатков данной технологии является подверженность сайтов, использующих JSONP, подделке межсайтовых запросов со стороны вредоносных web-страниц, вследствие чего внедрение JSONP зачастую рассматривается как небезопасный способ кроссдоменного взаимодействия [3]. Кроме того, JSONP не позволяет передавать данные с использованием POST-запросов.

2. Использование технологии предоставления доступа к ресурсам CORS (Cross-origin resource sharing) [4]. Согласно данной технологии серверу интегрируемого Web-приложения необходимо указать список доверенных клиентских приложений с помощью параметра Access-Control-Allow-Origin (например, Access-Control-Allow-Origin: http://example.com http://another.example.com:8080). Использование данного метода более безопасно по сравнению с технологией JSONP и позволяет использовать обычные запросы XMLHttpRequest. Основным недостатком данного метода является отсутствие поддержки со стороны Web-браузеров Microsoft Internet Explorer 6 и 7 и лишь частичная поддержка в Internet Explorer 8.

3. Собственные методы управления кроссдоменными запросами для каждого Web-браузера. Основными недостатками способа является отсутствие поддержки со стороны браузеров Microsoft Internet Explorer, а также необходимость отдельной реализации API для каждого браузера, что значительно затрудняет процесс разработки и поддержания функционирования программного обеспечения.

4. Использование фреймов (HTML-тег <iframe>). Данный подход предполагает создание на странице родительского приложения окна iframe, которое может обслуживать GET-запросы посредством смены его URL-адреса или POST-запросы посредством установки свойства form.target формы с данными в соответствии с идентификатором окна iframe.

5. Использование функции PostMessage. Данный механизм доступен как часть стандарта HTML5 [3] и поддерживается всеми совместимыми с данным стандартом браузерами. Основным недостатком способа является отсутствие поддержки со стороны браузеров Microsoft Internet Explorer.

Создание интерфейса API. На основании представленного выше обзора возможных решений в качестве основного метода взаимодействия был выбран метод использования функции PostMessage стандарта HTML5. Такой подход не требует отдельной реализации для различных браузеров и является стандартом, рекомендованным консорциумом W3C (World Wide Web Consortium) [3]. В качестве альтернативы для клиентских приложений, не поддерживающих стандарт HTML5, возможно использование одного из трех методов:

- обмен данными с использованием технологий Adobe Flash;
- обмен данными с использованием специальных HTML-файлов на клиентской и серверной сторонах;
- обмен данными с использованием hash-части строки URL элемента iframe.

Такой подход позволяет достичь поддержки программного интерфейса всеми современными Web-браузерами, выбирая оптимальный вариант взаимодействия для каждого конкретного случая.

Основой взаимодействия служат объекты классов сокета (специальные объекты, инициализируемые на стороне клиентского и серверного приложений), предоставляемые библиотекой JavaScript инструментов EasyXDM [5]. Рассмотрим механизм взаимодействия более подробно (рис. 1).

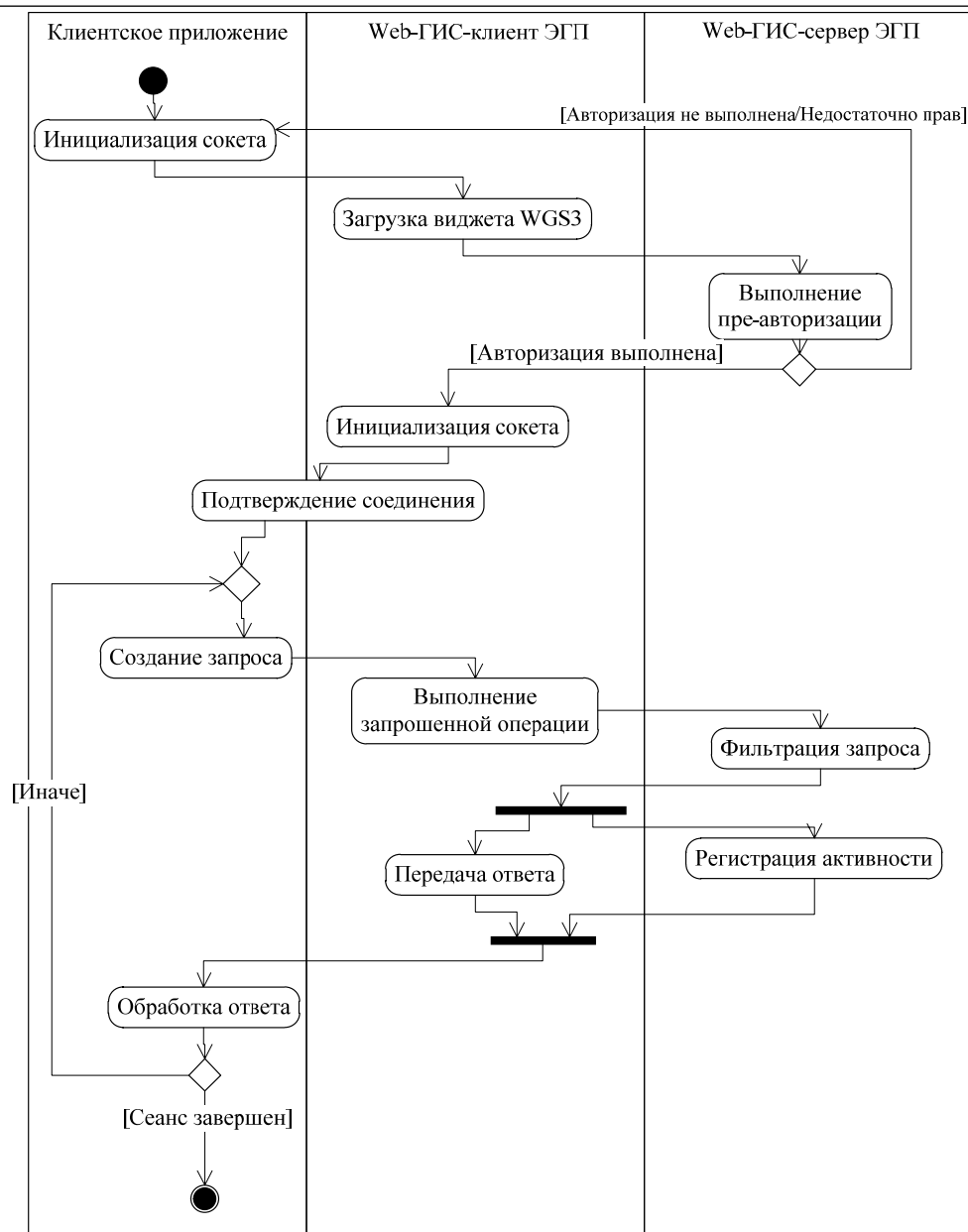


Рис. 1. Обобщенный алгоритм функционирования внешнего API ГИС ЭГП

Согласно представленному алгоритму соединение инициализируется внешней программной системой (клиентским приложением), которая создает сокет для установления соединения с ГИС ЭГП. Основным параметром сокета является адрес URL (Universal Resource Locator) программного обеспечения Web-ГИС-клиента ЭГП, которое будет использовано в качестве виджета. В качестве параметров GET-запроса данного адреса указываются имя и хэш-функция пароля учетной записи пользователя, которые необходимы для выполнения авторизации в системе, например:

`http://gis/mapwidget?user={USER_NAME}&auth={SECRET_KEY}`, где `SECRET_KEY = md5({IP-адрес}. {USER_PASSWORD})`.

Таким образом, обеспечивается безопасность передачи учетных данных пользователя по недоверенной сети.

После выполнения инициализации сокета на стороне внешней системы создается экземпляр облегченной версии клиентского Web-приложения ГИС ЭГП, реализующей обработку запросов через сокет. Выполняется подтверждение установки соединения, после чего возможно взаимодействие систем посредством передачи сообщений.

Обмен сообщениями. Разработанный программный интерфейс предполагает обработку двух основных видов запросов:

- запросы на выделение области территории генерального плана;
- запросы на выделение отдельных объектов генерального плана.

Каждый из данных типов запросов, в свою очередь, содержит запросы на получение и установку текущих параметров отображения картографического описания генплана.

Запрос представляется в виде сообщения фиксированной структуры. Такое сообщение является строкой в формате нотации JSON (строковым представлением объекта JavaScript).

Сообщение имеет следующий формат: {type: message_type, content: message_content}, где message_type – числовая константа, определяющая тип запроса, message_content – параметры запроса. Параметры запроса включают в себя контекст карты (список активных слоев, масштаб и центральную точку), а также данные, позволяющие идентифицировать выделенную область генплана или его отдельный объект. Система поддерживает четыре значения message_type:

- 1) получение выделенной области генерального плана и его текущего контекста;
- 2) установка выделенной области генерального плана и его текущего контекста;
- 3) получение идентификатора выделенного объекта и текущего контекста карты;
- 4) установка идентификатора выделенного объекта и текущего контекста карты.

Передача контекста карты. Контекст карты служит для однозначного определения видимой области территории генерального плана и списка активных слоев карты. Область территории однозначно определяется посредством указания координат центральной точки карты и ее масштаба. Для указания списка слоев используется циклический избыточный код 32-битных полиномов (сгс32). Каждый слой идентифицируется контрольной суммой сгс32 его основных характеристик (наименования, группы слоев, расположения в репозитории геоданных). Такой подход позволяет избежать зависимости функционирования API от текущих характеристик базы данных пространственных описаний и значений идентификаторов пространственных слоев, которые могут изменяться с течением времени.

Таким образом, контекст картографического объекта представляет собой строку следующего вида: <координата X центра карты>,< координата Y центра карты > | <масштаб карты> | <контрольная сумма активного слоя 1>,<контрольная сумма активного слоя 2>,...,<контрольная сумма активного слоя n>.

Дополнительно к перечисленным характеристикам контекст может также включать в себя описание выделенной области: «|<тип выделенной области>| <характеристика области 1>, <характеристика области 2>, ..., <характеристика области t>», где <тип выделенной области> может принимать два значения: «radius» – выделение окружностью и «polygon» – выделение полигональным объектом. Характеристиками выделенной области являются координаты по оси абсцисс, и ординат точки центра окружности и радиус для выделения окружностью и координаты вершин многоугольника для полигонального выделения.

Использование API ГИС ЭГП при интеграции со сторонними приложениями. В качестве примера использования API геоинформационной системы используем задачу разработки программного обеспечения организации документооборота электронного генерального плана промышленного предприятия.

В общем виде процесс организации документооборота включает в себя все действия по обработке входящей, исходящей и внутренней корреспонденции: прием и первичная обработка входящих документов, составление, согласование и утверждение проектов документов, формируемых внутри организации, формирование поручений и контроль их исполнения, отправка исходящих документов. Особенности процессов документооборота зависят от характеристик автоматизируемых бизнес-процессов. Так, при автоматизации процессов по управлению инженерной инфраструктурой предприятия СЭД должна решать следующие характерные для данной предметной области задачи:

- обеспечение регистрации всех типов документов, находящихся в общем потоке документооборота ЭГП;
- ведение электронного архива документов ЭГП;
- организация взаимодействия с другими профильными подсистемами программного комплекса ведения ЭГП;
- организация автоматизированного контроля исполнения документов;
- ведение технологии электронного взаимодействия между подразделениями организации (пользователями ЭГП);

- мониторинг документа – определение стадии, на которой находится рассмотрение того или иного документа;
- связь между документами различного уровня исполнения;
- хранение информации обо всех документах, обеспечив возможность доступа к любой стадии работы с документами;
- ведение системы отчетности по исполнению документов и исполнительской дисциплине сотрудников организации.

Решение обозначенных задач требует наличия информационной связи между объектами инженерной инфраструктуры предприятия и документами, циркулирующими в общем потоке документооборота. При этом можно выделить два вида связей: 1) связь документа с конкретным объектом генерального плана, при этом объект характеризуется уникальным идентификатором в системе ЭГП; 2) связь документа с некоторой пространственной областью генерального плана, которая характеризуется набором точек (полигоном). Для отражения данных связей в СЭД может быть использована дополнительная таблица базы данных со структурой:

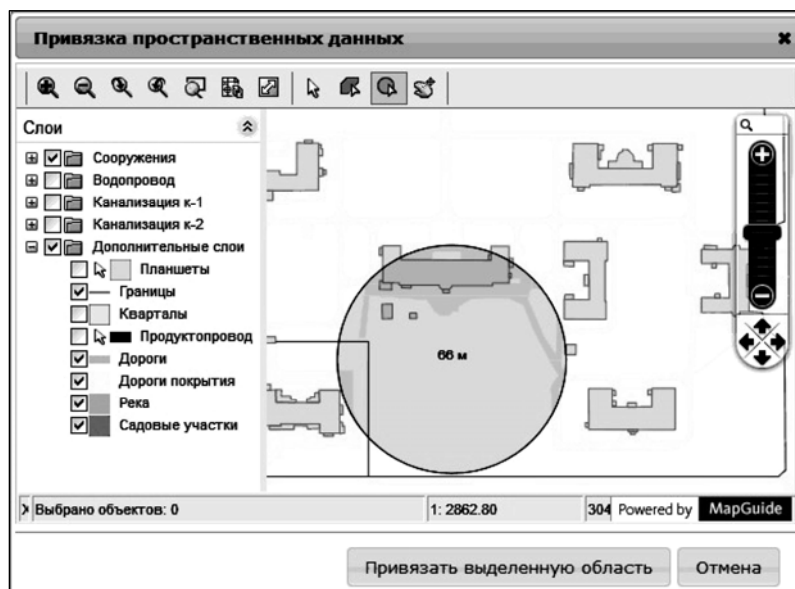
Структура таблицы связи документов СЭД и пространственных данных ЭГП

PK*	Название поля	Тип данных	Комментарий
+	ID	NUMBER(10)	Уникальный идентификатор связи
	LinkType	NUMBER(2)	Код типа связи
	DocumentID	NUMBER(10)	Уникальный идентификатор документа
	GIS_ObjectID	NUMBER(10)	Уникальный идентификатор объекта ЭГП
	GIS_Area	BLOB	Описание области генерального плана
	GIS_Context	BLOB	Контекст ЭГП
	AddedAt	DATE	Время создания связи

Примечание. PK* – Primary Key (первичный ключ таблицы).

Необходимо отметить наличие информации о контексте электронной карты в приведенной структуре. Информация о контексте не является характеристикой связи между документом и объектом ЭГП, однако наличие данной информации обеспечивает однозначность определения видимой области генерального плана при отображении электронной карты. С точки зрения интерфейса пользователя системы электронного документооборота связь между документом и объектом генерального плана создается с использованием виджета ГИС ЭГП, демонстрация которого представлена на рис. 2.

Рис. 2. Привязка пространственных данных к документу в системе электронного документооборота



Для интеграции данного пользовательского элемента в интерфейс СЭД создан специальный класс `GEOSelectorControl`, который инкапсулирует в себе логику взаимодействия с API ГИС ЭГП и отображение электронной карты в системе электронного документооборота.

В качестве примера связи пространственных данных и данных документационного обеспечения можно привести процесс получения выкопировки генерального плана. При регистрации заявки на выкопировку в СЭД пользователю необходимо указать информацию об участке территории для формирования выкопировки. Другим примером является добавление нового пространственного объекта на электронный генеральный план и регистрации в СЭД соответствующего документооснования.

Заключение. Результатами проделанной работы являются разработанные в рамках проекта W4GIS модель взаимодействия и программный модуль для интеграции ГИС ведения электронного генерального плана в пользовательский интерфейс сторонних Web-приложений. Реализованная технология взаимодействия опирается на стандарты консорциума W3C и поддерживается всеми современными Web-браузерами.

Выполнение данной работы проводилось при финансовой поддержке Министерства образования и науки Российской Федерации в рамках мероприятия 2.4 федеральной целевой программы «Исследования и разработки по приоритетным направлениям развития научно-технологического комплекса России на 2007–2013 годы», проект «Разработка Web-ориентированных геоинформационных технологий формирования и мониторинга электронного генерального плана инженерной инфраструктуры», государственный контракт № 07.524.11.4013 от 03 ноября 2011 г.

Литература

1. Same Origin Policy [Электронный ресурс]. – Режим доступа: http://www.w3.org/Security/wiki/Same_Origin_Policy, свободный (дата обращения: 21.11.2012).
2. RIAspot JSON P for Cross Site XHR [Электронный ресурс]. – Режим доступа: <http://www.riaspot.com/blogs/entry/JSONP-for-Cross-Site-XHR>, закрытый (дата обращения: 21.09.2012).
3. Json-p.org. The problem [Электронный ресурс]. – Режим доступа: <http://json-p.org>, свободный (дата обращения: 19.01.2013).
4. Cross-Origin Resource Sharing [Электронный ресурс]. – Режим доступа: <http://www.w3.org/TR/cors/>, свободный (дата обращения: 18.01.2013).
5. EasyXDM. Неме [Электронный ресурс]. – Режим доступа: <http://easyxdm.net>, свободный (дата обращения: 19.01.2013).

Гриценко Юрий Борисович

Канд. техн. наук, доцент каф. автоматизации обработки информации (АОИ) ТУСУРа

Тел.: 8(382-2) 90-01-80

Эл. почта: ubg@muma.tusur.ru

Жуковский Олег Игоревич

Канд. техн. наук, доцент каф. АОИ ТУСУРа

Тел.: 8(382-2) 90-01-80

Эл. почта: ol@muma.tusur.ru

Лазарев Иван Васильевич

Мл. науч. сотрудник каф. АОИ ТУСУРа

Тел.: +7-913-812-46-79

Эл. почта: liv@ms.tusur.ru

Милихин Михаил Михайлович

Студент факультета систем управления ТУСУРа

Тел.: +7-952-807-90-51

Эл. почта: milikhin@gmail.com

Сенченко Павел Васильевич

Канд. техн. наук, доцент каф. АОИ, декан факультета систем управления ТУСУРа

Тел.: 8 (383-2) 70-15-46

Эл. почта: pvs@tusur.ru

Gritsehko Yu.B., Zhukovskiy O.I., Lazarev I.V., Milikhin M.M., Senchenko P.V.

Interface between GIS technology of electronic master plan with other software systems

We investigate the possibilities of software integration of electronic information technology of a master plan for physical infrastructure with related information systems. The methods of cross domain interaction are described. We considered a variant using an API for application integration.

Keywords: geographic information systems, electronic master plan, integration of information systems, API, workflow.

УДК 004.8

И.А. Ходашинский, И.В. Горбунов, Д.С. Синьков

Алгоритмы генерации структур двухкритериальных Парето-оптимальных нечетких аппроксиматоров

Предложены алгоритмы генерации структуры нечетких аппроксиматоров, оптимизированных по следующим двум критериям: точность и сложность. Приведены результаты исследований полученных аппроксиматоров на реальных данных из репозитория KEEL, проведено сравнение полученных результатов с аналогами.

Ключевые слова: нечеткий аппроксиматор, Парето-оптимальность, генерация структуры, метаэвристики.

Традиционно разработка нечетких систем, основанных на правилах, направлена на оптимизацию показателей качества, например, максимизацию точности в задачах классификации или минимизацию ошибки в задачах аппроксимации. Указанные цели часто достигаются в ущерб понятности проектируемой нечеткой системы (НС). Решение проблемы нахождения компромисса между точностью и понятностью является предметом исследования настоящей работы.

Если в НС каждая входная переменная представлена небольшим числом нечетких термов, каждому терму можно придать осмысленное лингвистическое значение, например: «очень малое», «малое», «среднее», «большое», «очень большое». Кроме того, когда уменьшается число правил и число условий в правилах, интерпретируемость НС улучшается. Отметим также, что несложная и хорошо интерпретируемая НС более проста в настройке, требует меньше памяти и времени вывода, чем более сложная НС.

Не существует универсального способа измерения сложности или интерпретируемости моделей [1–3]. В нашей работе сложность определена как сумма числа правил и числа термов нечеткой системы. Кроме того, на нечеткую систему наложены следующие ограничения:

- число термов, которыми описывается каждая входная переменная, находится в разумных пределах (как правило, от 2 до 9, за исключением НС с одним входом);
- функции принадлежности (ФП) нечетких термов выпуклы и нормализованы, т.е. каждая ФП имеет значение, равное единице по крайней мере в одной точке в области определения;
- область определения полностью покрыта функциями принадлежности, т.е. по крайней мере, одна ФП получает значение, не равное нулю в любой точке области определения;
- ФП различимы, т.е. две ФП не принимают очень близких значений на области определения;
- в работе используются глобально определенные функции принадлежности, это означает, что определенные один раз функции принадлежности используются во всех правилах;
- в базе нет правил, у которых одинаковые antecedentes, но различные консеквенты.

Поскольку точность и сложность являются противоречивыми критериями, генерируется не одна оптимальная НС, а набор таких систем. Из сгенерированного набора выбирается множество недоминируемых решений, обозначенное как Парето-множество, характерным свойством которого является оптимальное соотношение между критериями точности и сложности.

Постановка задачи. Нечеткий аппроксиматор задается своей базой правил; i -е правило имеет следующий вид:

$$\text{ЕСЛИ } x_1=A_{1i} \text{ И } x_2=A_{2i} \text{ И } \dots \text{ И } x_n=A_{ni} \text{ ТО } y=r_i,$$

где A_{ij} – лингвистический терм, которым оценивается переменная x_i ; r_i – действительное число, которым оценивается выход y .

Нечеткий аппроксиматор осуществляет отображение $f: \mathfrak{R}^n \rightarrow \mathfrak{R}$:

$$f(\mathbf{x}) = \frac{\sum_{i=1}^R \mu_{A_{1i}}(x_1) \cdot \mu_{A_{2i}}(x_2) \cdot \dots \cdot \mu_{A_{ni}}(x_n) \cdot r_i}{\sum_{i=1}^R \mu_{A_{1i}}(x_1) \cdot \mu_{A_{2i}}(x_2) \cdot \dots \cdot \mu_{A_{ni}}(x_n)},$$

где \mathbf{x} – входной вектор, R – число правил; n – число входных переменных; μ_{Aij} – функция принадлежности, определяемая набором своих параметров, например, треугольная – тремя параметрами, трапецевидная – четырьмя, гауссова и параболическая – двумя.

Нечеткий аппроксиматор может быть представлен как

$$y = f(\mathbf{x}, \boldsymbol{\theta}),$$

где y – скалярный выход аппроксиматора; $\boldsymbol{\theta} = \|\theta_1, \dots, \theta_M\|$ – вектор параметров; $N = \sum_{i=1}^n a \times b_i$; a – чис-

ло. Пусть дано множество обучающих данных (таблица наблюдений) $\{(\mathbf{x}_p, t_p), p = 1, \dots, m\}$, тогда среднеквадратическая функция ошибки ($RMSE$), являющаяся численным критерием адекватности модели, вычисляется по следующей формуле:

$$E(\boldsymbol{\theta}) = \frac{1}{m} \sqrt{\sum_{p=1}^m (t_p - f(\mathbf{x}_p, \boldsymbol{\theta}))^2}.$$

Проблема построения Парето-оптимального множества нечетких аппроксиматоров с разными соотношениями между значениями их точности и сложности основано на оптимизации заданной функции в многомерном пространстве, координаты которого соответствуют параметрам $\boldsymbol{\theta}$ нечеткого аппроксиматора. Схематично технология построения Парето-оптимального множества нечетких аппроксиматоров приведена на рис. 1.

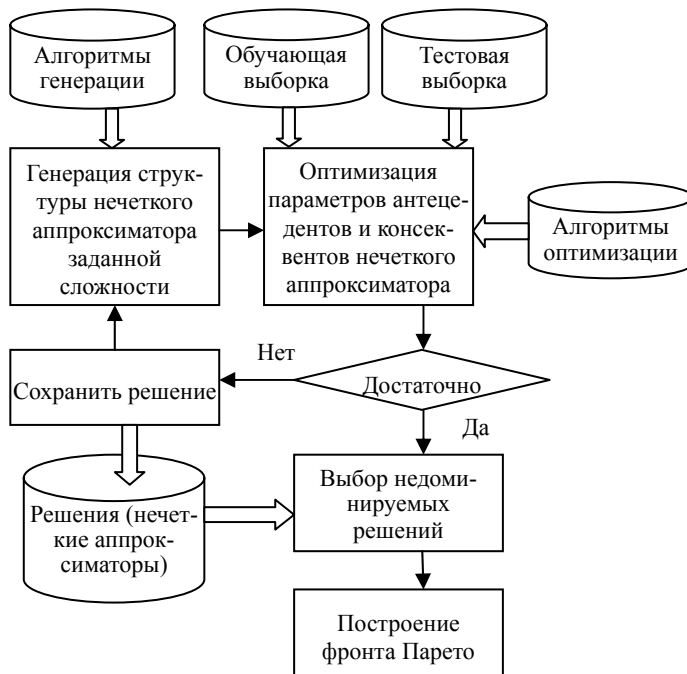


Рис. 1. Схема построения Парето-оптимального множества нечетких аппроксиматоров

алгоритмы муравьиной колонии [6–9]; 3) алгоритм пчелиной колонии [10, 11]; 4) алгоритм перемещения бактерий [12, 13]. Для оптимизации консеквентов правил применяется метод наименьших квадратов [14].

Оптимизация консеквентов правил в предлагаемой технологии выполняется следующими методами: 1) методом поиска ближайшего соседа в таблице наблюдений; 2) адаптированным методом наименьших квадратов.

Конечный пользователь может выбрать наиболее подходящее решение из окончательного фронта Парето путем поиска компромисса между сложностью и точностью в зависимости от своих предпочтений.

Сильные стороны предлагаемого подхода:

1) пользователь заранее может определить желаемый уровень сложности, все получаемые нечеткие системы будут отвечать данному требованию;

2) поскольку сложность не является оптимизируемым параметром, это позволяет использовать алгоритмы однокритериальной оптимизации.

Недостаток данного подхода заключается в его ориентации только на дискретное описание критерия сложности нечеткой системы.

Алгоритмы генерации структуры нечеткого аппроксиматора

Алгоритм генерации структуры на основе Gath-Geva метода кластеризации

Рассмотрим алгоритм генерации структуры на основе модифицированного Gath-Geva метода кластеризации [15]. Кластеры, полученные применением алгоритма, преобразуются в нечеткие термы, количество таких термов, относящихся к каждой переменной, фиксировано и равно количеству кластеров. Собственно алгоритм приведен ниже.

Вход: Таблица наблюдений $\{x_p, t_p\}$, количество входных параметров l , количество кластеров c , экспоненциальный вес m , требуемая точность работы алгоритма кластеризации ε .

Выход: θ – база правил аппроксиматора.

\mathbf{u} – матрица разбиения содержит c – строк, n – колонок, ограничения: $\sum_{i=1}^c \mathbf{u}_{ie} = 1, 0 < \sum_{e=1}^n \mathbf{u}_{ie} < n$;

\mathbf{v}_i – вектор координат центра i -го кластера;

\mathbf{d} – матрица расстояний между кластерами и элементами выборки содержит c – строк, n – колонок;

\mathbf{A}_i – матрица размером $l \times l$ нечеткой ковариации между i -м кластером и выборкой \mathbf{x} ;

\mathbf{S} – матрица дисперсий элементов таблицы наблюдения относительно центров векторов по каждому из входных параметров. Матрица содержит строк по количеству кластеров c , столбцов по количеству входных параметров l .

Шаг 1. Произвольное заполнение матрицы разбиения \mathbf{u} , $\sum_{e=1}^n \mathbf{u}_{ie} = 1$.

Шаг 2. Расчет векторов центров кластеров \mathbf{v}_i , $\mathbf{v}_i = \frac{\sum_{e=1}^n \mathbf{u}_{ie}^m x_e}{\sum_{e=1}^n \mathbf{u}_{ie}^m}$.

Шаг 3. Вычисление расстояний между кластером и каждым элементом выборки

$$\mathbf{A}_i = \frac{\sum_{e=1}^n \mathbf{u}_{ie}^m (\mathbf{x}_e - \mathbf{v}_j)(\mathbf{x}_e - \mathbf{v}_j)^T}{\sum_{e=1}^n \mathbf{u}_{ie}^m}, \quad p_i = \frac{1}{p} \left(\sum_{e=1}^n \mathbf{u}_{ie} \right), \quad \mathbf{d}_{ie}^2 = \frac{\det(\mathbf{A}_i)^{\frac{1}{2}}}{p_i} \exp\left(\frac{1}{2}(\mathbf{v}_i - \mathbf{x}_e)^T (\mathbf{A}_i)^{-1} (\mathbf{v}_i - \mathbf{x}_e)\right).$$

Шаг 4. Расчет новой матрицы разбиения;

Шаг 4.1. ЕСЛИ $\mathbf{d}_{ie} = 0$ ТО $\mathbf{u}_{ie}^* = 1$, переход на шаг 5.

Шаг 4.2. ЕСЛИ $\mathbf{d}_{ie} < 0$ ТО $\mathbf{u}_{ie}^* = 0$, переход на шаг 5.

Шаг 4.3. ЕСЛИ $\mathbf{d}_{ie} > 0$ ТО $\mathbf{u}_{ie}^* = \left[\sum_{k=1}^c \left(\frac{\mathbf{d}_{ie}}{\mathbf{d}_{ke}} \right)^{\frac{2}{m-1}} \right]^{-1}$, переход на шаг 5.

Шаг 5. ЕСЛИ $\|\mathbf{u}^* - \mathbf{u}\| < \varepsilon$ ТО переход на шаг 7, ИНАЧЕ переход на шаг 6.

Шаг 6. $\mathbf{u} = \mathbf{u}^*$ переход на шаг 2.

Шаг 7. Расчет размеров проекции кластера для определения параметров ФП

$$S_{ij} = \frac{\sum_{k=1}^n (\mathbf{u}_{ki})^m \cdot (x_{kj} - v_{ij})^2}{\sum_{k=1}^n (\mathbf{u}_{ki})^m}.$$

Шаг 8. Построение функций принадлежности A_{ij} , соответствующих каждому нечеткому терму:

$$A_{ij} = \text{Gauss}(\mathbf{v}_{ij}, \sqrt{S_{ij}}).$$

Шаг 9. Формирование нечетких правил вида

R_i : ЕСЛИ $x_1=A_{i1}$ И $x_2=A_{i2}$ И $x_3=A_{i3}$ И ... И $x_n=A_{in}$ ТО $r_i = none$,

описывающих каждый кластер.

Шаг 10. Инициализация консеквента r_i правила R_i методом поиска ближайшего соседа в таблице наблюдений.

Алгоритм генерации базы правил нечеткой системы равномерным разбиением и перебором

Вход: таблица наблюдений $\{x_p, t_p\}$, $Type$ – тип функции принадлежности, вектор **Count_terms** с указанием количества термов, требуемых по каждому признаку.

Выход: θ – начальная база правил аппроксиматора.

Алгоритм:

Шаг 1. Для каждого i -го входного параметра **делать**:

Шаг 1.1. Рассчитать длину одного терма как

$distance = \max(x_{pi}) - \min(x_{pi}) / (Count_terms_i - 1)$.

Шаг 1.2. Задать базовую точку построения терма $term_base = \min(x_{pi})$.

Шаг 1.3. Пока $term_base \leq \max(x_{pi})$.

Шаг 1.3.1. Создание терма A_{iq} типа $Type$, накрывающего интервал $[term_base - distance, term_base + distance]$.

Шаг 1.3.2. $term_base = term_base + distance$.

End Пока;

End делать (i).

Шаг 2. Формирование базы правил.

Шаг 2.1. Создание правил путем перебора термов каждого с каждым по всем входным параметрам i

R_q : ЕСЛИ $x_1=A_{1q}$ И $x_2=A_{2q}$ И $x_3=A_{3q}$ И ... И $x_n=A_{nq}$ ТО $r_q = none$.

Шаг 2.2. Инициализация консеквента r_q правила R_q методом ближайшего соседа;

Шаг 2.3. $\theta := \theta \cup R_q$.

Достоинство алгоритма заключается в простоте его реализации, полученная база правил хорошо поддается оптимизации.

Недостатком алгоритма является подверженность его эффекту «проклятия размерности».

Алгоритм генерации базы правил нечеткой системы исключением неэффективных правил

Вход: таблица наблюдений $\{x_p, t_p\}$, $Shrink_rule$ – количество правил, которые необходимо исключить, $Type$ – тип функции принадлежности, **Count_terms** – вектор с указанием количества термов для каждого признака.

Выход: θ – усеченная база правил аппроксиматора.

Шаг 1. Построить базу правил θ алгоритмом генерации базы правил нечеткой системы равномерным разбиением и перебором ($\{x_p, t_p\}$, $Type$, **Count_terms**).

Шаг 2. Сгенерировать единичный вектор **Init_rule_struct** размером, равным количеству правил в базе правил θ

$Init_rule_struct_i = 1, i = 1..|\theta.R|$.

Шаг 3. Заменить $Shrink_rule$ единиц в начале вектора **Init_rule_struct** на нули.

Шаг 4. Сгенерировать все различные варианты перестановок с повторением **Case_of_rule_struct** из вектора **Init_rule_struct** методом лексикографического генератора.

Шаг 5. Скопировать θ в θ' ;

Шаг 6. Для всех j -х вариантов структуры правил в

Case_of_rule_struct **делать**:

Шаг 6.1. Для всех k -х элементов структуры

Case_of_rule_struct_j **делать**:

Шаг 6.1.1. ЕСЛИ **Case_of_rule_struct_{jk}** = 0 ТО

$\theta'.R = \theta'.R \setminus \theta'.R_k$;

End делать (k).

Шаг 6.2. Для всех l -х термов в θ' **Do**:

Шаг 6.2.1. ЕСЛИ $\theta'.term_l \notin \theta'.R_k, k = 1..|\theta'.R|$ ТО $\theta'.term = \theta'.term \setminus \theta'.term_l$;

End делать (l).

Шаг 6.3. $\theta' =$

Метод геометрической коррекции параметров $\PhiП(\{x_p, t_p\}, \theta')$;

Шаг 6.4. **Case_FS_j** = θ' ;

Шаг 6.5. Рассчитать ошибку для каждого варианта базы правил Case_FS_j
 $\text{Error_of_FS}_j = \text{RMSE}(\{x_p, t_p\}, \text{Case_FS}_j)$;

End делать (j).

Шаг 7. Вернуть такой Case_FS_p что $p = \arg \min_p (\text{Error_of_FS}_p)$.

Алгоритм формирует нечеткие системы с малой ошибкой при невысокой сложности, но при больших сложностях системы алгоритм практически бесполезен.

Алгоритм генерации базы правил с заданной структурой

Вход: таблица наблюдений $\{x_p, t_p\}$, *Type* – тип функции принадлежности, *Count_terms* – вектор с указанием количества термов для каждого признака.

Выход: θ – начальная база правил аппроксиматора.

Шаг 1. Отсортировать *Count_terms* в лексикографическом порядке по возрастанию.

Шаг 2. Сгенерировать все различные варианты перестановок с повторением *Case_of_struct* из вектора *Count_terms* методом лексикографического генератора;

Шаг 3. Для j -го варианта перестановки *Case_of_struct* делать:

Шаг 3.1. Сгенерировать варианты начальной базы правил

$\text{Case_of_FS}_j = \text{алгоритм генерации базы правил нечеткой системы перебором}(\{x_p, t_p\}; \text{Type}; \text{Case_of_struct}_j)$.

Шаг 3.2. $\text{Error_of_FS}_j = \text{RMSE}(\{x_p, t_p\}, \text{Case_of_FS}_j)$.

End делать (j);

Шаг 4. Вернуть такой Case_of_FS_p , что $p = \arg \min_p (\text{Error_of_FS}_p)$.

Алгоритм генерирует аппроксиматор с ошибкой, сравнимой с алгоритмом генерации базы правил нечеткой системы перебором, при меньшем количестве нечетких термов и правил в базе. В меньшей степени, по сравнению с алгоритмом генерации базы правил нечеткой системы перебором, но подвержен влиянию эффекта «проклятия размерности».

Алгоритм генерации базы правил делением

Суть алгоритма заключается в следующем. На начальном этапе генерируется терм, накрывающий все пространство определения входной переменной, и формируются правила из этого набора термов. Данный шаг позволяет отказаться от методов коррекции базы правил, потому что не менее одного правила будет использовано для работы аппроксиматора. Далее используется алгоритм деления пополам. Этот этап повторяется, пока не будет сгенерировано нужное количество правил. Рекурсивное описание алгоритма представлено ниже.

Вход: таблица наблюдений $\{x_p, t_p\}$, *min_input* – вектор минимумов входных параметров, *max_input* – вектор максимумов входных параметров, *Type* – тип функции принадлежности, база правил аппроксиматора θ , требуемое количество правил *need_rules*.

Выход: θ – база правил аппроксиматора.

count_rules – счетчик правил в базе правил θ .

Шаг 1. **ЕСЛИ** *count_rules* \leq *need_rules*, **ТО** шаг 2, **ИНАЧЕ** **Выход**.

Шаг 2. Для каждого i -го входного параметра создать терм A_{iq} типа *Type*, накрывающего интервал $[\text{min_input}_i, \text{max_input}_i]$.

Шаг 3. Создать правило

R_q : **ЕСЛИ** $x_1 = A_{1q}$ **И** $x_2 = A_{2q}$ **И** $x_3 = A_{3q}$ **И** ... **И** $x_n = A_{nq}$ **ТО** $r_q = \text{none}$.

Шаг 4. Инициализация консеквента r_q правила R_q методом *ближайшего соседа*($\{x_p, t_p\}, R_q$).

Шаг 5. $\theta := \theta \cup \{R_q\}$.

Шаг 6. Задать вектор *center_input*, таким образом, что его элементы равны среднему арифметическому соответствующих элементов *min_input* и *max_input*

$\text{center_input}_i = (\text{min_input}_i + \text{max_input}_i) / 2, i = 1..|\text{min_input}|$.

Шаг 7. Вызов алгоритма *генерации_базы_правил_делением*(*min_input*; *center_input*; *Type*; θ ; *need_rules*).

Шаг 8. Вызов алгоритма *генерации_базы_правил_делением*(*center_input*; *max_input*; *Type*; θ ; *need_rules*).

Сравнение с аналогами. Были проведены исследования различных вариантов использования алгоритмов генерации и оптимизации. Каждый из алгоритмов генерации работал с алгоритмом оптимизации. Исследование алгоритмов проводилось при решении задач аппроксимации идеальных

данных и данных, описывающих реальные процессы, представленных в репозитории KEEL (Knowledge Extraction Evolutionary Learning, <http://www.keel.es>). Характеристики данных представлены в табл. 1. Все входные и выходные переменные – вещественные числа. Каждая выборка разделена на пять наборов, из которых строится обучающая и тестовая выборки, содержащие 80 и 20% данных соответственно. Разделение проводилось таким образом, чтобы каждый набор попал во все тестовые выборки ровно один раз.

Таблица 1

Описание данных			
Название данных	Количество образцов	Количество входных переменных	Краткое описание
<i>Diabetes</i>	43	2	Прогнозирование развития сахарного диабета у инсулинозависимых детей
<i>ELE2</i>	1066	4	Проблема оценки стоимости обслуживания городских электрических сетей

В работах [17, 18] ошибка представлена как $MSE/2$, поэтому в таблице сравнений оригинальное значение умножено на 2. Из таблицы видно, что оптимизация нашими алгоритмами дает лучшие результаты на всех выборках, кроме тестовой выборки по набору данных *Diabetes*.

Заключение. Интерпретируемость и транспарентность нечетких систем являются их важнейшим преимуществом в интерактивных приложениях. Для обеспечения этих свойств в работе предложены алгоритмы генерации, накладывающие определенные ограничения на проектируемые нечеткие системы. Алгоритмы позволяют находить компромисс между сложностью, выраженной через число нечетких правил и нечетких термов, и точностью модели. Сравнительный анализ разработанных алгоритмов с аналогами показал их высокую эффективность.

Таблица 2

Сравнение эффективности алгоритмов						
Алгоритм	Правила		Обучающая выборка		Тестовая выборка	
	Число	СКО	<i>MSE</i>	СКО	<i>MSE</i>	СКО
<i>Diabetes</i>						
Наши алгоритмы	25	0	0,02502	0,0116	0,5963	0,1918
Wang–Mendel [16]	18,6	1,4	0,2284	0,0425	1,4024	0,6890
COR-BWAS [16]	18,6	1,4	0,1750	0,0250	1,4587	0,7091
Thrift [16]	46,2	0,7	0,0745	0,0098	0,8783	0,3575
Pittsburgh [16]	15	2,9	0,1040	0,0182	0,9509	0,7881
Fuzzy-GAP [16]	10	0	0,1429	0,0376	0,5014	0,3014
Pitts-DNF min [16]	1,6	0,5	0,4162	0,1231	0,4539	0,1288
Pitts-DNF med [16]	5,4	0,5	0,1296	0,0136	0,3213	0,1922
Pitts-DNF max [16]	9,6	1,2	0,1066	0,0150	0,6340	0,5276
<i>ELE 2</i>						
Наши алгоритмы	37,3	12,3	13510	2611	17235	2420
Wang–Mendel [16]	65	0	112270	1498	112718	4685
COR-BWAS [16]	65	0	102664	1080	102740	4321
Thrift [16]	524,6	6,4	146305	12991	168472	20135
Pittsburgh [16]	240	21,1	210717	32027	265130	30161
Fuzzy-GAP [16]	33	0	279166	90017	290062	89155
Pitts-DNF min [16]	12,2	0,7	202943	43684	212018	44616
Pitts-DNF med [16]	18,6	1,4	86930	3955	99310	12996
Pitts-DNF max [16]	32,4	6,6	70207	1658	88017	8968
DynMO GFS [17]	25	–	18732	–	20858	–
NSGA-II_RB [18]	29	–	34230	–	39668	–
NSGA-II_KB [18]	29	–	26272	–	31174	–
PAES_RB [18]	30	–	15453	–	30906	–
PAES_KB [18]	30	–	22086	–	25212	–

Работа выполнена при финансовой поддержке РФФИ (проект №12-07-00055), РГНФ (проект №12-06-12008) и в соответствии с Госзаданием 7.701.2011.

Литература

1. Gacto M.J. Interpretability of linguistic fuzzy rule-based systems: An overview of interpretability measures / M.J. Gacto, R. Alcalá, F. Herrera // *Information Sciences*. – 2011. – Vol. 181. – P. 4340–4360.
2. Cordon O. A historical review of evolutionary learning methods for Mamdani-type fuzzy rule-based systems: Designing interpretable genetic fuzzy systems // *International Journal of Approximate Reasoning*. – 2011. – Vol. 52. – P. 894–913.
3. Interpretability assessment of fuzzy knowledge bases: A co-intension based approach / C. Mencar, C. Castiello, R. Cannone, A.M. Fanelli // *International Journal of Approximate Reasoning*. – 2011. – Vol. 52. – P. 501–518.
4. Ходашинский И.А. Идентификация параметров нечетких моделей типа сингтон на основе алгоритма роящихся частиц // *Информационные технологии*. – 2009. – № 6. – С. 8–11.
5. Ходашинский И.А. Идентификация параметров нечетких систем на основе адаптивного алгоритма роящихся частиц / И.А. Ходашинский, Д.С. Синьков // *Информационные технологии*. – 2011. – № 8. – С. 2–5.
6. Ходашинский И.А., Дудин П.А. Параметрическая идентификация нечетких моделей на основе гибридного алгоритма муравьиной колонии // *Автоматрия*. – 2008. – Том 44, № 5. – С. 24–35.
7. Ходашинский И.А. Идентификация нечетких систем на основе прямого алгоритма муравьиной колонии / И.А. Ходашинский, П.А. Дудин // *Искусственный интеллект и принятие решений*. – 2011. – № 3. – С. 26–33.
8. Ходашинский И.А. Идентификация нечетких систем на основе непрерывного алгоритма муравьиной колонии / И.А. Ходашинский, П.А. Дудин // *Автоматрия*. – 2012. – Т. 48, № 1. – С. 63–71.
9. Khodashinskii I.A. Identification of fuzzy systems using a continuous ant colony algorithm / I.A. Khodashinskii, P.A. Dudin // *Optoelectronics, Instrumentation and Data Processing*. – 2012. – Vol. 48, № 1. – P. 54–61.
10. Ходашинский И.А. Алгоритмы муравьиной и пчелиной колонии для обучения нечетких систем / И.А. Ходашинский, И.В. Горбунов, П.А. Дудин // *Доклады Том. гос. ун-та систем управления и радиоэлектроники*. – 2009. – № 2 (20). – С. 157–161.
11. Ходашинский И.А. Оптимизация параметров нечетких систем на основе модифицированного алгоритма пчелиной колонии / И.А. Ходашинский, И.В. Горбунов // *Мехатроника, автоматизация, управление*. – 2012. – №10. – С. 15–20.
12. Ходашинский И.А. Построение нечетких аппроксиматоров на основе метода перемещения бактерий / И.А. Ходашинский, Н.Н. Земцов, Р.В. Мещеряков // *Известия высших учебных заведений. Физика*. – 2012. – Т. 55, № 3. – С. 57–61.
13. Khodashinskii I.A. Construction of fuzzy approximators based on the bacterial foraging method / I.A. Khodashinskii, N.N. Zemtsov, R.V. Meshcheryakov // *Russian Physics Journal*. – 2012. – Vol. 55, № 3. – P. 301–305.
14. Ходашинский И.А. Идентификация нечетких систем на базе алгоритма имитации отжига и методов, основанных на производных // *Информационные технологии*. – 2012. – № 3. – С. 14–20.
15. Gath I. Unsupervised optimal fuzzy clustering / I. Gath, A.B. Geva // *IEEE Transactions on Pattern Analysis and Machine Intelligence*. – 1989. – Vol. 7. – P. 773–781.
16. Casillas J. Learning consistent, complete and compact sets of fuzzy rules in conjunctive normal form for regression problems / J. Casillas, P. Martinez, A.D. Benitez // *Soft Computing*. – 2009. – Vol. 13. – P. 451–465.
17. Pulkkinen P. Dynamically Constrained Multiobjective Genetic Fuzzy System for Regression Problems / P. Pulkkinen, H.A. Koivisto // *IEEE Transactions on Fuzzy Systems*. – 2010. – Vol. 18. – P. 161–177.
18. A Multiobjective Evolutionary Approach to Concurrently Learn Rule and Data Bases of Linguistic Fuzzy-Rule-Based Systems / R. Alcalá, P. Ducange, F. Herrera et al. // *IEEE Transactions on Fuzzy Systems*. – 2009. – Vol. 17. – P. 1106–1122.

Ходашинский Илья Александрович

Д-р техн. наук, профессор каф. комплексной информационной безопасности
электронно-вычислительных систем (КИБЭВС) ТУСУРа

Тел.: 8 (382-2) 41-34-26

Эл. почта: hodashn@rambler.ru

Горбунов Иван Викторович

Аспирант каф. КИБЭВС ТУСУРа

Тел.: 8 (382-2) 41-34-26

Эл. почта: noby.Ardor@gmail.com

Синьков Дмитрий Сергеевич

Аспирант каф. КИБЭВС ТУСУРа

Тел.: 8 (382-2) 41-34-26

Эл. почта: express@sibmail.com

Hodashinsky I.A., Gorbunov I.V., Sinkov D.S.

Algorithms for generating structures of two-criterion Pareto optimal fuzzy approximators

This paper proposes an approach, enabling one to build Pareto optimal fuzzy approximators. We present algorithm generation of fuzzy approximator structure. Simulation results confirmed that algorithms are a useful tool for tuning fuzzy systems to achieve better performance.

Keywords: fuzzy approximators, Pareto optimality, generating structure, metaheuristics.

**Требования к подготовке рукописей статей,
представляемых для публикации в журнале
«Доклады Томского государственного университета
систем управления и радиоэлектроники»**

1. Оригинал на бумажном носителе должен полностью соответствовать электронному варианту.

2. Электронный вариант должен быть представлен в виде файла, названного по-русски фамилией первого автора, на дискете или диске в формате Word 2003. Предпочтительнее представить его по электронной почте.

3. Текст статьи должен быть отпечатан через один интервал (множитель 1,05) шрифтом Times New Roman 10,5 кегля на одной стороне листа белой писчей бумаги формата А4 с полями шириной 25 мм, без помарок и вставок. Размер статьи со всеми атрибутами должен быть, как правило, не более пяти страниц.

Шаблон статьи размещен на сайте:

http://www.tusur.ru/ru/science/tusur_reports_magazine/template.dot.

4. Одни и те же символы в тексте, формулах, таблицах и рисунках должны быть единообразными по написанию. Русские и греческие символы набираются прямым шрифтом, а латинские – курсивом, кроме слов, их сокращений, имен функций, программ, фирм и химических формул.

5. Все употребляемые обозначения и сокращения должны быть пояснены.

6. Единицы измерения физических величин должны соответствовать Международной системе единиц (СИ).

7. Таблицы и рисунки должны иметь тематические заголовки (не повторяющие фразы-ссылки на них в тексте). (Рис. 1. Название рисунка; Таблица 1. Название таблицы). Большие блоки расшифровки условных обозначений лучше приводить в тексте. Подписи и надписи – Times New Roman, 10 пт. На все рисунки и таблицы должны быть ссылки в тексте (... на рис. 3, ... в табл. 2).

8. Рисунки и фотографии должны быть **черно-белыми**, четкими, контрастными, аккуратными, сгруппированными. Графики – не жирно, сетка – четко. Единицы измерения – на русском. Десятичная запятая (не точка).

Рисунки могут быть выполнены в программах CorelDraw, Illustrator, Word, Visio и должны давать возможность внесения исправлений.

9. Иллюстрации, разрешением не менее 300 dpi, дублируются отдельными файлами. Если это невозможно, должны быть предоставлены оригиналы иллюстраций, пригодные для полиграфического исполнения. Масштаб изображения – наиболее мелкий (при условии читаемости).

10. Формулы должны быть набраны в формульном редакторе (Equation или MathType) программы Word.

11. На все источники, указанные в списке литературы, должны быть ссылки по тексту (нумерация в порядке упоминания, например, [1, 2], [5–7]). Описание источников должно соответствовать ГОСТ 7.1-2003 и ГОСТ Р 7.0.5-2008 и содержать всю необходимую для идентификации источника информацию, а именно: *для неперIODических изданий* – фамилию и инициалы автора, полное название работы, место издания, название издательства, год издания, количество страниц; *для периодических изданий* – фамилию, инициалы автора, полное название работы, название журнала, год выпуска, том, номер, номера страниц.

12. Статья должна иметь (в порядке следования): УДК; И.О. Фамилии авторов; заглавие; аннотация (не реферат); ключевые слова; основной текст статьи; список библиографий под подзаголовком "Литература"; сведения об авторах; далее на английском языке: Фамилии авторов И.О., заглавие статьи, аннотацию, ключевые слова. Сведения об авторах включают в себя фамилию, имя, отчество, ученую степень, ученое звание, должность, место работы, телефон, электронный адрес.

Бумажный вариант рукописи статьи должен быть подписан авторами и иметь сопроводительное письмо на бланке организации (для сторонних авторов).

Плата за публикацию статей не взимается.

Материальные претензии авторов, связанные с распространением материалов их статей после опубликования, не принимаются.

Авторы несут полную ответственность за содержание статей и за последствия, связанные с их публикацией.

Примечание. Адрес для переписки: vnmas@main.tusur.ru Тел.: 7-382-2-51-21-21