

СИСТЕМЫ И СРЕДСТВА ИНФОРМАТИКИ

**Научный журнал Российской академии наук
(издается под руководством Отделения нанотехнологий
и информационных технологий РАН)**

Издается с 1989 года
Журнал выходит ежеквартально

Учредители:

**Российская академия наук
Федеральный исследовательский центр
«Информатика и управление» Российской академии наук**

РЕДАКЦИОННЫЙ СОВЕТ

академик РАН И. А. Соколов — председатель Редакционного совета
академик РАН Г. И. Савин
академик РАН А. Л. Стемпковский
член-корреспондент РАН Ю. Б. Зубарев
профессор Ш. Долев (S. Dolev, Beer-Sheva, Israel)
профессор Ю. Кабанов (Yu. Kabanov, Besancon, France)
профессор М. Никулин (M. Nikulin, Bordeaux, France)
профессор В. Ротарь (V. Rotar, San-Diego, USA)
профессор М. Финкельштейн (M. Finkelstein, Rostok, Germany)

РЕДАКЦИОННАЯ КОЛЛЕГИЯ

академик РАН И. А. Соколов — главный редактор
профессор, д.ф.-м.н. С. Я. Шоргин — заместитель главного редактора
д.т.н. В. Н. Захаров проф., д.ф.-м.н. В. Ю. Королев
проф., д.т.н. В. Д. Ильин проф., д.г.-м.н. Р. Б. Сейфуль-Мулюков
проф., д.ф.-м.н. Л. А. Калининченко проф., д.т.н. И. Н. Синицын
д.т.н. В. А. Козмидади к.т.н. А. В. Филин
проф., д.т.н. К. К. Колин к.ф.-м.н. С. А. Христочевский

Редакция

профессор, д.г.-м.н. Р. Б. Сейфуль-Мулюков
к.ф.-м.н. Е. Н. Арутюнов
С. Н. Стригина (ответственный секретарь)

© Федеральный исследовательский центр «Информатика
и управление» Российской академии наук, 2016

Журнал включен в базу данных Russian Science Citation Index (RSCI),
интегрированную с Web of Science

Журнал входит в систему Российского индекса научного цитирования (РИНЦ)
Журнал включен в базу данных CrossRef (систему DOI — Digital Object Identifier),
в базу данных Ulrich's periodicals directory
и в информационную систему «Общероссийский математический портал Math-Net.Ru»

Журнал реферируется в «Реферативном журнале» ВИНТИ
и в системе Google Scholar

Журнал включен в сформированный Минобрнауки России Перечень рецензируемых научных изданий, в которых должны быть опубликованы основные научные результаты диссертаций на соискание ученой степени кандидата наук, на соискание ученой степени доктора наук

<http://www.ipiran.ru/journal/collected>

СИСТЕМЫ И СРЕДСТВА ИНФОРМАТИКИ

Том 26 № 1 Год 2016

СОДЕРЖАНИЕ

Разрежённые буфера — использование механизма виртуальной памяти для уменьшения расходования физической памяти и экономии процессорного времени И. М. Адамович, Д. В. Земсков	4
Структурно-функциональная модель управления потоками пациентов региональной консультативной поликлиники В. Г. Азанов	13
Алгоритм и программный комплекс для анализа характеристик выходного излучения плазменного релятивистского генератора сверхвысоко частоты С. Е. Андреев	30
Экспериментальный анализ одного метода кластеризации и ранжирования многомерных данных нейронной сетью Кохонена В. И. Аникин, О. В. Аникина, А. А. Карманова	44
Об одной реализации автоматизированного средства аттестации А. К. Горшенин, А. С. Кучин	62
Построение клиентских .NET-приложений в распределенных схмотехнических системах автоматизированного проектирования В. Н. Гридин, В. И. Анисимов, А. Д. Ахмад	76
Проблемы обеспечения семантической геоинтероперабельности и согласования понимания семантики геоданных С. К. Дулин, Н. Г. Дулина, Д. А. Никишин	86
Некоторые вопросы практической реализации концепции SDN В. Б. Егоров	109

СИСТЕМЫ И СРЕДСТВА ИНФОРМАТИКИ

Том 26 № 1 Год 2016

СОДЕРЖАНИЕ

Информационная безопасность ситуационных центров А. А. Зацаринный, В. И. Королёв	121
Принципы создания системы мониторинга и анализа мирового потока научно-технической информации Ю. П. Калинин, Ал-др А. Хорошилов, Ал-ей А. Хорошилов	139
Методика расчета надежности международного термоядерного экспериментального реактора и оптимизация действий по снижению рисков Г. М. Коновалов	166
Информационная система поддержки принятия процессуальных решений Р. Р. Рзаев, Ф. Б. Агаев, А. И. Гоюшов, З. Р. Джамалов	182
Нормальные и ортогональные субоптимальные фильтры для нелинейных стохастических систем на многообразиях И. Н. Сеницын	199
Математическое обеспечение аналитического моделирования нормальных процессов в стохастических системах со сложными дробно-рациональными нелинейностями И. Н. Сеницын, В. И. Сеницын, И. В. Сергеев, Э. Р. Корепанов, В. В. Белоусов, В. С. Шоргин	227
Об авторах	248
Правила подготовки рукописей статей	251
Requirements for manuscripts	255

РАЗРЕЖЁННЫЕ БУФЕРА — ИСПОЛЬЗОВАНИЕ МЕХАНИЗМА ВИРТУАЛЬНОЙ ПАМЯТИ ДЛЯ УМЕНЬШЕНИЯ РАСХОДОВАНИЯ ФИЗИЧЕСКОЙ ПАМЯТИ И ЭКОНОМИИ ПРОЦЕССОРНОГО ВРЕМЕНИ

И. М. Адамович¹, Д. В. Земсков²

Аннотация: Рассматриваются разрежённые буфера — диапазоны адресов виртуальной памяти компьютера, для которых частично или полностью отсутствует отображение в физическую память. Такие буфера предназначены для использования в подпрограммах для передачи в них входных или получения из них выходных данных таким образом, что при чтении или записи подпрограммой данных по адресу, для которого не выделена физическая память, происходит вызов зарегистрированной подпрограммы-обработчика исключительной ситуации, выделяющей физическую память для фрагмента виртуальной памяти, к которому произошло обращение, а в случае чтения данных — заполняет выделенный фрагмент необходимыми подпрограмме данными. Действия, производимые подпрограммой-обработчиком, не влияют на исполнение подпрограммы, т. е. происходят незаметно для нее. Такой механизм позволяет избежать ненужного выделения физической памяти в случае, когда количество записанных или считанных подпрограммой данных меньше размера буфера, а в случае чтения данных — сэкономить процессорное время, необходимое для заполнения части буфера данными, которые не будут использованы подпрограммой.

Ключевые слова: информатика; виртуальная память; буфера памяти; подпрограммы; уменьшение расходования памяти; экономия процессорного времени

DOI: 10.14357/08696527160101

1 Термины и определения

Подпрограмма — фрагмент программного кода, к которому можно обратиться из другого места программы. В разных языках программирования может также называться процедурой, функцией, методом. Иногда используется общий термин «вызываемый блок кода».

Аргументы подпрограммы — значения, передаваемые подпрограмме, предназначенные для задания входных данных или их месторасположений, влияющих

¹Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, adam@amsd.com

²Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, dima@amsd.com

на поведение подпрограммы, а также задающие месторасположения для выходных данных.

Виртуальная память — технология управления памятью в компьютерах, реализованная аппаратными и программными средствами, с помощью которых адресам памяти, используемым программой (виртуальным адресам), ставятся в соответствие физические адреса памяти компьютера (так называемое отображение виртуального адреса на физический).

Страничная память — способ организации виртуальной памяти, при котором единицей отображения виртуальных адресов на физические является регион постоянного размера (так называемая страница).

Резервирование памяти — выделение свободного диапазона адресов в виртуальной памяти без назначения ему диапазона адресов физической памяти.

Выделение памяти — назначение диапазону виртуальных адресов диапазона адресов физической памяти.

Отсутствие доступа к странице памяти — ситуация при обращении программы по адресу памяти, не отображенной в физическую память.

Буфер — непрерывный диапазон адресов памяти, предназначенный для чтения из него подпрограммой данных, записи в него данных или того и другого.

Выходной буфер — буфер, предназначенный для записи подпрограммой выходных данных для их последующего использования в вызывающей программе.

Входной буфер — буфер, предназначенный для передачи подпрограмме входных данных.

Разрежённый буфер — буфер, для которого частично или полностью отсутствует отображение в физическую память.

Программное исключение — ситуация, прерывающая нормальный ход исполнения программы и требующая исполнения дополнительного программного кода (подпрограммы-обработчика исключения).

Обработчик исключения — подпрограмма, вызываемая средствами операционной системы при возникновении программного исключения.

Регистрация обработчика исключения — задание адреса подпрограммы, которая будет вызываться средствами операционной системы при возникновении исключения конкретного типа.

Обработчик отсутствия доступа к странице памяти — подпрограмма-обработчик исключения, которая вызывается при возникновении ситуации отсутствия доступа к странице памяти, получающая при вызове адрес соответствующей страницы.

Контроллер буфера — подпрограмма, вызываемая в описываемой реализации при возникновении исключения отсутствия доступа к странице памяти, относящейся к конкретному буферу.

Диспетчер контроллеров буферов — подпрограмма, получающая в качестве аргумента адрес памяти, принадлежащий буферу, и вызывающая соответствующий этому буферу контроллер.

2 Введение

Любая программа использует множество подпрограмм, большое количество которых принимает в качестве исходных или выдает в качестве результирующих данные, хранящиеся в оперативной памяти и имеющие переменную длину.

Наиболее характерным примером таких данных служат символьные строки, без использования которых трудно представить себе законченную программу. Помимо использования символьных строк для представления текста они используются для задания имен файлов в файловой системе и т. п. Бинарные, т. е. нетекстовые, данные переменной длины используются во множестве алгоритмов, к каковым относятся алгоритмы компрессии, шифрования данных и др. Важным свойством таких данных является то, что необходимость их присутствия в оперативной памяти ограничена по времени на период исполнения использующих их подпрограмм. Несмотря на то что подпрограммы осуществляют доступ к разным фрагментам таких данных одновременно, вызывающая программа предоставляет их перед выполнением подпрограммы в полном объеме, если эти данные являются входными, если же эти данные — выходные, то их обработка вызывающей программой откладывается до момента их полного формирования и окончания исполнения подпрограммы.

Зачастую вызывающей программе неизвестно как количество реально используемых подпрограммой входных данных, так и количество выходных данных, выданных подпрограммой. В таком случае подпрограмме предоставляются входные данные или выделяется память под выходные данные максимально возможной необходимой длины, что приводит к излишнему расходованию памяти и ресурсов компьютера для генерации входных данных.

В многозадачной системе отрицательное влияние ненужного расходования памяти и процессорного времени растет, также негативно сказывается увеличение времени выполнения подпрограмм, так как оперативная память под данные занята на время их выполнения.

Предлагаемый механизм предназначен для уменьшения количества выделяемой под данные для подпрограмм физической памяти и накладных расходов на генерацию входных данных за счет использования виртуальной памяти операционных систем, а именно того их свойства, что фрагменты адресного пространства памяти можно зарезервировать без выделения соответствующих им фрагментов физической памяти, а выделение таковых производить только при обращении подпрограммы по зарезервированному адресу.

Предложенный в статье механизм может быть реализован средствами множества существующих операционных систем (ОС), однако в качестве примера рассматривается его возможная реализация в семействе ОС Windows начиная с версии Windows XP. Данное семейство ОС в статье называется ОС Windows, без дополнительного указания ОС Windows XP как минимальной поддерживаемой версии.

3 Обработка исключительных ситуаций в разных операционных системах

Обработка исключений реализуется механизмом регистрации средствами ОС подпрограмм-обработчиков исключений, могущих возникнуть при исполнении процесса. В разных ОС такие подпрограммы могут быть как привязаны к определенному исключению или группе исключений, так и не иметь такой привязки, являясь обработчиками исключений любого типа.

Количество возможных зарегистрированных обработчиков также зависит от ОС и может варьироваться от одного до произвольного числа таковых.

При возникновении исключения управление передается зарегистрированной для данного типа исключения подпрограмме-обработчику или, если зарегистрированных обработчиков несколько, каждому из них последовательно.

По окончании обработки исключения обработчик имеет две основные возможности:

- (1) сообщить системе, что исключение не обработано. В таком случае система продолжит последовательно вызывать оставшиеся зарегистрированные обработчики;
- (2) сообщить системе, что исключение обработано. В таком случае система продолжит исполнение прерванного кода с повторного исполнения инструкции, вызвавшей исключение.

В ОС Windows этот механизм называется Vectored Exception Handling (VEH), что переводится как «векторная обработка исключений», и имеет следующие свойства:

- позволяет зарегистрировать неограниченное количество обработчиков исключений;
- регистрируемые обработчики исключений не имеют привязки к типу исключения, а получают информацию о типе исключения из передаваемых при вызове подпрограммы-обработчика системой данных;
- не привязаны к потоку исполнения процесса.

Регистрация обработчика в ОС Windows осуществляется вызовом функции API `AddVectoredExceptionHandler`.

4 Механизм действия разрежённых буферов

Механизм действия разрежённых буферов, представленный на диаграмме последовательности, состоит из следующих этапов:

- (1) программа резервирует средствами ОС диапазон адресов под буфер максимально необходимого подпрограмме размера;
- (2) если контроллеру буфера для работы необходимы дополнительные данные, программа создает и/или инициализирует их;

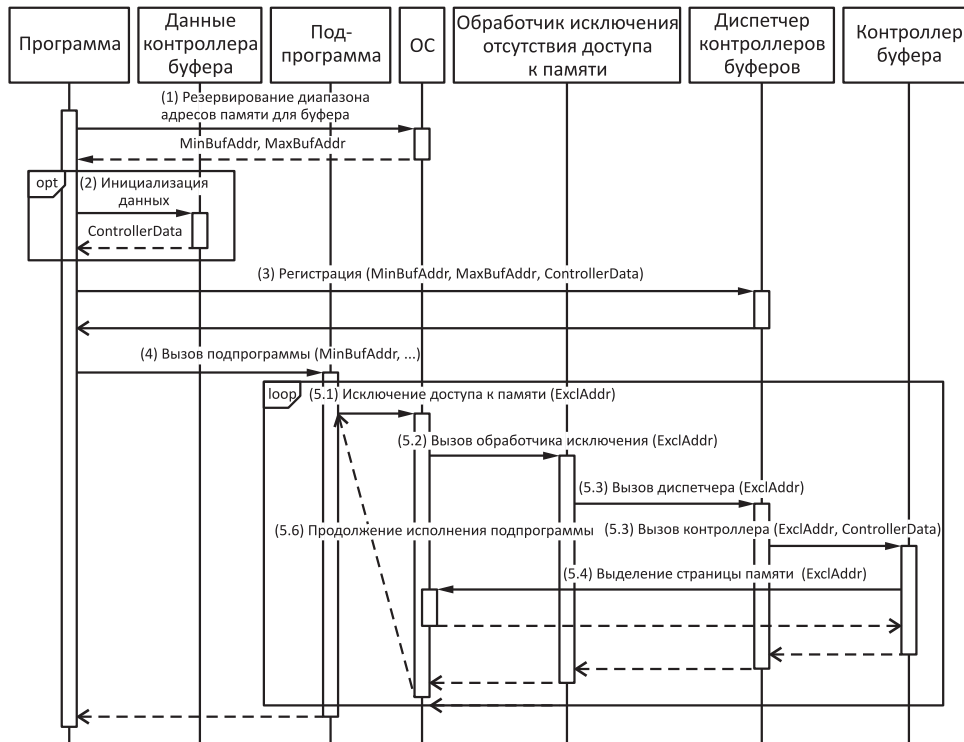


Диаграмма последовательности

- (3) программа регистрирует у диспетчера контроллеров буферов минимальный и максимальный адреса буфера и соответствующий контроллер буфера и, при необходимости, данные для его работы;
- (4) программа вызывает подпрограмму, передавая ей буфер для чтения или записи данных;
- (5) во время исполнения подпрограммы при каждой ее попытке чтения из буфера или записи в буфер данных по адресу, не имеющему выделенной физической страницы памяти, выполняются следующие действия:
 - (5.1) возникает ситуация исключения и управление передается ОС;
 - (5.2) операционная система вызывает зарегистрированный обработчик исключения доступа к памяти и передает ему адрес памяти, при обращении к которому произошло исключение;
 - (5.3) обработчик исключения доступа к памяти вызывает через диспетчера контроллеров буферов соответствующий адресу исключения контроллер, передавая ему в качестве аргументов адрес исключения

и зарегистрированные для этого контроллера данные, необходимые для его работы;

- (5.4) контроллер выделяет средствами ОС страницу памяти, содержащую адрес, обращение к которому вызвало исключение, и если буфер является входным, то заполняет эту страницу данными;
- (5.5) контроллер возвращает управление диспетчеру контроллеров буферов; тот, в свою очередь, возвращает управление обработчику исключения доступа к памяти; последний возвращает управление операционной системе с указанием, что ошибка исправлена;
- (5.6) операционная система восстанавливает контекст исполнения подпрограммы и продолжает ее исполнение, начиная с повторного исполнения инструкции процессора, вызвавшей возникновение ситуации исключения.

Существуют более сложные сценарии работы с буферами, например:

- (1) при последовательном чтении из буфера постепенное освобождение начала буфера одновременно с выделением страниц в конце буфера;
- (2) при последовательной записи в буфер постепенное использование данных с головы буфера одновременно с выделением и заполнением страниц в конце буфера.

5 Примеры использования

В качестве одного из примеров использования предлагаемого механизма рассмотрим работу с API (Application Programming Interface) функциями файловой системы ОС Windows, которые принимают в качестве аргумента символьную строку переменной длины, определяющую имя файла.

К таковым относятся функции открытия, переименования, удаления файла и им подобные. Внутренняя реализация таких функций работает со строками в кодировке UTF-16 [1] и, несмотря на то что существуют аналогичные функции, принимающие в качестве аргументов имена файлов в однобайтной кодировке, в таких функциях производится их перекодировка в формат UTF-16.

Поскольку имена файлов в однобайтной кодировке не могут содержать все допустимые в UCS (Universal Coded Character Set) [2] символы, удобно использовать кодировку с переменной шириной символов, например UTF-8 [3] или ACE [4]. Так как API-функции файловой системы ОС Windows напрямую с такими кодировками не работают, необходимо перед вызовом этих функций производить преобразование таких строк в UTF-16. При таком преобразовании количество байт для хранения преобразованной строки заранее не известно, хотя и может быть определено предварительным сканированием исходной строки, что требует дополнительного процессорного времени.

В ОС Windows существует ограничение на максимальную длину имени файла, которое составляет 32 767 символов, или 65 534 байт, если используется кодировка UTF-16.

Таким образом, можно было бы иметь заранее выделенный буфер максимальной длины и использовать его для преобразования имени файла из многобайтной кодировки в UTF-16. Однако в многопоточной программе может возникнуть одновременная необходимость использования такого буфера в нескольких потоках исполнения, что повлечет за собой необходимость ожидания потоком освобождения буфера. Использование отдельного выделенного буфера максимального размера для каждого потока приведет к нерациональному использованию памяти.

В предлагаемом механизме используется один или несколько временных буферов, каждый из которых является разреженным, т. е. память под буфер максимально допустимого размера зарезервирована, но не выделена или выделена частично (например, только размером в одну страницу в начале буфера). При передаче такого буфера подпрограмме транскодирования имени файла в качестве выходного физическая память под буфер будет выделяться только при необходимости, при попытке записи очередной порции транскодированного имени файла подпрограммой в невыделенную область буфера.

Другим примером использования механизма может быть чтение из текстового файла строки, поскольку длина такой строки определяется нахождением символа (или символов) конца строки.

Обычно на максимальную длину строки для чтения из файла накладывается, хотя и достаточно большое, ограничение — порядка 65 000 байт. Аналогично первому примеру подпрограмма чтения из файла получает разреженный выходной буфер и значение его максимального размера и пишет в него строку из файла длиной, не превышающей этого размера, при этом не выделенная на момент записи в буфер память выделяется прозрачно для программы.

Еще одним примером может служить ситуация, когда подпрограмме должен быть передан входной буфер, количество затребованных подпрограммой данных из которого заранее не известно. Рассмотрим текстовый редактор, текст в котором хранится в кодировке с переменной длиной символа, например UTF-8 или АСЕ. Пользователю текстового редактора необходимо найти в тексте с текущей позиции курсора подстроку, используя регулярное выражение. Существующие библиотеки регулярных выражений не поддерживают напрямую поиск в тексте с кодировкой с переменной длиной символа (по крайней мере, авторам такие библиотеки неизвестны). Такой поиск, хотя теоретически и мог бы быть реализован, но достаточно сложен в реализации. Некоторые библиотеки регулярных выражений предоставляют возможность такого поиска путем предварительного неявного преобразования исходного текста в UTF-16, однако при этом преобразуется весь текст, по которому мог бы вестись поиск, несмотря на то что часть (возможно, большая) преобразованного текста подпрограммой не будет использована (та часть текста, которая находится после найденной строки). Описываемый механизм для такого поиска будет работать следующим образом:

1. Порождается разрежённый буфер длины, достаточной для преобразования всего текста, по которому будет производиться поиск, в UCS. Порожденный буфер не содержит ни одной выделенной страницы физической памяти.
2. Данный буфер подается в качестве входного подпрограмме поиска по регулярному выражению.
3. В процессе чтения подпрограммой данного буфера невыделенной страницы памяти происходит вызов средствами ОС предварительно установленного для данного буфера обработчика невыделенной страницы.
4. Обработчик выделяет страницу физической памяти для адреса памяти, обращение к которому привело к вызову обработчика, раскодирует очередную порцию исходного текста в формат UCS и возвращает управление системе, которая продолжает исполнение подпрограммы поиска с того места, где оно прервалось, незаметно для подпрограммы.

В результате происходит транскодирование исходного текста «по требованию», т. е. только той части, которая требуется подпрограмме поиска для нахождения заданной строки по регулярному выражению.

Литература

1. RFC 278 — UTF-16, an encoding of ISO 10646. <https://tools.ietf.org/html/2781>.
2. ISO/IEC 10646 — Information technology — Universal Coded Character Set (UCS). http://standards.iso.org/ittf/PubliclyAvailableStandards/c063182JISO_IEC_10646_2014.zip.
3. RFC 3629 — UTF-8, a transformation format of ISO 10646. <https://tools.ietf.org/html/3629>.
4. *Адамович И. М., Земсков Д. В.* Настраиваемая схема кодирования символов переменным числом октетов — ACE // Системы и средства информатики, 2014. Т. 24. № 4. С. 124–134.

Поступила в редакцию 17.09.15

SPARSE BUFFERS — USING THE VIRTUAL MEMORY MECHANISM TO REDUCE PHYSICAL MEMORY AND CPU TIME USAGE

I. M. Adamovich and D. V. Zemskov

Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilova Str., Moscow 119333, Russian Federation

Abstract: This article discusses sparse buffers — address ranges in computer’s virtual memory, for which mapping to the physical memory is partially or completely absent. Such buffers are intended for use in subroutines to transfer to

them input data or to obtain from them output data so that when a subroutine reads or writes data at addresses for which no physical memory is allocated, a registered exception handler is called that allocates physical memory for virtual memory fragment to which access has been made, and in the case of reading data fills that memory with the necessary data. Actions made by a subroutine handler do not affect the execution of the subroutine, i. e., are transparent to it. This mechanism allows avoiding unnecessary allocation of physical memory when the amount of read or written data is less than the provided buffer's size and, in the case of reading the data, allows saving the CPU time required to fill the buffer with portions of data that would not be used by the subroutine.

Keywords: informatics; virtual memory; memory buffer; subroutines; reducing memory usage; reducing CPU time usage

DOI: 10.14357/08696527160101

References

1. RFC 2781 — UTF-16, an encoding of ISO 10646. Available at: <https://tools.ietf.org/html/2781> (accessed August 16, 2015).
2. ISO/IEC 10646 — Information technology — Universal Coded Character Set (UCS). Available at: http://standards.iso.org/ittf/PubliclyAvailableStandards/c063182.ISO.IEC_10_646_2014.zip (accessed August 16, 2015).
3. RFC 3629 - UTF-8, a transformation format of ISO 10646. Available at: <https://tools.ietf.org/html/3629> (accessed August 16, 2015).
4. Adamovich, I. M., and D. V. Zemskov. 2014. Nastraivaemaya skhema kodirovaniya simvolov peremennym chislom oktetov — ACE [Adjustable variable-length character encoding scheme — ACE]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 24(4):124–134.

Received September 17, 2015

Contributors

Adamovich Igor M. (b. 1934) — Candidate of Science (PhD) in technology, Head of Department, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; adam@amsd.com

Zemskov Dmitry V. (b. 1969) — leading programmer, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; dima@amsd.com

СТРУКТУРНО-ФУНКЦИОНАЛЬНАЯ МОДЕЛЬ УПРАВЛЕНИЯ ПОТОКАМИ ПАЦИЕНТОВ РЕГИОНАЛЬНОЙ КОНСУЛЬТАТИВНОЙ ПОЛИКЛИНИКИ

*В. Г. Азанов*¹

Аннотация: Рассмотрена проблема доступности и оперативности оказания консультативной узкоспециализированной медицинской помощи для пациентов региональной поликлиники, обслуживающей обширный регион Российской Федерации с низкой средней плотностью населения (на примере Красноярского края). Предлагается решать данную проблему с помощью медицинской информационной системы (МИС) «Виртуальная поликлиника», в которой реализованы следующие принципы управления потоками удаленных пациентов: дифференциация пациентов по видам консультативной помощи (очное или заочное консультирование), заочное формирование маршрутного листа по оказанию очных консультативных услуг пациенту внутри региональной поликлиники, дистанционное заочное консультирование. Представлено обоснование предлагаемого подхода в виде результатов структурно-функционального моделирования предметной области по методологии SADT (Structured Analysis and Design Technique). Приведены результаты моделирования управления потоками пациентов в виде диаграмм потоков данных в графической нотации IDEF0 «как было» (AS IS) и «как стало» (TO BE). Выявлены особенности дистанционного заочного консультирования и планирования индивидуального маршрута пациента в поликлинике регионального уровня (на примере краевой клинической больницы (ККБ) г. Красноярска). Дано описание системы «Виртуальная поликлиника», где отражены состав и функции системы, представление данных, категории пользователей, методические и финансовые аспекты внедрения и средства мониторинга степени доступности оказываемой консультативной помощи. Продемонстрирован опыт внедрения МИС «Виртуальная поликлиника» в региональной консультативной поликлинике и 134 медицинских организациях (МО) Красноярского края, отправляющих пациентов на консультацию в ККБ.

Ключевые слова: медицинские информационные системы; дистанционное консультирование; управление потоками пациентов; индивидуальная маршрутизация пациентов; виртуальная поликлиника

DOI: 10.14357/08696527160102

1 Постановка задачи

1.1 Особенности решаемой задачи

В 2011 г. перед Министерством здравоохранения Красноярского края встала проблема сокращения очередей и повышения оперативности оказания узкоспеци-

¹Краевая клиническая больница, Красноярск, azanov@mail.ru

ализированной медицинской помощи пациентам, направляемым в региональную консультативную поликлинику при ККБ.

Были установлены основные причины появления данной проблемы:

- необоснованность направления пациентов на консультацию в ККБ из других МО региона и, как следствие, дополнительная нагрузка на региональную консультативную поликлинику;
- недостаточный набор предварительных обследований пациента на момент очной консультации в региональной поликлинике и, как результат, отказ в консультативном обслуживании пациента с рекомендацией дополнительного обследования по месту жительства;
- отсутствие средств поддержки оптимального планирования маршрута каждого отдельного пациента в региональной поликлинике с целью уменьшения времени его обслуживания, т. е. времени закрытия случая обращения.

Региональная консультативная поликлиника ККБ располагается в столице Красноярского края г. Красноярске. Объем консультативной помощи, оказываемой этой поликлиникой, составляет 278 000 посещений в год. Красноярский край является географически обширным регионом: максимальная протяженность его с севера на юг — 3000 км, максимальная ширина — 1500 км. Поэтому имеют место сложности, связанные с приездом пациентов на очную консультацию и различными возможностями пациентов по проживанию в г. Красноярске на период консультативной помощи. В этих условиях обоснованность направления пациента в консультативную поликлинику ККБ и сокращение времени его пребывания в г. Красноярске особенно важно. Согласно существующим в Красноярском крае нормам консультативной медицинской помощи, направить пациента на консультацию в региональную поликлинику может только врач одной из муниципальных или частных МО. Самостоятельное прямое обращение пациента в региональную поликлинику предусмотрено только в жизнеугрожающих для этого пациента состояниях.

1.2 Пути решения задачи

Для решения указанной выше проблемы на уровне Министерства здравоохранения Красноярского края было решено ввести новые принципы управления потоками пациентов в региональной поликлинике с целью снижения очередности и повышения оперативности оказания консультативной узкоспециализированной медицинской помощи. Для этого был предусмотрен обязательный предварительный анализ сопроводительной медицинской документации на предмет обоснованности направления, полноты обследований и обеспечения возможности формирования оптимального маршрута пациента при очном приеме в региональной поликлинике. Кроме того, было принято решение об использовании для этих целей МИС, способной выполнять следующие функции:

- (1) дифференциация пациентов по видам консультативной помощи (очное или заочное консультирование): проверка обоснованности направлений на очную консультацию;
- (2) поддержка заочных консультаций в режиме онлайн на основе прикрепленных файлов с сопроводительной медицинской документацией;
- (3) заочное формирование маршрутного листа пациента, в котором определен оптимальный маршрут внутри региональной поликлиники с указанием дат и времени посещения кабинетов от диагностики до врачебных приемов, минимизирующий промежуток времени от первого очного посещения до закрытия случая обращения;
- (4) автоматическая сверка данных полиса обязательного медицинского страхования с сервисом территориального Фонда обязательного медицинского страхования;
- (5) интеграция с уже существующей МИС ведения расписания приема врачей поликлиники ККБ;
- (6) поддержка контроля стандартов обследований в соответствии с приказом Министерства здравоохранения Красноярского края № 68-орг от 17.02.2011;
- (7) поддержка проведения мониторинга доступности узкоспециализированной медицинской помощи на объемах поданных заявок по оказанию консультативной медицинской помощи;
- (8) формирование истории заявок в разрезе пациентов.

В соответствии с решением Министерства здравоохранения Красноярского края на реализацию и внедрение всего проекта было отведено 3 мес. Оперативно был проведен анализ существующих систем, способных поддерживать перечисленные выше функции: Карельской МИС, МЕДИАЛОГ, ИНТЕРИН, ДОКА+, ДЖЕМИС, БАРС, АКСИМЕД, ЭВЕРЕСТ и многих других [1–3]. Анализ проводился на основе опросных листов, заполненных представителями разработчиков указанных систем, с обсуждением текущих функциональных возможностей и необходимой доработки этих систем под требуемый набор функций. Полностью соответствующей всем необходимым функциям МИС обнаружено не было, сроки доработки существующих МИС с целью введения новых функциональных возможностей превышали выделенное время на реализацию и внедрение проекта. Поэтому было принято решение о разработке МИС с требуемым набором функций.

Разработанная МИС получила название «Виртуальная поликлиника». В рамках реализации проекта было выполнено структурно-функциональное моделирование предметной области, осуществлено проектирование, реализация и внедрение системы «Виртуальная поликлиника».

Структурно-функциональное моделирование предметной области выполнено по методологии SADT. Методология структурного анализа и проектирова-

ния SADT представляет собой технологию моделирования и проектирования программного и информационного обеспечения для управления крупномасштабными распределенными системами, включающими в себя людей, технические и коммуникационные средства [4, 5]. Системы управления потоками удаленных пациентов по оказанию консультативной медицинской помощи относятся именно к таким системам — распределенным медицинским системам организационного типа [6–8].

Результаты структурно-функционального моделирования бизнес-процессов региональной консультативной поликлиники ККБ г. Красноярска послужили обоснованием выбранных проектных решений и принципов управления потоками удаленных пациентов ККБ в системе «Виртуальная поликлиника». Результатом структурно-функционального моделирования прежде всего являются модели управления потоками пациентов, выполненные в виде иерархически связанных между собой диаграмм потоков данных в графической нотации IDEF0 «как было» (AS IS) и «как стало» (TO BE), учитывающие особенности очного и заочного консультирования пациентов в поликлинике регионального уровня.

Далее в статье приводятся и анализируются основные диаграммы потоков (всего модель содержит 17 диаграмм).

2 Модели управления потоками пациентов

2.1 Модель консультативного приема «как было» (AS IS)

Существовавший ранее порядок направления пациентов на консультацию к узким специалистам региональной клиники ККБ предполагал обязательное оформление бумажного направления в консультативную поликлинику лечащим врачом пациента. Пациент с бумажным направлением приходил в регистратуру поликлиники ККБ, где пациенту выдавали талон на первичное посещение необходимого специалиста.

Прежний порядок консультативного приема можно описать диаграммой потоков «как было» (AS IS). Вид этой диаграммы для случая «первичный прием» представлен на рис. 1. Согласно этой диаграмме при первичном приеме врач вначале определяет перечень необходимых обследований и консультаций и далее пациент с этим перечнем посещает регистратуру, чтобы получить талоны на диагностические процедуры и консультативные приемы.

По завершении нужных обследований и консультаций проводился повторный прием у врача. Вид диаграммы для случая «повторный прием» представлен на рис. 2.

Из анализа моделей потоков данных, представленных на рис. 1 и 2, видно, что для пациентов трудности возникали по следующим причинам: отклонение приема в регистратуре по причине отсутствия необходимого набора обследований; пациент уже прибыл в г. Красноярск и обратился в регистратуру региональной поликлиники, однако попасть на консультацию смог только через несколько

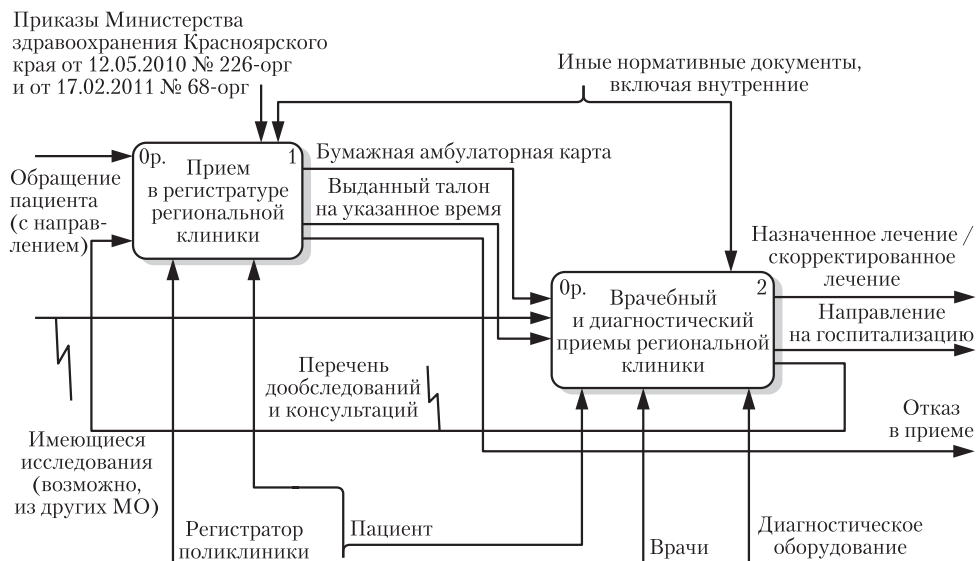


Рис. 1 Модель ранее действовавшего порядка консультативного приема для случая «первичный прием»

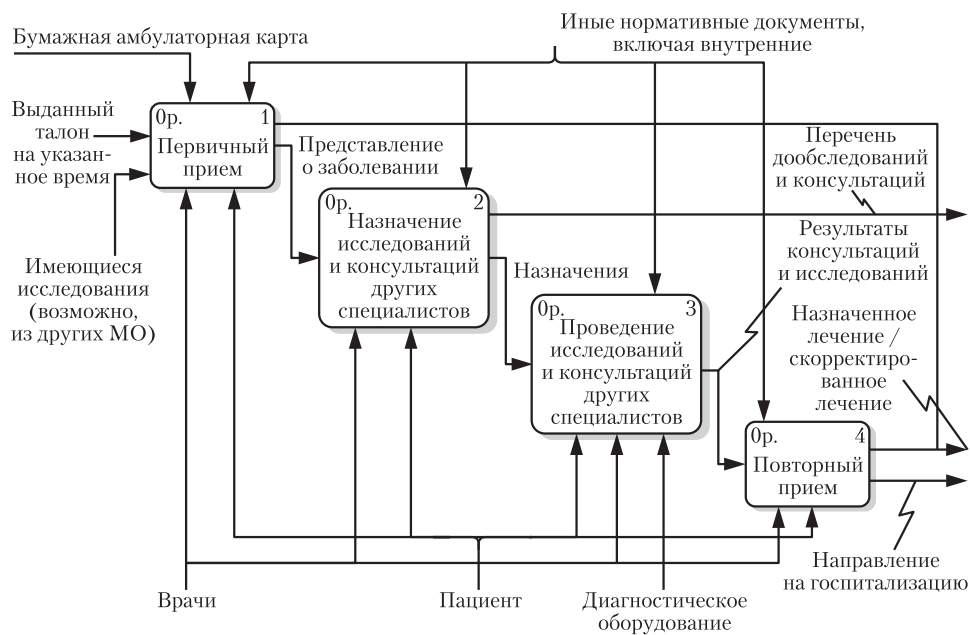


Рис. 2 Модель ранее действовавшего порядка консультативного приема для случая «повторный прием»

дней. Таким образом, самые узкие места прежней системы управления потоками пациентов — это отсутствие возможности заочного анализа сопроводительной документации и невозможность заранее спланировать маршрут пациента в региональной консультативной поликлинике. Устранение этих узких мест и стало целью создания системы «Виртуальная поликлиника».

2.2 Модель консультативного приема «как стало» (ТО BE)

Для создания системы «Виртуальная поликлиника» были построены модели процессов, описывающих взаимодействие удаленных МО с региональной консультативной поликлиникой ККБ.

На диаграмме, представленной на рис. 3, показаны два типа процессов: формирование направления удаленной МО и обработка заявки региональной поликлиникой (заочно).

Процессы, представленные на рис. 3, описывают:

- предварительный анализ (до приезда пациента на очный прием в региональную поликлинику) сопроводительной медицинской документации пациента,

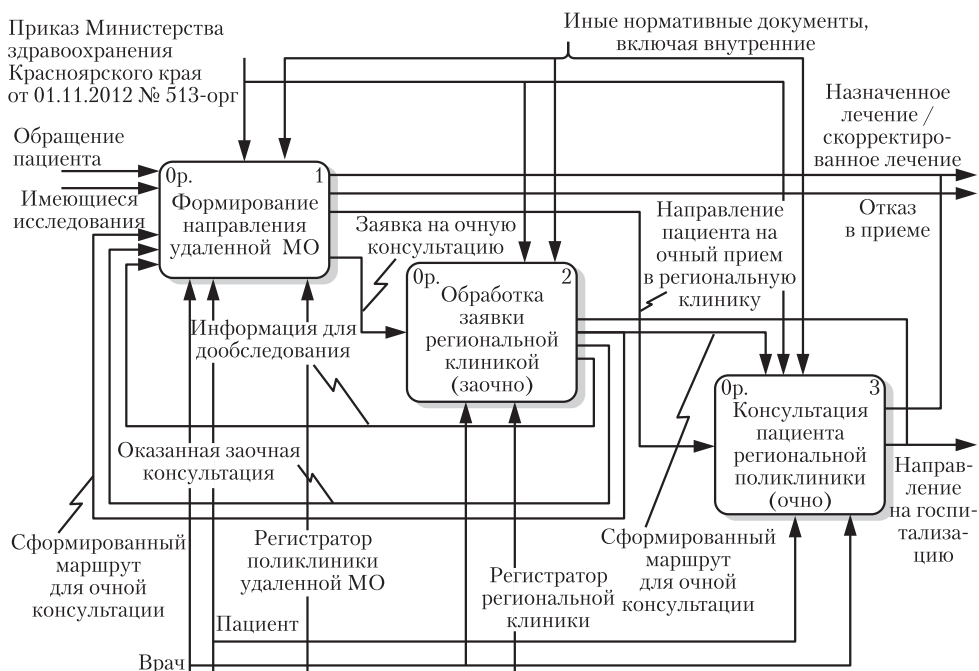


Рис. 3 Модель взаимодействия удаленной МО с региональной консультативной поликлиникой

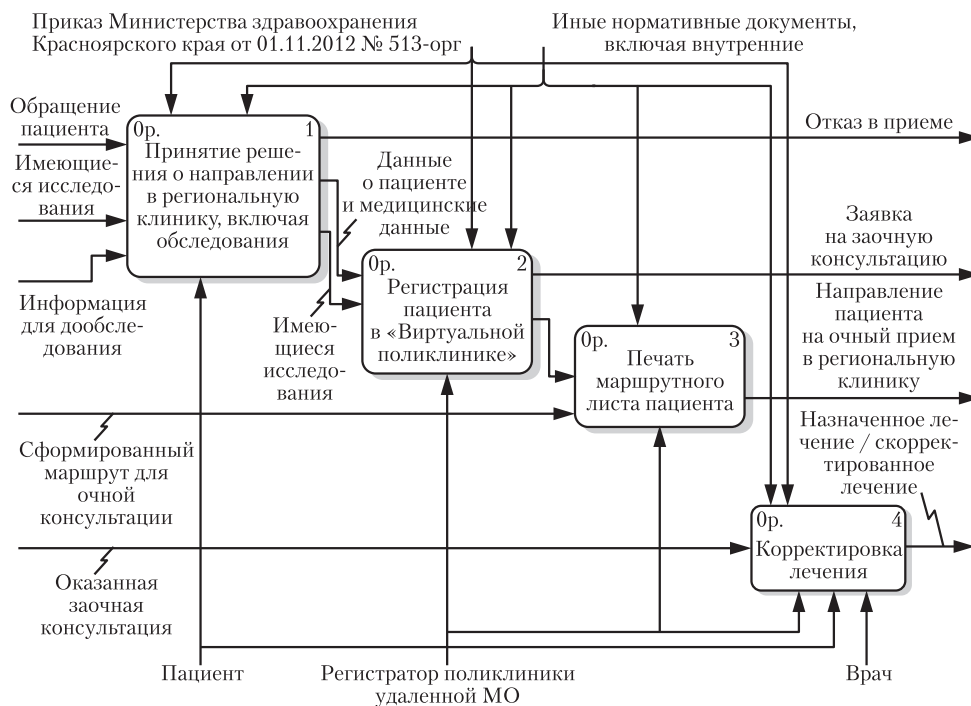


Рис. 4 Модель процесса формирования направления от удаленной МО

проверку обоснованности направления на очную консультацию и возможность заочного консультирования пациента в офлайн-режиме;

- возможность заочного формирования маршрутного листа пациента, в котором определен оптимальный маршрут внутри региональной поликлиники с указанием дат и времени посещения кабинетов от диагностики до врачебных приемов.

Процесс формирования направления удаленной МО (рис. 4) описывает следующую ситуацию: при возникновении потребности врача некоторой удаленной МО в консультативной помощи для пациента эта МО подает соответствующую заявку на защищенный ресурс «Виртуальная поликлиника». В заявке содержится основная информация о пациенте: паспортные данные, сопроводительная медицинская документация, отсканированная или выгруженная из МИС (при наличии).

Как видно из рис. 3 и 4, в некоторых случаях в МО возвращается информация для коррекции лечения пациента по месту жительства, при этом не требуется очного присутствия пациента. Однако допускается направление пациента на

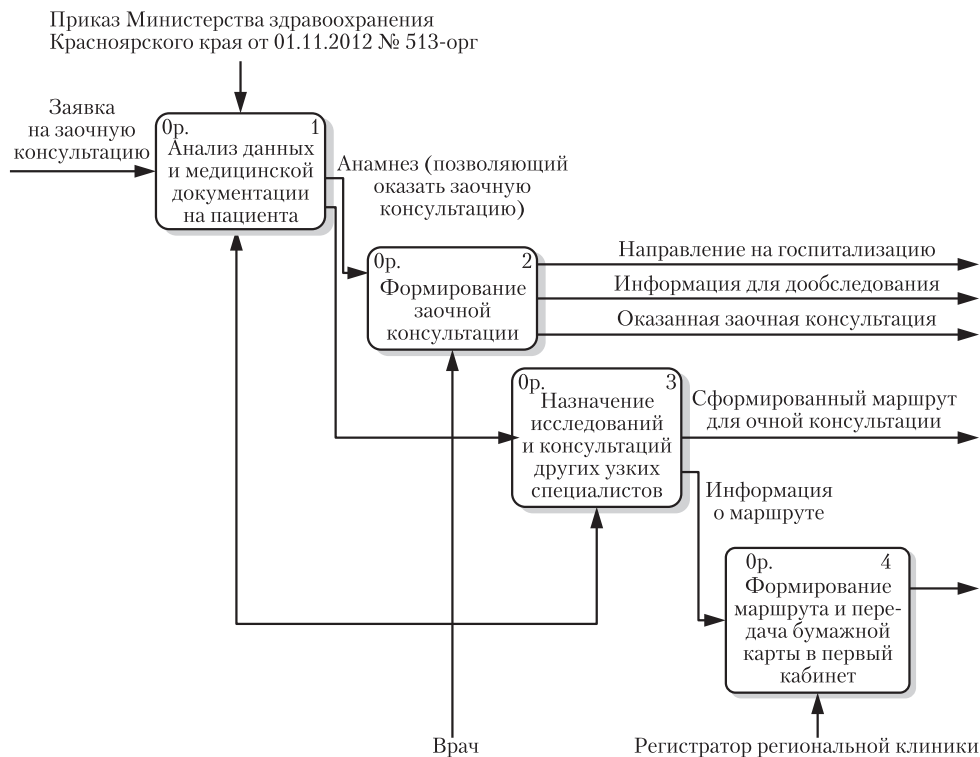


Рис. 5 Модель процесса обработка заявки в региональной клинике (заочно)

очный прием в региональную консультативную клинику. Возможность заочного консультирования снижает нагрузку на региональную поликлинику, вызванную непосредственным посещением пациентов, а для пациентов устраняет в ряде случаев приезд в ККБ.

Описанный на рис. 5 процесс заочной обработки заявок от удаленных МО отражает:

- подготовку и проведение заочной консультации. Такая консультация дает возможность сократить число необоснованных очных визитов пациентов в несложных диагностических и лечебных случаях;
- заочное формирование маршрутного листа для каждого пациента внутри региональной поликлиники и передачу бумажной карты в первый кабинет посещения.

Следует заметить, что порядок оказания медицинских услуг, отраженный в представленных моделях (рис. 3–5), не противоречит Федеральному закону «Об основах охраны здоровья граждан в Российской Федерации» [9].

3 Описание системы «Виртуальная поликлиника»

3.1 Алгоритм обработки заявок

Модель процесса обработки заявки региональной клиникой (см. рис. 5) определяет следующий двухэтапный алгоритм обработки направлений на оказание консультативной медицинской помощи.

Этап 1. В региональной поликлинике заявка поступает к врачу-эксперту, который анализирует приложенную к заявке документацию и на ее основании принимает решение о наличии показаний для очной консультации.

Если очная консультация не показана, то выполняется дистанционная корректировка лечения соответствующим врачом-специалистом, случай обращения в поликлинику ККБ закрывается. Если же очная консультация все же необходима, то врач-эксперт определяет, все ли необходимые диагностические обследования выполнены. Если требуется провести дополнительное обследование в учреждении по месту жительства, то врач указывает эту информацию в заявке и отправляет ее на доработку. В случае достаточности обследований и показаний к очной консультации формируется состав маршрута пациента, где указываются все дополнительные обследования и консультации узких специалистов, которые необходимо выполнить в ККБ.

Этап 2. Далее заявка становится доступна регистратору ККБ. Работник регистратуры формирует в полуавтоматическом режиме оптимальный маршрут консультаций и дополнительных обследований, сводя все посещения на один день, если это возможно. Для формирования маршрутной карты пациента используются специально разработанные программные средства, интерфейс которых изображен на рис. 6.

Заявка возвращается в МО с указанием для пациента информации о датах и времени приема врачей и диагностических процедур. Каждое утро работник регистратуры региональной поликлиники обеспечивает доставку бумажных амбулаторных карт в соответствующий кабинет приема пациента. Пациент при очном приеме в поликлинике ККБ минует регистратуру и идет сразу же в первый кабинет приема, указанный в его маршрутной карте.

В статье [10] предложен подход к формированию маршрутной карты пациента, где ему предоставляется возможность самому определять начальное время приема. Однако в условиях Красноярского края данный подход значительно усложнит процесс маршрутизации пациента и затруднит его автоматизацию.

В системе «Виртуальная поликлиника» учет пожелания пациента о начальной дате приема осуществляется лишь на этапе формирования заявки: врач удаленной МО уточняет желаемую дату и МО вносит ее в направление на консультирование. Далее эта информация рассматривается как рекомендательная при формировании маршрутного листа пациента.

Расписание

Дата подобрана автоматически. В этот день есть свободные талоны на всех специалистов (Период поиска: наименьшая дата из желаемых плюс 30 дней)

Ортопед Дата: 23.10.2012

Каб. 231			08:30	08:45	09:00	09:15	09:30	09:45	10:00	10:15	10:30	10:45	11:00
	11:15	11:30	11:45	12:00	12:15	12:30	12:45	13:00					

Примечание:

Уважаемый пациент! ПРИ НЕОБХОДИМОСТИ, в соответствии с назначениями врача - консультанта, в день консультации, Вам будут проведены исследования: развернутый анализ крови, общий анализ мочи, рентгенография, биохимические исследования, при госпитализации консультация кардиолога и электрокардиография. Обследования проводятся натощак. На руках иметь свежие рентген-снимки.

Сохранить

Дата подобрана автоматически. В этот день есть свободные талоны на всех специалистов (Период поиска: наименьшая дата из желаемых плюс 30 дней)

Эндокринолог Дата: 23.10.2012

Каб. 240						09:15	09:30	09:45	10:00	10:15	10:30	10:45	11:00
	11:15	11:30	11:45	12:00									
Каб. 240			14:30	14:45	15:00	15:15	15:30	15:45	16:00	16:15	16:30	16:45	
Каб. 340	08:00	08:15	08:30		09:00	09:15	09:30	09:45	10:00	10:15	10:30		11:00
	11:15	11:30	11:45	12:00									

Примечание:

Уважаемый пациент! При необходимости в соответствии с назначениями врача - консультанта в день назначения Вам будут проведены исследования: развернутый анализ крови, общий анализ мочи, биохимические исследования, гормональные исследования, УЗИ щитовидной железы.

Сохранить

Рис. 6 Интерфейс для формирования маршрутного листа пациента

3.2 Категории пользователей

Представленные выше модели управления потоками пациентов определили для системы «Виртуальная поликлиника» следующие категории пользователей:

- сотрудники удаленных МО, которые обеспечивают оформление и подачу заявки на консультирование и сопроводительной документации;
- представители страховых компаний. В случае отсутствия обязательного медицинского страхования у пациента они проводят подготовку документов и выдают полис пациенту, когда тот очно посещает региональную поликлинику;
- работники регистратуры региональной поликлиники. Обеспечивают формирование маршрутного листа для пациентов, которым показана очная консультация;
- врачи-эксперты отделений ККБ. Проводят предварительный анализ сопроводительной документации на пациента;
- кураторы районов Красноярского края. Организуют взаимодействие других категорий пользователей и решают вопросы, требующие особого внимания.

Схема взаимодействия пользователей системы «Виртуальная поликлиника» приведена на рис. 7.

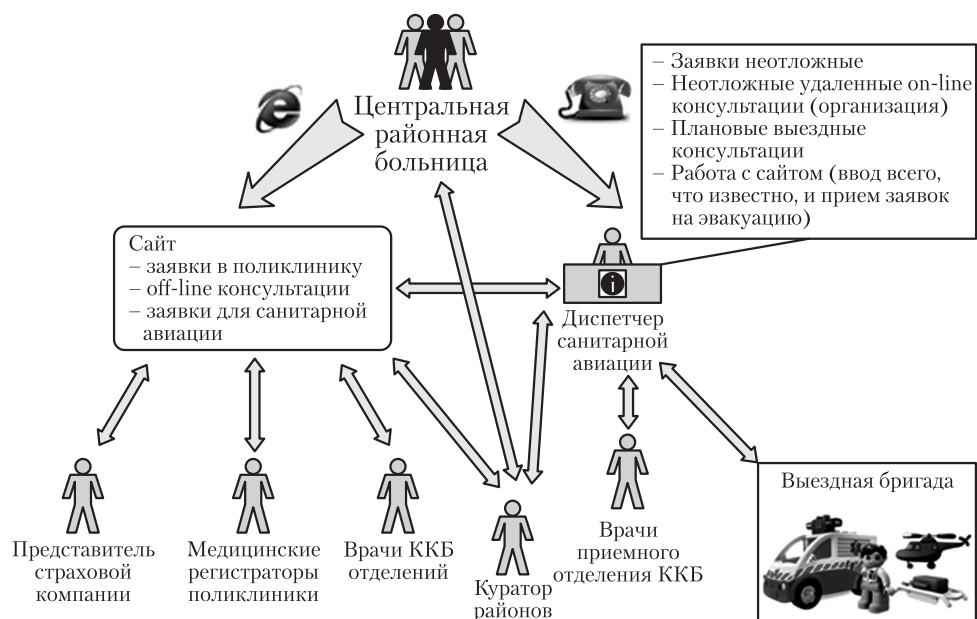


Рис. 7 Схема взаимодействия пользователей

3.3 Состав и функции системы

Система «Виртуальная поликлиника» реализует все перечисленные в разд. 1 данной статьи функции, включая:

- поддержку заочного консультирования с текстовым обменом сообщениями между врачом-экспертом и лечащим врачом удаленной МО;
- построение маршрутной карты пациента с оптимизацией времени приема;
- формирование обширного перечня отчетной документации, в том числе результатов мониторинга доступности консультативной помощи ККБ;
- организация оперативной связи со всеми 134 МО Красноярского края для обмена необходимой информацией методического характера.

Формируемые системой результаты мониторинга доступности консультативной помощи ККБ (рис. 8) в реальном масштабе времени дают ответ на вопрос, когда (по истечении скольких дней) пациент попадет на прием к указанному врачу-специалисту, если заявка будет подана в настоящий момент времени. Эти результаты могут быть использованы всеми пользователями системы.

Система «Виртуальная поликлиника» включает следующие подсистемы:

- подсистема формирования заявки;
- подсистема заочного консультирования;

5.9	Среднее ожидание (дней)	21.11.2015 21:35:57	Количество неотвеченных заявок:	580			
3	Аллерголог	7	Гастроэнтеролог	10	Гастроэнтеролог-хирург	9	Гематолог
10	Гинеколог	0	Гинеколог-эндокринолог	6	Иммунолог	3	Инвазивный кардиолог
10	Кардиоаритмолог	3	Кардиолог	3	Кардиохирург	7	ЛОП
1	Микрохирург	2	Невропатолог	7	Нейрохирург	6	Нефролог
1	Ожоговый хирург	6	Ортопед	8	Подиатр	10	Проктолог
1	Профпатолог	3	Пульманолог	3	Ревматолог	7	Сосудистый хирург
6	Сурдолог	1	Терапевт Центра остеопороза	8	Травматолог	7	Уролог
13	Эндокринолог	13	Эндокринолог-хирург	9	Эпилептолог		

Рис. 8 Результаты мониторинга доступности консультативной помощи

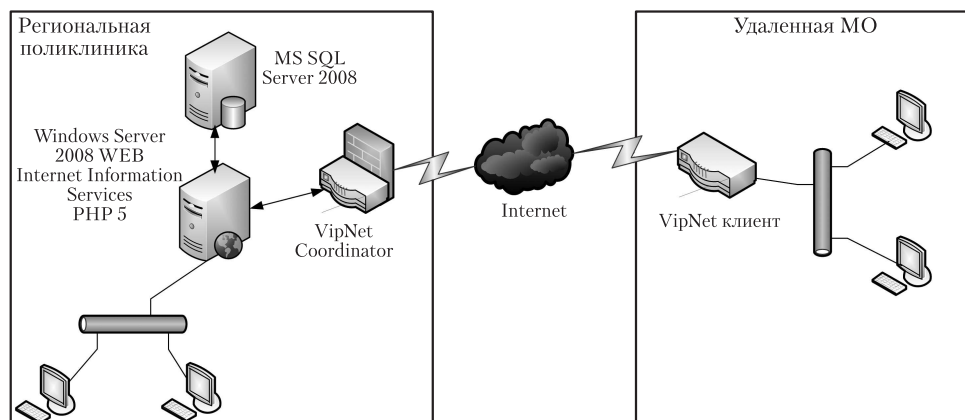


Рис. 9 Схема взаимодействия составных частей системы «Виртуальная поликлиника»

- подсистема формирования отчетов и итогов мониторинга доступности;
- подсистема сверки полиса обязательного медицинского страхования;
- подсистема формирования маршрутной карты пациента для очного консультирования в ККБ.

Система реализована в виде веб-сервиса на языке программирования PHP версии 5. Использованы система управления базами данных MS SQL Server 2008, веб-сервер IIS под управлением операционной системы Windows Server 2008 редакции WEB. Защита передачи персональных данных при передаче через сеть Интернет обеспечена с помощью программных и аппаратных средств VipNET. Взаимодействие основных составляющих системы изображено на рис. 9.

3.4 Представление данных

Физическая модель данных, использованная в системе «Виртуальная поликлиника», представлена на рис. 10.

Эта модель данных позволяет хранить паспортные данные и сопроводительную медицинскую документацию пациента, справочник МО, данные о различных пользователях системы, данные о переписке между врачами удаленных МО и консультантами, выданные талоны на прием, маршруты пациента и другую информацию.

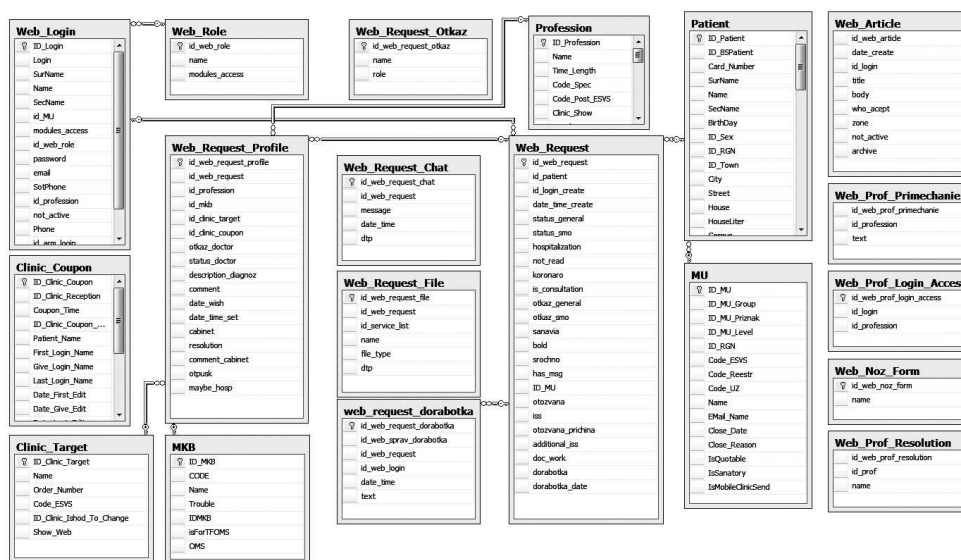


Рис. 10 Физическая модель данных

4 Результаты внедрения

На сегодняшний день система «Виртуальная поликлиника» внедрена в 134 МО Красноярского края. Внедрение системы потребовало введения единой терминологии. Были разработаны новые стандарты оформления результатов обследования по различным нозологиям и целям консультирования, которые были утверждены Министерством здравоохранения Красноярского края.

Для обеспечения заочного консультирования и обработки заявок на консультирование в настоящее время задействованы 20 врачей-экспертов из числа специалистов ККБ. В их задачи входит: принятие решения о необходимости очной консультации, в случае возможности — проведение заочной консультации, формирование рекомендаций для лечащего врача удаленной МО, в случае необходимости — принятие решения о срочной госпитализации.

С помощью системы «Виртуальная поликлиника» начиная с 2012 г. ежегодно проводятся консультации по десяткам тысяч уникальных пациентов, т. е. без учета повторных обращений этих пациентов. На рис. 11 представлены объемные показатели по оказанию консультативной медицинской помощи за 10 мес. каждого года (для корректного сравнения по годам). Эти показатели свидетельствуют о востребованности и эффективности разработанной системы.

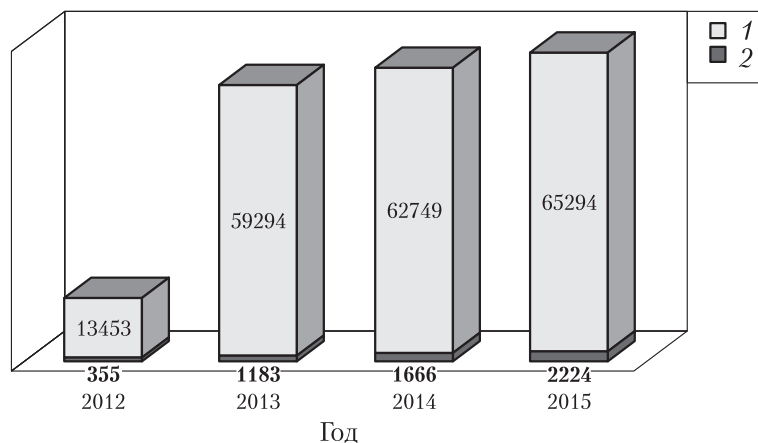


Рис. 11 Число уникальных пациентов, получивших консультативную помощь через систему «Виртуальная поликлиника»: 1 — количество заявок; 2 — количество заочных консультаций

5 Заключение

Внедрение системы «Виртуальная поликлиника» позволило:

- уменьшить число необоснованных очных посещений пациентов региональной поликлиники за счет заочного консультирования (доля заочных консультаций от общего количества поданных заявок составила 3,4%);

– сократить время пребывания пациентов в региональной поликлинике (для 42% пациентов посещение поликлиники ограничилось лишь одним днем).

Эксплуатация системы «Виртуальная поликлиника» показала, что необходимо дальнейшее ее развитие в следующих направлениях:

- электронное сопровождение пациента по маршруту, указанному в маршрутной карте, и при необходимости корректировка маршрутной карты;
- расширение функций личного кабинета пациента ККБ с возможностью доступа к системе «Виртуальная поликлиника» по согласованию маршрутной карты.

Это будет способствовать дальнейшему расширению доступности и оперативности оказания узкоспециализированной медицинской помощи для пациентов региональной консультативной поликлиники Красноярского края.

Литература

1. Ассоциация развития медицинских информационных технологий (АРМИТ). <http://www.armit.ru>.
2. Дубровин А. А., Жилина Н. М. Проблемы и пути решения организации доступа населения к информационным ресурсам в лечебно-профилактических учреждениях при внедрении комплексной медицинской информационной системы записи к врачу через Интернет // Врач и информационные технологии, 2012. № 4. С. 56–61.
3. Гулиев Я. И. Основные аспекты разработки медицинских информационных систем // Врач и информационные технологии, 2014. № 5. С. 10–19.
4. Марка Д. А., МакГоуэн К. Методология структурного анализа и проектирования SADT. — М.: Метатехнология, 1993. 231 с.
5. Методология функционального моделирования. — М.: Изд-во стандартов, 2001. 50 с.
6. Ириков В. А., Тренев В. Н. Распределенные системы принятия решений. — М.: Наука, 1999. 285 с.
7. Бурков В. Н., Коргин Н. А., Новиков Д. А. Введение в теорию управления организационными системами. — М.: Либроком, 2009. 264 с.
8. Паклин Н. Б., Орешков В. И. Бизнес-аналитика: от данных к знаниям. — СПб.: Питер, 2013. 706 с.
9. Об основах охраны здоровья граждан в Российской Федерации: Федеральный закон № 323-ФЗ от 21.11.2011. <http://www.rg.ru/2011/11/23/zdorovie-dok.html>.
10. Илюшин Г. Я., Лиманский В. И. Построение системы управления потоками пациентов // Системы и средства информатики, 2015. Т. 25. № 1. С. 186–197.

Поступила в редакцию 22.12.15

STRUCTURAL AND FUNCTIONAL MODEL OF PATIENT FLOW CONTROL IN A REGIONAL CONSULTATION CLINIC

V. G. Azanov

IT Department, Krasnoyarsk Regional Clinic Hospital, 3A Partizana Zheleznyaka Str., Krasnoyarsk 660022, Russian Federation

Abstract: The paper deals with the problem of the availability and efficiency of advising highly specialized care for patients of regional health centers, serving a vast region of the Russian Federation with low average density of population by the example of the Krasnoyarsk region. It is proposed to solve this problem with the help of a medical information system “Virtual clinic,” which implements the following principles of flow control of remote patients: differentiation of patients by type of advice (full-time or part-time consulting), distance forming a route sheet to provide full-time advice to the patient of intraregional clinics, and remote distance counseling. The paper presents the rationale of the proposed approach as a result of the structural and functional domain modeling methodology SADT (Structured Analysis and Design Technique). The paper contains the results of the simulation of patient flow control in the form of a data flow diagram in a graphical notation IDEF0 AS IS and TO BE. The paper reveals the features of the remote absentee counseling and individual route planning of the patient in the clinic at the regional level by the example of the Krasnoyarsk Regional Hospital. The paper contains the description of the “Virtual clinic,” which reflects the composition and functions of the system, data presentation, user category, methodical and financial aspects of implementation, and monitoring tools accessibility provide advice. The paper describes the experience of the introduction of the “Virtual clinic” medical information system in the regional consultative clinics and 134 medical organizations of the Krasnoyarsk region sending patients for consultation to the regional hospital.

Keywords: medical information systems; remote consulting; patient flow control; individual patient routing; virtual clinic

DOI: 10.14357/08696527160102

References

1. Assotsiatsiya razvitiya meditsinskikh informatsionnykh tekhnologiy (ARMIT) [Association of medical information technology development]. Available at: <http://www.armit.ru/> (accessed February 10, 2011).
2. Dubrovin, A. A., and N. M. Zhilina. 2012. Problemy i puti resheniya organizatsii dostupa naseleniya k informatsionnym resursam v lechebno-profilakticheskikh uchrezhdeniyakh pri vnedrenii kompleksnoy meditsinskoy informatsionnoy sistemy zapisi k vrachu cherez Internet [Problems and ways of organization of access to information resources in health-care institutions in the implementation of an integrated health

- information system records to the doctor over the Internet]. *Vrach i Informatsionnye Tekhnologii* [Doctor and Information Technology] 4:56–61.
3. Guliev, Y. I. 2014. Osnovnye aspekty razrabotki meditsinskikh informatsionnykh sistem [Basic aspects of medical information systems development]. *Vrach i Informatsionnye Tekhnologii* [Doctor and Information Technology] 5:10–19.
 4. Marka, D. A., and M. G. Klement, eds. 1993. *Metodologiya strukturnogo analiza i proektirovaniya SADT* [Methodology of structure analysis and design SADT]. Moscow: Meta Technology. 231 p.
 5. RD 50.1.028-2001. 2001. Metodologiya funktsional'nogo modelirovaniya [Methodology of functional modeling]. Moscow: IPK Standart Publs. 50 p.
 6. Irikov, V. A., and V. N. Trenev, eds. 1999. *Raspredelemnnye sistemy prinyatiya resheniy* [The distributed decision marketing systems]. Moscow: Nauka. 285 p.
 7. Burkov, V. N., N. A. Korgin, and D. A. Novikov. 2009. *Vvedenie v teoriyu upravleniya organizatsionnymi sistemami* [Introduction to theoretical management organization systems]. Moscow: Librokom. 264 p.
 8. Paklin, N. B., and V. I. Oreshkov, eds. 2013. *Biznes-analitika: Ot dannykh k znaniyam* [Business-analytics: From data to knowledge]. St. Petersburg: Piter. 706 p.
 9. Ob osnovakh okhrany zdorov'ya grazhdan v Rossiyskoy Federatsii: Federal'nyy zakon No.323-FZ ot 21.11.2011 [About basis of health of citizens in the Russian Federation: Federal Law No.323-FZ of 21.11.2011]. Available at: <http://www.rg.ru/2011/11/23/zdorovie-dok.html> (accessed December 10, 2015).
 10. Ilyushin, G. Y., and V. I. Limanskiy. 2015. Postroenie sistemy upravleniya potokami patsientov [Creation of patient flow control system]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 25(1):186–197.

Received December 22, 2015

Contributor

Azanov Vitaliy G. (b. 1983) — manager of IT department, Krasnoyarsk Regional Clinic Hospital, 3A Partizana Zheleznyaka Str., Krasnoyarsk 660022, Russian Federation; azanov@mail.ru

АЛГОРИТМ И ПРОГРАММНЫЙ КОМПЛЕКС ДЛЯ АНАЛИЗА ХАРАКТЕРИСТИК ВЫХОДНОГО ИЗЛУЧЕНИЯ ПЛАЗМЕННОГО РЕЛЯТИВИСТСКОГО ГЕНЕРАТОРА СВЕРХВЫСОКОЙ ЧАСТОТЫ*

*С. Е. Андреев*¹

Аннотация: Предложена методика обработки результатов эксперимента по генерации излучения сверхвысокой частоты (СВЧ) плазменным релятивистским СВЧ-генератором. Для получения информации о спектре полученного СВЧ-импульса с высоким разрешением по частоте или времени предлагается совместное использование вейвлет-анализа и цифровой фильтрации. Программа написана в среде программирования MATLAB 7.0. Шаг дискретизации сигнала, исходящего от генератора, равен 50 пс. Плазменный релятивистский генератор (ПРГ) на ускорителе Sinus 550-80 обладает рядом уникальных свойств: работа в режиме генерации одиночного импульса или в частотно-периодическом режиме (до 50 имп./с); генерация импульсов СВЧ-излучения длительностью до 80 нс с частотами излучения от 2 до 25 ГГц; генерация как монохроматического излучения, так и излучения с многочисленными гармониками в спектре. Начальными параметрами, в частности плотностью плазмы, можно изменять режим работы ПРГ и этим регулировать характеристики излучения: центральные частоты, ширину спектра и т. д. Для выбора перечисленных режимов генерации поставлена задача — исследование зависимости поведения спектров выходного излучения от времени.

Ключевые слова: системы обработки численных данных; алгоритмы; временной ряд; преобразование Фурье; вейвлет-преобразование; СВЧ-излучение; релятивистский плазменный СВЧ-генератор

DOI: 10.14357/08696527160103

1 Введение

В последнее время большое внимание уделяется разработке и созданию мощных СВЧ-генераторов с широким диапазоном частот излучения. Одно из решений данной задачи — использование плазменных релятивистских источников СВЧ-излучения.

Согласно теории и экспериментальным исследованиям [1–3], плазменный релятивистский СВЧ-генератор обладает возможностью перестройки частоты СВЧ-излучения в широком диапазоне. В ранних теоретических работах не

* Работа выполнена при поддержке РФФИ (проект 14-08-01126-а).

¹ Институт общей физики РАН; Московский государственный университет информационных технологий, радиотехники и электроники, funkmonk@rambler.ru

указывался характер спектра выходного сигнала ПРГ, однако в работе [4] экспериментально наблюдался дискретный спектр излучения плазменного СВЧ-генератора.

Регулировка частоты выходного сигнала ПРГ происходит за счет изменения плотности плазмы. В частности, в работе [5] удалось добиться перестройки частоты в 7 раз (от 4 до 28 ГГц) на уровне мощности около 50 МВт за промежутки времени 30 мкс благодаря изменению плотности плазмы. Однако в процессе работы (генерации) ПРГ происходят изменения параметров генерации и вследствие этого меняются характеристики выходного СВЧ-излучения. Для более глубокого понимания физических процессов, протекающих в ПРГ во время генерации, и оценки возможности их использования необходимо проанализировать влияние начальных параметров эксперимента на выходное излучение ПРГ. Для этого надо получить спектры выходного СВЧ-излучения, показывающие зависимость спектральной плотности не только от частоты, но и от времени.

В данной статье описан метод и программа для проведения такого анализа: алгоритмы, схема сбора экспериментальных данных и пример результатов обработки. Исследование спектров выходного СВЧ-излучения с высоким разрешением по частоте и времени позволяет определять гармонический состав излучения и динамику отдельных гармоник. В основу метода получения динамических спектров положена цифровая фильтрация и вейвлет-преобразование.

2 Метод получения и структура экспериментальных данных

В зависимости от режима работы на выходе плазменного релятивистского СВЧ-генератора регистрируется либо одиночный СВЧ-импульс, либо серия СВЧ-импульсов (в частотно-периодическом режиме работы). Регистрация излучения производится широкополосной вибраторной антенной. Схема эксперимента приведена в [6]. Для измерений в двух диапазонах частот 1,5–5 и 5–25 ГГц использовались вибраторные антенны, различающиеся только геометрическими параметрами. Получаемый сигнал после прохождения по коаксиальной линии передачи и ослабления аттенюаторами регистрируется осциллографом Tektronix TDS-7404/Tektronix DPO-71604c и передается на ЭВМ для последующей обработки. Схема сбора данных приведена на рис. 1.



Рис. 1 Схема сбора экспериментальных данных

Характеристики осциллографов

Осциллограф	Полоса, ГГц	Число выборок, $10^{10}/с$	Шаг дискретизации, пс	АЦП, бит
TDS 7404 [7]	4	2	50	8
DPO 71604с [8]	16	5	20	8

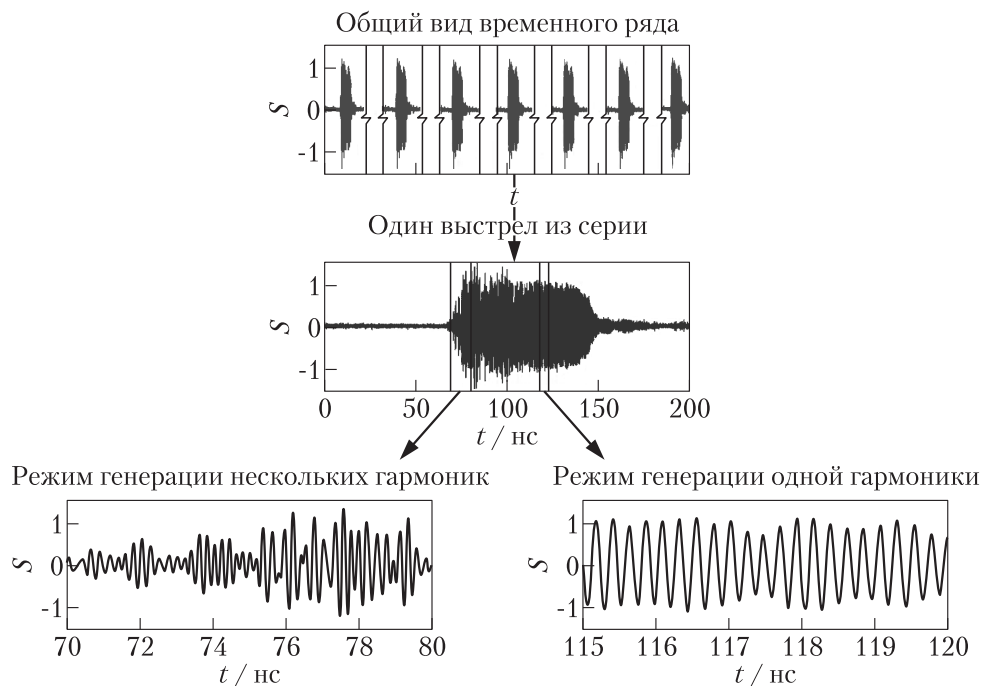


Рис. 2 Структура экспериментальных данных (вид представления выходного излучения ПРГ)

Характеристики осциллографов представлены в таблице.

В результате использования данной схемы сбора данных было получено множество временных рядов, описывающих выходное излучение ПРГ. В частотно-периодическом режиме работы ПРГ количество СВЧ-импульсов в одной серии равнялось 20 (установка позволяет варьировать их от 1 до 50). Вследствие заданного шага дискретизации используемых осциллографов (Tektronix TDS-7404/Tektronix DPO-71604с) количество точек, отводимое под один импульс, составило 4000 или 8000. Структура экспериментальных данных для серии импульсов представлена на рис. 2.

3 Методы обработки экспериментальных данных

Выходное излучение ПРГ представляет собой широкополосный сигнал с некратными гармониками (до 20) в спектре. Для исследования его динамических характеристик были использованы спектры с разрешением по времени, а их параметры определялись исходя из черенковского взаимодействия плазмы и релятивистского электронного пучка (РЭП), т. е. физических особенностей ПРГ.

Частотный интервал между соседними гармониками спектра СВЧ-излучения определяется параметрами эксперимента, а именно: скоростью электронов и длиной взаимодействия РЭП с плазмой. Данный частотный интервал в исследуемой серии экспериментов составлял от 100 до 300 МГц.

После тестирования множества методов обработки экспериментальных данных были выбраны наиболее оптимальные для конкретных данных с последующей реализацией в среде программирования MATLAB:

- (1) цифровая фильтрация,
- (2) вейвлет-преобразование.

Метод получения спектров с разрешением по времени с использованием цифровой фильтрации основан на использовании набора фильтров с импульсными характеристиками

$$h_n(f, t) = |\text{sinc}(10^8 t)| \sin(2\pi ft), \quad (1)$$

где f — центральная частота полосы фильтра (частота сканирования); t — время.

Далее в порядке возрастания частоты сканирования производится свертка этих импульсных характеристик с регистрируемым сигналом:

$$W(f, t) = x(t) * h(f, t), \quad (2)$$

где $x(t)$ — СВЧ-импульс.

Используемые цифровые фильтры хорошо локализованы в частотной области, что позволяет получить хорошее разрешение по частоте. Постоянное значение полосы пропускания во всем диапазоне частот, так же как и в оконном преобразовании Фурье, сохраняет разрешающую способность по частоте во всем диапазоне частот. Кроме того, в отличие от оконного преобразования Фурье, этот метод позволяет выделить наиболее стабильные частоты, время изменения амплитуды которых превышает ширину главного лепестка. При вейвлет-преобразовании разрешающая способность зависит от выбранного масштаба базисной функции. Поэтому была выбрана фильтрация сигнала, а не вейвлет-преобразование. Временная диаграмма импульсной характеристики фильтра для частоты сканирования 2 ГГц представлена на рис. 3.

Амплитудно-частотные характеристики цифрового фильтра представлены на рис. 4.

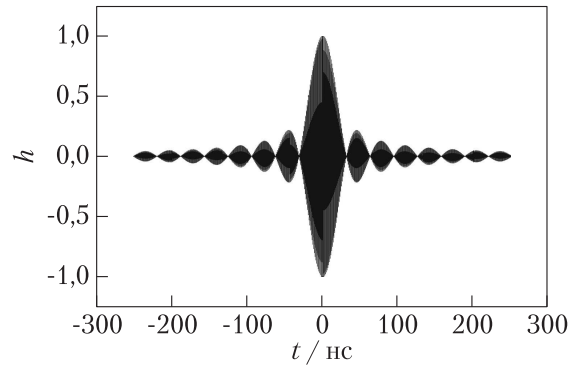


Рис. 3 Временная диаграмма импульсной характеристики фильтра с центральной частотой 2 ГГц

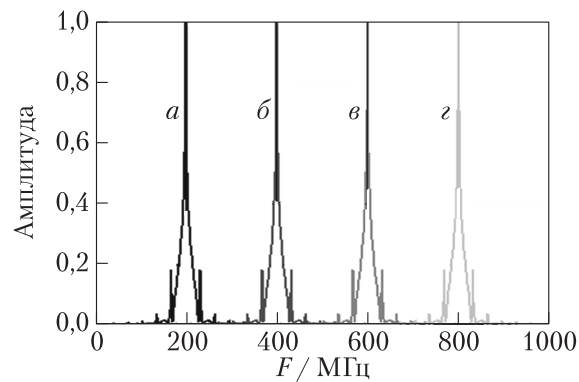


Рис. 4 Амплитудно-частотные характеристики используемого фильтра для разных частот сканирования: *а* — 200 МГц; *б* — 400; *в* — 600; *г* — 800 МГц

Второй метод получения спектров с разрешением по времени основан на использовании вейвлет-преобразования с базисной функцией:

$$\Psi_M(\omega, t) = e^{-(\omega t)^2/40} \sin(5\omega t) \quad (3)$$

или

$$\Psi_M(\omega, t) = e^{-(\omega t)^2/40} e^{j5\omega t} \quad (4)$$

(опционально). Эти базисные функции используемого вейвлета представляют из себя вариации вейвлета Морле. Вейвлет-преобразование более подробно описано в [9, 10]. Базисная функция (3) используется для ускорения работы программы, так как требует в два раза меньше вычислений, чем базисная функция (4).

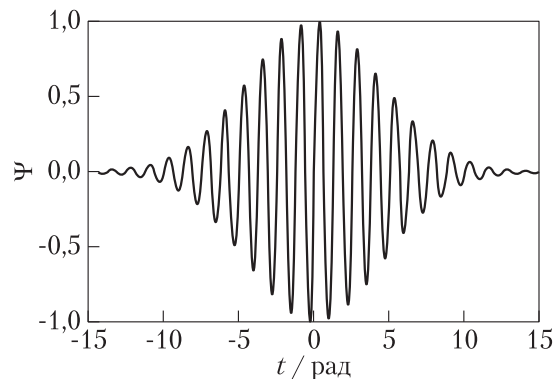


Рис. 5 Базисная функция используемой модификации вейвлета Морле

У них по сравнению с вейлетом Морле изменен коэффициент спада. Данные базисные функции отвечают условиям допустимости, ограниченности и условию самоподобия, а также отвечают условию локализации. Таким образом, они полностью удовлетворяют всем условиям, необходимым для вейвлет-преобразования.

Временная диаграмма базисной функции представлена на рис. 5.

4 Программа обработки экспериментальных данных

Разработанная программа предназначена для расчета динамического спектра с помощью фильтрации и дальнейшего вейвлет-преобразования. Для начала работы необходимо указать путь к необходимым файлам для обработки, задать число точек, отводимое на один выстрел, и шаг дискретизации (если в файле отсутствуют данные о временных характеристиках). Как результат этого шага появляются графики самого временного ряда и спектр, полученный прямым преобразованием Фурье. Переключение между этими графиками осуществляется с помощью переключателей. Далее производится задание начальных параметров обработки, таких как

- (1) число точек, отводимое на каждый импульс;
- (2) начальная частота обработки;
- (3) конечная частота обработки;
- (4) шаг частотной сетки;
- (5) необходимость сглаживания спектра для цифровой фильтрации;
- (6) вид базисной функции вейвлета;
- (7) необходимость сглаживания спектра для вейвлет-преобразования.

После задания начальных параметров обработки данных производится чтение одного из файлов с временными рядами и производится его разбиение на отдельные импульсы. Для серии импульсов разбиение производится по заранее заданному числу точек, отводимому на один временной ряд. В случае одиночного импульса разбиение пропускается — программа переходит непосредственно к обработке. После получения временных рядов к ним применяются методы обработки, описанные выше. Эти методы адаптированы к компьютерным вычислениям путем дискретизации. Свертка для цифровой фильтрации имеет вид:

$$\gamma_{k,j} = \sum_{l=j-\Delta t}^{j+\Delta t} x_l h_k(l dt - j dt),$$

где k — номер элемента по частотной сетке; j — временной сдвиг свертки; Δt — ширина фильтра; dt — шаг дискретизации; x — временной ряд; $h_k(t)$ — фильтр из набора (1).

Далее производится дискретное вейвлет-преобразование. Преобразование осуществляется сверткой, полученной путем адаптации формулы непрерывного вейвлет-преобразования к дискретному варианту:

$$\gamma_{k,j} = \sum_{l=j-\Delta t}^{j+\Delta t} x_l \Psi(2\pi dfk(l dt - j dt)),$$

где k — номер элемента по частотной сетке; j — временной сдвиг вейвлета; Δt — ширина вейвлета; df — шаг частотной сетки; dt — шаг дискретизации; x — временной ряд; $\Psi(t) = e^{-t^2/40} \sin(5t)$ — вейвлет (3).

Из получаемых спектров выделяются максимумы по времени — по ним производится линейная аппроксимация. После окончания вычисления спектра одного импульса строится контурный график с последующей записью в файл и начинает обрабатываться следующий импульс. После обработки всей серии импульсов из одного файла программа переходит к следующему файлу, если он задан.

Алгоритм данной программы представлен на рис. 6. Алгоритм получения спектра одного импульса с помощью вейвлет-преобразования изображен на рис. 7. Алгоритм получения спектра методом фильтрации аналогичен алгоритму получения спектров с помощью вейвлет-преобразования. Интерфейс данной программы представлен на рис. 8.

В результате обработки серии экспериментов по генерации СВЧ-излучения ПРГ было получено множество результатов, некоторые из которых представлены на рис. 9. Из анализа этих графиков видно, что

- предлагаемая фильтрация позволяет получить хорошее разрешение по частоте (40 МГц), выделяет наиболее стабильные составляющие спектра, однако подавляет компоненты с быстро изменяющейся частотой;

Главная программа

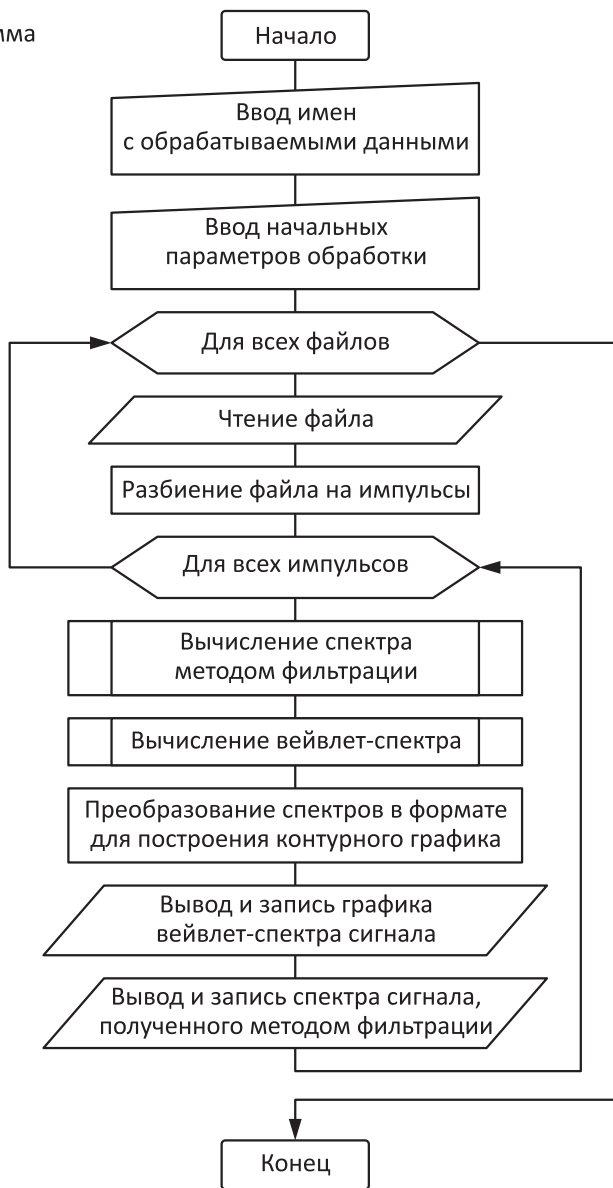


Рис. 6 Главная программа

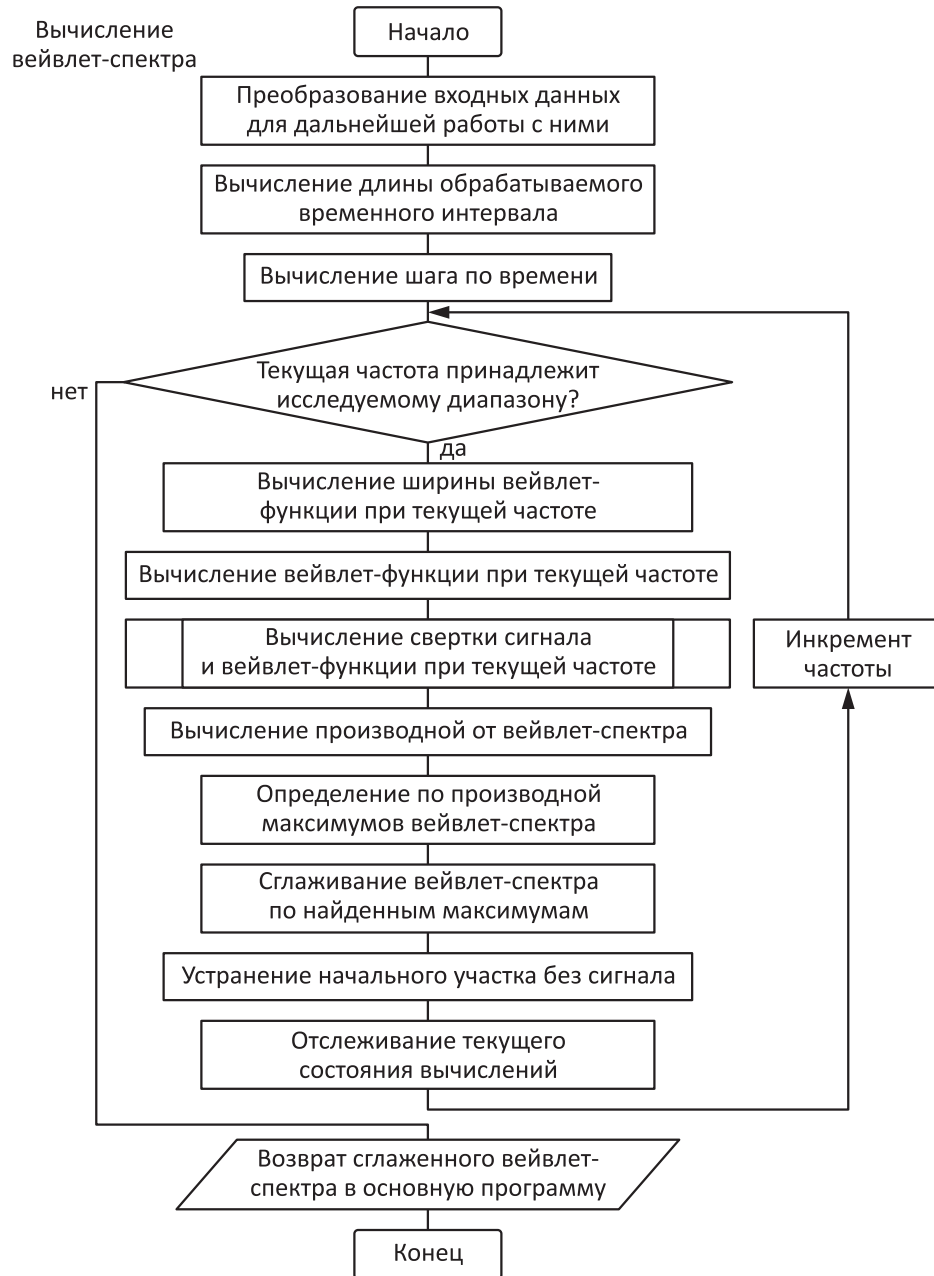


Рис. 7 Процесс получения вейвлет-спектра

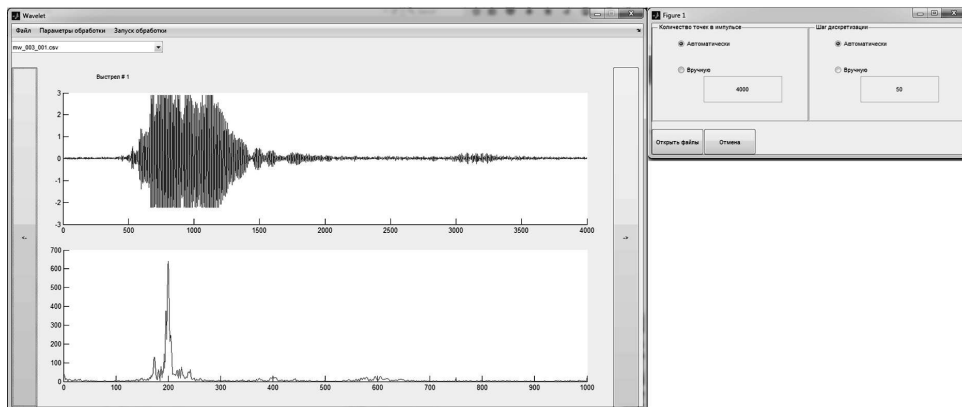


Рис. 8 Интерфейс программы обработки

- предлагаемая модификация вейвлета Морле обладает хорошим разрешением по времени (36 нс), что позволяет определять моменты начала и окончания процессов для каждой компоненты спектра и изменение в их амплитудах. Однако из-за плохого разрешения по частоте «размывает» линии спектра на графике, что не позволяет их идентифицировать по частоте.

Совместное использование обоих методов позволяет с хорошим разрешением по времени и частоте получить информацию о поведении спектра во времени.

5 Заключение

Благодаря разработанной программе удалось проанализировать временную динамику спектра выходного сигнала ПРГ, увидеть его гармонический состав, а также время начала и окончания генерации присутствующих в спектре гармоник. В результате обработки с помощью предложенной программы нескольких тысяч СВЧ-импульсов, полученных в ходе экспериментов с ПРГ в различных режимах работы, установлено:

- при одинаковых начальных условиях эксперимента ПРГ может генерировать излучение с устойчивыми параметрами от импульса к импульсу;
- при использовании ПРГ можно плавно менять частоту выходного СВЧ-излучения (см., например, рис. 9, а);
- ширина спектра СВЧ-импульсов может варьироваться от 50 МГц до 2 ГГц (см., например, рис. 9) в зависимости от начальных условий генерации, что говорит о возможности эффективного управления шириной спектра излучения.

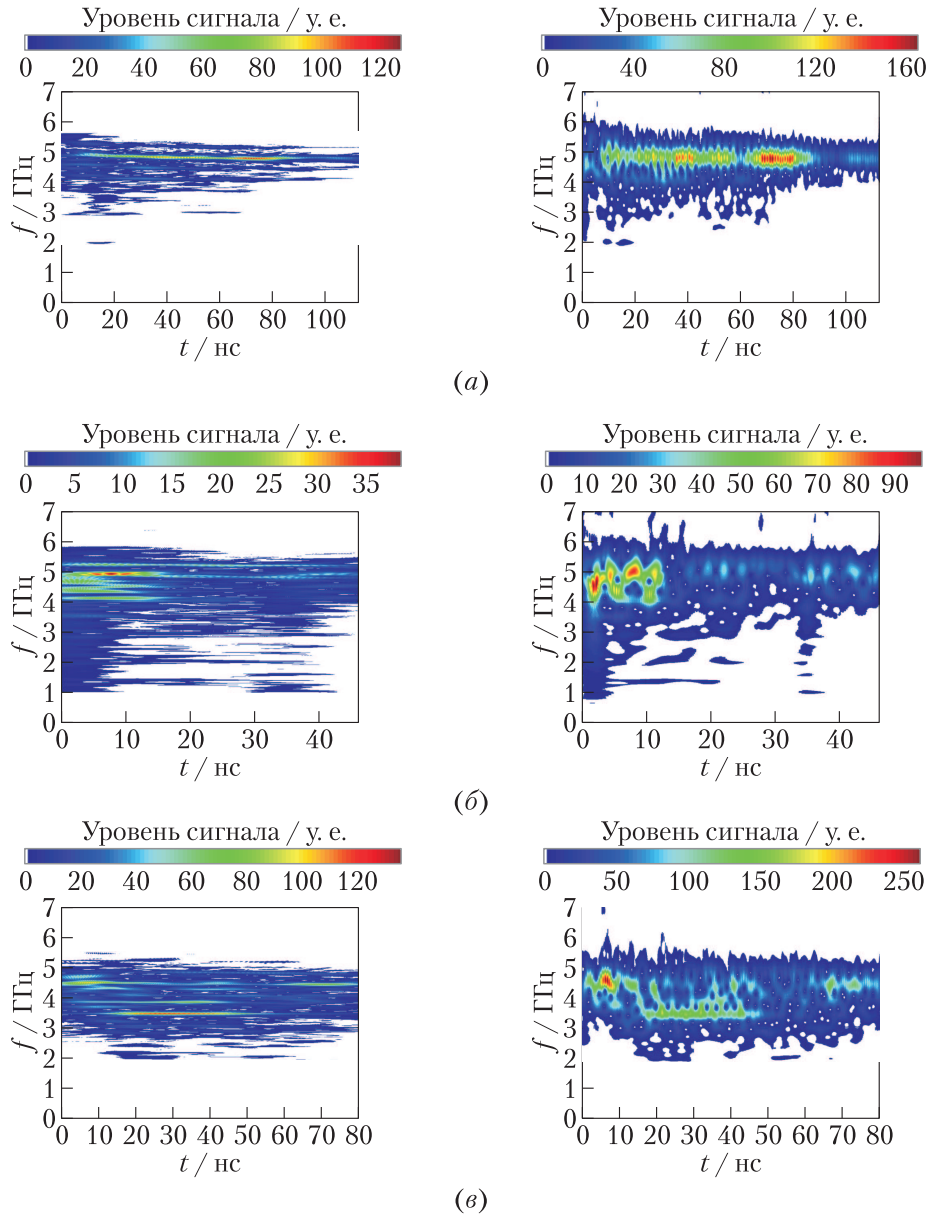


Рис. 9 Пример спектров, полученных в результате применения разработанной программы: (а) одночастотный режим генерации; (б) переход от многочастотного режима генерации к режиму генерации двух гармоник; (е) многочастотный режим генерации. Слева — спектры, полученные методом фильтрации; справа — спектры, полученные методом вейвлет-преобразования

Автор выражает благодарность Н. Г. Гусейну-заде, Н. Н. Скворцовой и Д. К. Ульянову за активную помощь.

Литература

1. Кузелев М. В., Рухадзе А. А., Стрелков П. С., Шкварунец А. Г. Релятивистская сильноточная плазменная СВЧ-электроника: преимущества, достижения, перспективы // Физика плазмы, 1987. Т. 13. № 11. С. 1370–1382.
2. Кузелев М. В., Лоза О. Т., Рухадзе А. А., Стрелков П. С., Шкварунец А. Г. Плазменная релятивистская СВЧ-электроника // Физика плазмы, 2001. Т. 27. № 8. С. 710–733.
3. Кузелев М. В., Рухадзе А. А., Стрелков П. С. Плазменная релятивистская СВЧ-электроника // Под ред. А. А. Рухадзе. — М.: МГТУ им. Н. Э. Баумана, 2002. 543 с.
4. Богданкевич И. Л., Иванов И. Е., Лоза О. Т., Рухадзе А. А., Стрелков П. С., Тараканов В. П., Ульянов Д. К. Тонкая структура спектров излучения плазменного релятивистского СВЧ-генератора // Физика плазмы, 2002. Т. 28. № 8. С. 748–757.
5. Ульянов Д. К. Спектры плазменного релятивистского СВЧ-генератора: Дис. . . . канд. физ.-мат. наук. — М.: ИОФ РАН, 2000.
6. Андреев С. Е., Ульянов Д. К. Метод управления спектром плазменного релятивистского СВЧ-генератора в частотно-периодическом режиме // Прикладная физика, 2014. № 4. С. 26–30.
7. Digital Phosphor Oscilloscopes: TDS7000 Series. — Advanced Test Equipment Rentals, 2001. <http://www.atecorp.com/ATECorp/media/pdfs/data-sheets/TektronixTDS7154.Datasheet.pdf>.
8. DPO/DSA70000 Series Data Sheet: Digital Phosphor Oscilloscopes/Digital Serial Analyzers. — Tektronix, 2008. <http://www.testequipmentconnection.com/specs/71604.pdf>.
9. Mallat S. A wavelet tour of signal processing. — Amsterdam: Elsevier, 2008. 832 p.
10. Скворцова Н. Н., Шестаков О. В., Малахов Д. В. Методы численного анализа стохастических сигналов: Лекции по курсу «Прикладная радиофизика». — М.: МИРЭА, 2011. 108 с.

Поступила в редакцию 24.08.15

ALGORITHM AND PROGRAM COMPLEX FOR ANALYSIS OF THE CHARACTERISTICS OF PLASMA RELATIVISTIC ULTRAHIGH FREQUENCY GENERATOR OUTPUT RADIATION

S. E. Andreev^{1,2}

¹Prokhorov General Physics Institute, Russian Academy of Sciences, 38 Vavilov Str., Moscow 119991, Russian Federation

²Moscow Technological University (MIREA), 78 Vernadskogo Ave., Moscow 119454, Russian Federation

Abstract: The article proposes a method for processing experimental results obtained from plasma relativistic ultrahigh frequency (UHF) generator (PRG). For gathering information about a PRG output radiation spectrum with high time and frequency resolution, the authors use wavelet transformation and digital filtration of signal, presenting output radiation of PRG. The program has been written in the MATLAB language in the MATLAB 7.0 environment. The sampling interval of the PRG output signal is equal to 50 ps. The plasma relativistic UHF generator, built on Sinus 550-80 accelerator, has some unique properties: PRG has a single impulse mode and a multiple impulses mode (up to 50 pulses per second); PRG can generate pulses with durations of 80 ns and frequency band lying in 2 to 25 GHz area of spectrum; PRG can generate an UHF radiation in single frequency and multifrequency modes. By alteration of starting parameters such as plasma density, one can switch between different PRG work modes and, by that, one can regulate some of radiation characteristics such as carrier frequency or bandwidth or other. To make control of PRG output radiation possible, a complex of program and signal-gathering devices were created. The relation between different initial parameters of experiment and PRG output radiation spectral characteristics needs to be investigated.

Keywords: analysis of numerical data; algorithms; time series; Fourier transformation; wavelet transformation; UHF radiation; plasma relativistic UHF generator

DOI: 10.14357/08696527160103

Acknowledgments

The work was supported by the Russian Foundation for Basic Research (project 14-08-01126-a).

References

1. Kuzelev, M. V., A. A. Ruhadze, P. S. Strelkov, and A. G. Shkvarunec. 1987. Relyativistskaya sil'notochnaya plazmennaya SVCh-elektronika: Preimushchestva, dosti-
- 42 *Sistemy i Sredstva Informatiki* — Systems and Means of Informatics 2016 vol 26 no 1

- zheniya, perspektivy [Relativistic high-current plasma UHF-electronics: Advantages, achievements, prospects]. *Plasma Phys.* 13(11):1370–1382.
2. Kuzelev, M. V., A. A. Ruhadze, P. S. Strelkov, A. G. Shkvarunec, and O. T. Loza. 2001. Plazmennaya relyativistskaya SVCh-elektronika [Plasma relativistic UHF-electronics]. *Plasma Phys.* 27(8):710–733.
 3. Kuzelev, M. V., A. A. Ruhadze, and P. S. Strelkov. 2002. *Plazmennaya relyativistskaya SVCh-elektronika* [Plasma relativistic UHF-electronics]. Moscow: Bauman MSTU. 543 p.
 4. Bogdankevich, I. L., I. E. Ivanov, O. T. Loza, A. A. Ruhadze, P. S. Strelkov, V. P. Tarakanov, and D. K. Ul'yanov. 2002. Tonkaya struktura spektrov izlucheniya plazmennogo relyativistskogo SVCh-generatora [The plasma relativistic UHF-generator emission spectrum fine structure]. *Plasma Phys.* 28(8):748–757.
 5. Ul'yanov, D. K. 2000. Spektry plazmennogo relyativistskogo SVCh-generatora [The spectrum of the plasma relativistic UHF-generator]. PhD Diss. Moscow: GPI RAN.
 6. Andreev, S. E., and D. K. Ul'yanov. 2014. Metod upravleniya spektrom plazmennogo relyativistskogo SVCh-generatora v chastotno-periodicheskom rezhime [The spectrum control method of the plasma relativistic UHF-generator in a frequency-periodic work mode]. *Appl. Phys.* 4:26–30.
 7. Digital Phosphor Oscilloscopes: TDS7000 Series. — Advanced Test Equipment Rentals. Available at: <http://www.atecorp.com/ATECorp/media/pdfs/datasheets/TektronixTDS7154.Datasheet.pdf> (accessed February 20, 2016).
 8. DPO/DSA70000 Series Data Sheet: Digital Phosphor Oscilloscopes/Digital Serial Analyzers. Available at: <http://www.testequipmentconnection.com/specs/71604.pdf> (accessed February 20, 2016).
 9. Mallat, S. 2008. A wavelet tour of signal processing. Amsterdam: Elsevier. 832 p.
 10. Skvortsova, N. N., O. V. Shestakov, and D. V. Malakhov. 2011. *Metody chislennogo analiza stokhasticheskikh signalov: Lektsii po kursu "Prikladnaya radiofizika"* [The numerical data analysis methods of stochastic signals: "Applied radiophysics" lecture course]. Moscow: MIREA. 108 p.

Received August 24, 2015

Contributor

Andreev Sergey E. (b. 1990) — junior scientist, Prokhorov General Physics Institute, Russian Academy of Sciences, 38 Vavilov Str., Moscow 119991, Russian Federation; Moscow Technological University (MIREA), 78 Vernadskogo Ave., Moscow 119454, Russian Federation; funkmonk@rambler.ru

ЭКСПЕРИМЕНТАЛЬНЫЙ АНАЛИЗ ОДНОГО МЕТОДА КЛАСТЕРИЗАЦИИ И РАНЖИРОВАНИЯ МНОГОМЕРНЫХ ДАННЫХ НЕЙРОННОЙ СЕТЬЮ КОХОНЕНА

В. И. Аникин¹, О. В. Аникина², А. А. Карманова³

Аннотация: Предложена методика кластеризации и ранжирования данных нейронной сетью (НС) Кохонена, основанная на пространственно-корреляционных свойствах обучающей выборки вне зависимости от используемого алгоритма обучения сети. Экспериментально показана возможность применения метода линейного преобразования координат образцов выборки для кластеризации слабо коррелированных пространственно неразделимых данных. Продемонстрировано использование ранжирования для выделения приграничных образцов и определения степени их близости к соседствующему кластеру, что решает проблему нахождения границ кластеров в пространственно неразделимых данных. Обоснована необходимость многослойной кластеризации в случае неравномерного пространственного распределения данных. Методика кластеризации и ранжирования иллюстрируется примером анализа эмпирических данных о финансовой отчетности российских банков. Методика удобна в применении для выборок малого и среднего объема.

Ключевые слова: многомерная кластеризация; ранжирование; нейронная сеть Кохонена; линейное преобразование; матрица корреляции; методика

DOI: 10.14357/08696527160104

1 Введение

Кластеризация как метод анализа ставит перед собой цель получения структурной оценки данных посредством разделения исходного множества на группы сходных объектов. В методах кластеризации на основе систем искусственного интеллекта широкое распространение получила НС Кохонена [1].

Анализ литературы [2–4] показал, что в настоящий момент отсутствует аргументированный подход к кластеризации данных НС Кохонена в зависимости от пространственных и статистических свойств обучающей выборки. В большинстве исследований, основанных на различных модификациях сети Кохонена, отсутствуют требования предварительного анализа статистических свойств исходных данных и превалирует неоправданно упрощенный подход к поиску решения, что ведет к неверной трактовке полученных результатов из-за недопонимания физического процесса обучения НС. В исследованиях не учитывается, что ста-

¹Поволжский государственный университет сервиса, anikin_vi@mail.ru

²Тольяттинский государственный университет, blue-waterfall@yandex.ru

³ООО «НетКрэкер», turaeva.alexandra@gmail.com

статистические свойства исходной выборки напрямую влияют на процесс обучения сети и порой не позволяют однозначно выполнить кластеризацию, например из-за низкой коррелированности данных, что ведет к зависимости результатов кластеризации от начальных настроек алгоритма обучения. В конечном счете нарушается основополагающий принцип технологии Data Mining [5], утверждающий, что найденные в результате анализа закономерности должны определяться самими данными, а не используемыми методами и условиями проведения экспериментов.

2 Цель и задачи исследования, постановка эксперимента

Цель исследования — экспериментальный анализ применения методики кластеризации и ранжирования НС Кохонена, основанной на пространственно-корреляционных свойствах данных. Предложенная методика позволяет получить обоснованные и надежные результаты, не зависящие от начальных настроек алгоритмов обучения сети Кохонена.

Задачи экспериментального исследования:

1. Разработка методики кластеризации многомерных данных НС Кохонена с учетом пространственно-корреляционных свойств исходной выборки, демонстрация методики на примере кластеризации банков с помощью клеточной сети Кохонена.
2. Разработка методики ранжирования многомерных данных НС Кохонена с учетом пространственно-корреляционных свойств исходной выборки, демонстрация методики на примере ранжирования банков.
3. Сравнение результатов кластеризации банков клеточной НС Кохонена [6, 7] с результатами кластеризации классическим алгоритмом обучения [1], полученными на аналитической платформе Deductor [8].

Эксперименты по кластеризации и ранжированию учебных образцов НС Кохонена проводились на эмпирической выборке, содержащей данные об активах российских банков в виде четырех входных переменных: *Сумма активов*, *Собственные активы*, *Банковские активы*, *Средства в банке*. Объем выборки составлял 99 элементов.

Экспериментальная выборка взята из Deductor, где она используется для обучения самоорганизующихся карт Кохонена с помощью классического алгоритма [1]. Авторские оригинальные эксперименты, реализованные в среде MS Excel применением множественных итераций [9], используют эту выборку для работы с НС Кохонена, обучаемой клеточным автоматом (КА). Реализованная в Excel клеточная НС Кохонена может содержать до 30 связанных нейронов на двумерной решетке размером 5×6 узлов. В описываемых экспериментах активными являются нейроны с номерами 0, 1, 6 и 7, остальные нейроны не задействованы. На рис. 1 показаны использованные конфигурации обучающего КА размером 2×2 узла, из которых особо следует выделить четыре конфигу-

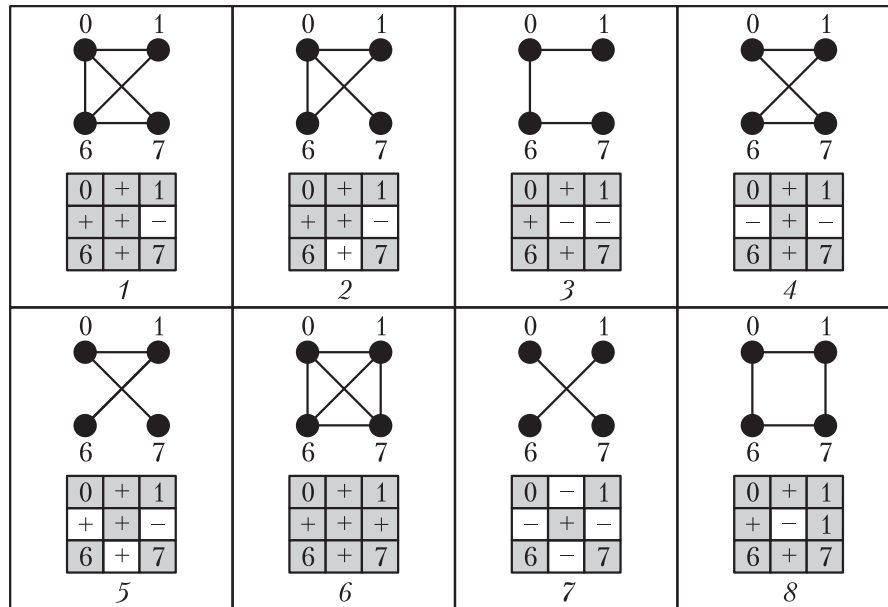


Рис. 1 Возможные конфигурации обучающих КА для НС Кохонена размером 2×2 узла: «+» — связь между нейронами включена, «-» — связь выключена

рации: 3 — одномерный обучающий КА; 6 — КА с окружением Мура; 7 — двусвязный КА; 8 — КА с окружением фон Неймана.

3 Метод линейного преобразования координат образцов обучающей выборки данных

Корреляция как статистическая характеристика наиболее важна при поиске зависимостей в данных. Ниже предлагается метод линейного преобразования координат образцов многомерной выборки, посредством которого удастся существенно повысить корреляцию координат учебных образцов. Высокий уровень корреляции, в свою очередь, гарантирует, что найденные кластеры будут стабильными вне зависимости от начальной настройки обучающего алгоритма.

Прямое и обратное линейное преобразование i -й координаты образца X исходной выборки данных определим соотношениями:

$$X'_i = \frac{(1/\alpha_i) \sum_{k=1}^m \alpha_k X_k - X_i}{m-1};$$

$$X_i = \frac{1}{\alpha_i} \sum_{k=1}^m \alpha_k X'_k - (m-1)X'_i,$$

где m — размерность многомерной выборки данных; i и k — индексы координат ее образцов; α_i — весовые коэффициенты ($\alpha_i \neq 0$) и, как легко показать, имеет место равенство:

$$\sum_{k=1}^m \alpha_k X'_k = \sum_{k=1}^m \alpha_k X_k,$$

т. е. взвешенная сумма координат каждого образца выборки есть инвариант рассматриваемого преобразования.

Применяя прямое линейное преобразование итерационно, получим асимптотическое уравнение

$$X'_i = \frac{(1/\alpha_i) \sum_{k=1}^m \alpha_k X'_k - X'_i}{m-1},$$

или

$$\alpha_i X'_i = \frac{1}{m} \sum_{k=1}^m \alpha_k X'_k.$$

Следовательно, асимптотикой данного итерационного преобразования является прямая линия в m -мерном входном пространстве, задаваемая уравнением:

$$\alpha_1 X'_1 = \alpha_2 X'_2 = \dots = \alpha_m X'_m,$$

причем элементы корреляционной матрицы координат преобразованной выборки данных асимптотически стремятся к ± 1 и без ограничения общности можно считать, что $|\alpha_i| \leq 1$.

С учетом обнаруженного инварианта линейного преобразования рейтинг Rt, определяемый как взвешенная сумма координат каждого образца выборки, определим выражением:

$$Rt = \frac{1}{m} \sum_{k=1}^m \alpha_k X_k = \frac{1}{m} \sum_{k=1}^m \alpha_k X'_k.$$

4 Методика ранжирования и кластеризации многомерных данных нейронной сетью Кохонена

Как отмечалось, новизна предложенной методики заключается в рассмотрении для целей кластеризации НС Кохонена пространственно-статистических свойств выборки как основополагающих критериев выбора плана проведения эксперимента. В методике главенствует такой важный статистический параметр, как корреляция, высокое значение которой является обязательным условием получения стабильных результатов кластеризации НС Кохонена вне зависимости от

применяемого алгоритма обучения последней. Таким образом, если кластеризация выполняется при низкой корреляции данных, то ее результаты не могут быть признаны надежными, а исходные данные требуют предварительной обработки перед экспериментом, например линейного преобразования координат.

Среди важных пространственных характеристик методика выделяет также распределение образцов во входном пространстве данных. Если данные пространственно разделимы, то кластеризация становится достаточно тривиальной задачей для большинства известных алгоритмов [10]. Кластеризация пространственно неразделимых данных является более сложной задачей из-за невозможности проведения четкой границы между кластерами и неоднозначности получаемого результата. Количество выделенных кластеров в этом случае определяется как пространственными характеристиками учебных образцов, так и конфигурацией НС Кохонена. Поэтому для пространственно неразделимых данных важно не столько количество выделенных групп, сколько возможность сравнительной оценки образцов друг с другом, что достигается путем их ранжирования по некоторому объективному критерию, например среднему значению координат. Ранжирование является независимым способом оценки относительно положения образцов как по всей области данных, так и в пределах выделенных кластеров на основе выбранного критерия рейтингования. Описанный выше метод линейного преобразования интересен в применении тем, что он позволяет повысить корреляцию, не меняя при этом порядка ранжирования.

Таким образом, предложенная методика подразумевает, что задача кластеризации пространственно неразделимых данных трансформируется в проведение ранжирования выборки с последующим разделением результатов ранжирования на нужное количество кластеров. После кластеризации сравнение рейтинга образцов позволяет выделить приграничные образцы и оценить степень их близости к соседнему кластеру, решая тем самым проблему нахождения границ кластеров в пространственно неразделимых данных, что служит также проверкой результатов кластеризации. Общее правило корректного группирования, когда каждый кластер сформирован образцами, для которых расстояние до центра своего кластера меньше, чем до центра соседних кластеров [1], работает и для показателя рейтинга, т. е. кластер должен быть сформирован образцами, для которых рейтинг максимально близок к рейтингу своего кластера (рассчитанному как взвешенная сумма рейтингов образцов кластера), а не соседних. Если образец по рейтингу ближе к одному кластеру, а по координатам к другому, то границы кластеров определены неверно.

Наравне с возможностью пространственного разделения второй важной пространственной характеристикой для данной методики выступает плотность распределения образцов. В случае неравномерно распределенных данных эффективно применение метода многослойной кластеризации [11], когда образцы из областей с разной плотностью распределения разделяются на группы в разных слоях сети, что позволяет избежать нецелесообразного дробления небольших кластеров на более мелкие.

Ниже приведена методика ранжирования и кластеризации многомерных данных НС Кохонена с учетом описанного влияния пространственно-статистических свойств данных.

1. Нормирование координат образцов исходной выборки данных к интервалу $[-1, 1]$.
2. Анализ возможности пространственного разделения образцов выборки с использованием двумерных графиков разброса, построенных по наиболее интересным парам координат:
 - в случае возможности пространственного разделения делаем попытку группирования данных (шаг 3);
 - в случае невозможности пространственного разделения всю выборку рассматриваем как одну группу и пропускаем шаг 3.
3. Кластеризация выборки методом группирования [6] с помощью самоорганизующейся карты Кохонена (Self-Organizing Map — SOM) [1]:
 - выбор размера и конфигурации SOM, обучающего алгоритма и его параметров;
 - обучение НС Кохонена и кластеризация данных обученной сетью в нескольких независимых реализациях опыта (при различных начальных параметрах алгоритма обучения) до получения стабильных результатов кластеризации;
 - группирование нейронов SOM с помощью U- и P-матриц Ултша [10] в физически значимые группы;
 - нормирование координат образцов в группах к интервалу $[-1, 1]$.
4. Ранжирование образцов в группах с использованием физически значимого критерия, основанного на взвешенной сумме координат образцов в группе (инварианте линейного преобразования).
5. Анализ пространственного распределения образцов в группах с помощью корреляционной матрицы, перцентилей координат образцов, двумерных графиков разброса, построенных по наиболее интересным парам координат образцов:
 - в случае относительно равномерного распределения образцов выполняется кластеризация выбранным алгоритмом обучения сети Кохонена (шаг 6);
 - в случае существенно неравномерного распределения образцов применяется метод многослойной кластеризации (шаг 7).
6. Проведение кластеризации в группах образцов НС Кохонена:
 - (а) уменьшение при необходимости размерности данных путем отбрасывания наименее коррелированных координат;
 - (б) повышение уровня корреляции образцов методом линейного преобразования для улучшения качества кластеризации;

- (в) выбор обучающего алгоритма сети Кохонена, задание начальных параметров;
 - (г) обучение сети Кохонена и кластеризация данных обученной сетью в нескольких независимых реализациях опыта (при разных начальных параметрах алгоритма обучения) до получения стабильных результатов кластеризации. При необходимости возможно дополнительное повышение корреляции.
7. Проведение многослойной кластеризации в группах включает в себя следующие шаги для каждого слоя кластеризации:
- (а) нормирование координат образцов в каждом слое к интервалу $[-1, 1]$;
 - (б) проведение кластеризации НС Кохонена (шаг 6) малой размерности для выделения не более 4–6 кластеров в одном слое, что необходимо для их отбора в текущем слое;
 - (в) отсеивание кластеров небольших размеров (по требованиям задачи) в каждом слое и добавление их в результаты кластеризации.
8. Проверка согласованности результатов ранжирования и кластеризации образцов в группах:
- в случае их несогласованности эксперимент требует пересмотра, так как это говорит о проблемах в выбранном плане эксперимента;
 - в случае их согласованности кластеризация считается верной и завершенной.

5 Применение методики кластеризации и ранжирования данных нейронной сетью Кохонена

Рассмотрим применение методики ранжирования и кластеризации сетью Кохонена на примере выборки данных из финансовой отчетности банков [8]. В дальнейшем будем интересоваться не собственно количественной информацией об активах банков, а пространственными и корреляционными свойствами образцов этой выборки. По этой причине координаты образцов будем далее обозначать через X_1 , X_2 , X_3 и X_4 , сохранив, однако, нумерацию образцов от 2 до 100 (для удобства сравнения результатов с получаемыми в Deductor).

Реализация плана эксперимента по кластеризации и ранжированию образцов выборки:

Шаг 1. Нормирование исходной выборки к интервалу $[-1; 1]$.

Шаг 2. Анализ выборки с точки зрения ее пространственного разделения. На графиках разброса, построенных по нескольким парам координат, видно, что данные пространственно не делимы (рис. 2).

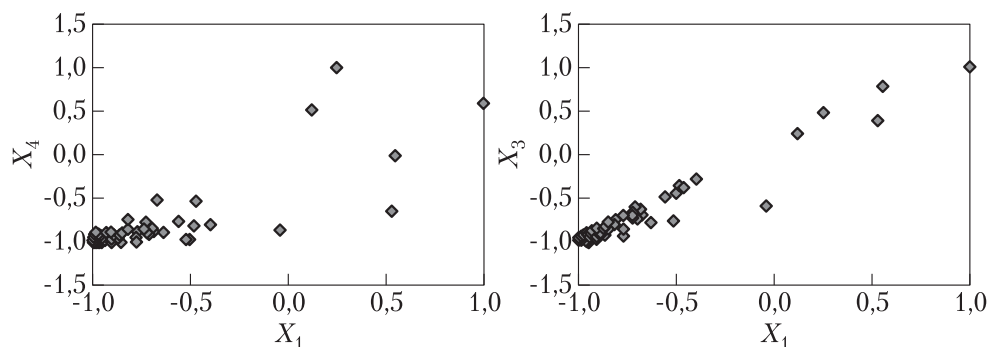


Рис. 2 Проекция многомерных нормированных данных исходной выборки по двум парам координат

Пространственная неразделимость данных делает нецелесообразным применение группирования с помощью самоорганизующейся карты Кохонена, поэтому шаг 3 пропущен.

Шаг 4. Рейтингование образцов учебной выборки по параметру $Rt = \text{mean}$, равному среднему значению нормированных координат образцов учебной выборки, что соответствует единичным весам координат образцов в инварианте линейного преобразования: $\alpha_i = 1, i = 1, \dots, 4$.

Шаг 5. Анализ выборки с точки зрения ее пространственного распределения. Из табл. 1 видно, что учебные образцы в основном сконцентрированы вокруг главной диагонали $(-1, -1, -1, -1)$ – $(1, 1, 1, 1)$ четырехмерного гиперкуба в нормированном входном пространстве, причем координаты 70% образцов отстоят от точки $(-1, -1, -1, -1)$ на небольшое расстояние $r \leq 0,18$.

Особенностью исходной выборки данных является резкая неоднородность пространственного распределения учебных образцов, что также подтверждает

Таблица 1 Перцентили исходной выборки данных

%	X_1	X_2	X_3	X_4	mean	r
0,1	-1,00	-0,99	-0,98	-1,00	-0,99	0,04
0,2	-0,99	-0,98	-0,97	-0,99	-0,98	0,05
0,3	-0,98	-0,98	-0,96	-0,99	-0,98	0,06
0,4	-0,97	-0,97	-0,95	-0,98	-0,97	0,07
0,5	-0,96	-0,96	-0,94	-0,98	-0,96	0,09
0,6	-0,95	-0,95	-0,93	-0,97	-0,95	0,12
0,7	-0,87	-0,94	-0,90	-0,95	-0,92	0,18
0,8	-0,76	-0,89	-0,77	-0,90	-0,81	0,43
0,9	-0,52	-0,80	-0,58	-0,80	-0,69	0,74
1,0	1,00	1,00	1,00	1,00	0,84	3,69

ется графиками (см. рис. 2) и говорит о невозможности применения стандартного подхода к кластеризации. В данном случае необходимо использовать метод многослойной кластеризации.

Шаг 7. Многослойная кластеризация данных. В первом слое данные уже нормированы. Матрица корреляции координат образцов исходной выборки и двух производных от них величин mean и r , где

$$\text{mean} = \frac{X_1 + X_2 + X_3 + X_4}{4};$$

$$r = \sqrt{(X_1 + 1)^2 + (X_2 + 1)^2 + (X_3 + 1)^2 + (X_4 + 1)^2},$$

Таблица 2 Матрица корреляции исходной выборки данных

	X_1	X_2	X_3	X_4	mean
X_1	1				
X_2	0,827	1			
X_3	0,973	0,812	1		
X_4	0,814	0,578	0,867	1	
mean	0,977	0,861	0,987	0,881	1
r	0,983	0,846	0,988	0,879	0,998

Таблица 3 Параметры обучения классической НС Кохонена [1] в Deductor Studio

№	Название	Значение
1	Входные столбцы	$X_i, i = 1, \dots, 4$
2	Выходные столбцы	Прибыль
3	Значимость столбцов	100%
4	Размер обучающего множества	100%
5	Размер тестового множества	0%
6	Считать пример распознанным, если ошибка меньше	0,05 по достижении 500-й эпохи
7	Способ начальной инициализации карты	Из собственных векторов
8	Число эпох, через которое нужно перемешивать строки	20
9	Скорость обучения	В начале обучения — 0,3; в конце обучения — 0,005
10	Радиус обучения	В начале обучения — 4; в конце обучения — 0,1
11	Функция соседства	Ступенчатая
12	Фиксированное количество кластеров (число нейронов в сети)	4

Таблица 4 Параметры обучения клеточной НС Кохонена [7] в Excel

№	Название	Значение
1	Входные столбцы	$X_i, i = 1, \dots, 4$
2	Выходные столбцы	Прибыль
3	Скорость обучения	1
4	Эпох взаимодействия	0,5
5	Число прогонов	30
6	Число нейронов в сети	4
7	Конфигурации обучающих клеточных автоматов	1–8 (см. рис. 1)

Таблица 5 Результаты кластеризации в 1-м слое сети Кохонена

Кластер	Образцы	Размер, %
1	2; 3; 4; 5; 6	5
2	7; 8; 9; 10; 11; 13	6
3	12; 14; 15; 16; 17; 18; 19; 20; 21; 22; 25; 26; 27; 29	14
4	Остальные учебные образцы	75

показывает (табл. 2) хорошую коррелированность выборки данных, что говорит об отсутствии необходимости линейного преобразования координат образцов в первом слое.

Настройки обучения постоянны для каждого слоя сети. Параметры обучения классической НС Кохонена в Deductor представлены в табл. 3.

Клеточная НС Кохонена в Excel обучалась при параметрах, приведенных в табл. 4.

Предварительные эксперименты показали, что обучение НС Кохонена классическим алгоритмом в Deductor Studio и КА в Excel дают одинаковые результаты, деля пространство учебных образцов на четыре кластера (табл. 5).

На рис. 3 представлены результаты кластеризации в 1-м слое сети Кохонена.

Левый столбец содержит номера учебных образцов, верхняя строка — номер реализации эксперимента (см. рис. 3, а) или номер конфигурации обучающего КА (рис. 3, б), цвет заливки определяет один из четырех кластеров, к которому принадлежит соответствующий образец. Здесь и далее образцы в кластерах перечисляются в порядке убывания их рейтингов Rt.

Хотя обучение НС Кохонена обоими алгоритмами и позволило выделить четыре одинаковых кластера, проблема заключалась в том, что 75% наиболее интересных образцов попали в четвертый кластер и по полученным результатам кластеризации о них мало что можно было сказать. Увеличение размерности НС не решало проблемы и приводило лишь к дроблению уже найденных кластеров. Поэтому схема дальнейших экспериментов по

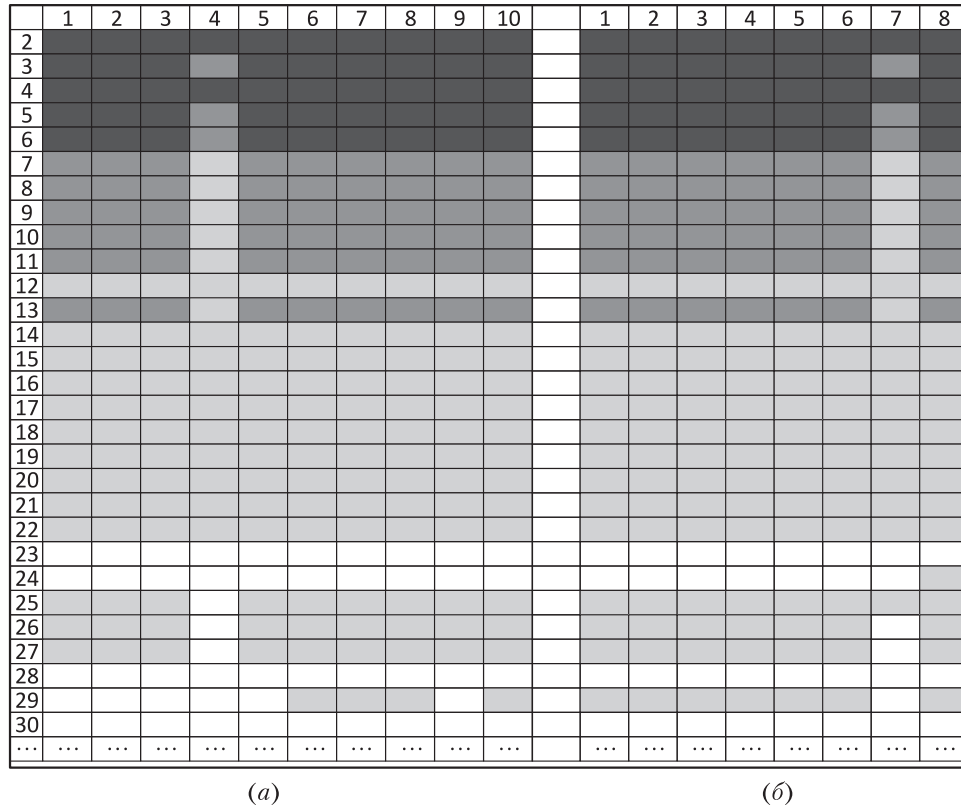


Рис. 3 Результаты кластеризации 99 учебных образцов НС Кохонена 1-го слоя классическим алгоритмом обучения (а) и КА (б)

многослойной кластеризации данных нейронной сетью Кохонена выглядела следующим образом:

1. Нейронная сеть Кохонена обучается в Excel и в Deductor на исходной нормированной выборке данных объемом 99 учебных образцов, что позволяет выделить два небольших стабильных кластера (табл. 6). Оставшиеся

Таблица 6 Отобранные кластеры в 1-м слое сети Кохонена

Кластер	Образцы	Размер, %
1	2; 3; 4; 5; 6	5
2	7; 8; 9; 10; 11; 13	6

- 88 учебных образцов образуют усеченную выборку данных, которая используется для обучения НС Кохонена 2-го слоя.
2. Точно такая же НС Кохонена 2-го слоя обучается в Excel и в Deductor на усеченной нормированной выборке данных объемом 88 учебных образцов, что позволяет отобрать еще два стабильных кластера (табл. 7). Оставшиеся 68 учебных образцов образуют усеченную выборку данных, которая используется для обучения НС Кохонена 3-го слоя.

Таблица 7 Отобранные кластеры во 2-м слое сети Кохонена

Кластер	Образцы	Размер, %
3	12; 14; 15; 16; 17; 18; 19; 20; 21; 22; 25	11
4	23; 24; 26; 27; 28; 29; 30; 32; 37	9

3. Нейронная сеть Кохонена 3-го слоя обучается в Excel и в Deductor Studio на выборке данных объемом 68 учебных образцов. Однако теперь однозначно кластеризовать выборку данных не удалось ни в Excel, ни в Deductor Studio (рис. 4), что объяснялось ее недостаточно высокой коррелированностью и, соответственно, более хаотичным пространственным распределением образцов.

Однако определенная организованность образцов в этой выборке все же была, что давало надежду на выявление закономерностей в данных. Действительно, из табл. 8 и 9 видно, что все коэффициенты в матрице корреляции образцов выборки положительны, а распределение учебных образцов по строкам таблицы перцентилей заметно неравномерно.

В соответствии с методикой невозможность выделения стабильных кластеров при различных начальных параметрах требует дополнительного повышения корреляции. Применение к координатам образцов выборки 3-го слоя линейного преобразования позволило повысить корреляцию и обеспечило надежную кластеризацию учебных образцов. Итогом многослойной кластеризации клеточной НС Кохонена в Excel стало выделение 8 кластеров в разных слоях сети (табл. 10). Образцы перечислены в порядке убывания их рейтинга. При этом, как и ожидалось, средние значения координат образцов по кластерам совпали с абсолютными координатами центров тяжести этих кластеров.

Проведение кластеризации в Deductor по описанной методике также дало надежные результаты, которые совпали с результатами клеточной сети Кохонена в Excel, что говорит об эффективности методики вне зависимости от алгоритма обучения нейронной сети и среды реализации.

Шаг 8. Проверкой правильности нахождения границ кластеров в пространственно неразделимых данных является согласованность результатов рейтингования с кластеризацией. Из табл. 10 видно, что образцы, перечисленные

Таблица 8 Матрица корреляции выборки данных объемом 68 образцов

	X_1	X_2	X_3	X_4	mean
X_2	0,094	1			
X_3	0,666	0,342	1		
X_4	0,203	0,182	0,365	1	
mean	0,717	0,577	0,861	0,621	1
r	0,750	0,565	0,871	0,555	0,989

Таблица 9 Перцентили выборки данных объемом 68 образцов

%	X_1	X_2	X_3	X_4	mean	r
0,1	-0,96	-0,84	-0,52	-0,96	-0,73	0,66
0,2	-0,92	-0,79	-0,44	-0,90	-0,67	0,83
0,3	-0,86	-0,72	-0,37	-0,84	-0,62	0,88
0,4	-0,74	-0,67	-0,28	-0,76	-0,58	0,98
0,5	-0,63	-0,59	-0,07	-0,66	-0,50	1,23
0,6	-0,59	-0,48	0,06	-0,61	-0,35	1,45
0,7	-0,47	-0,39	0,17	-0,56	-0,27	1,62
0,8	-0,34	-0,16	0,31	-0,51	-0,19	1,90
0,9	-0,09	0,04	0,55	-0,11	-0,01	2,29
1,0	1,00	1,00	1,00	1,00	0,38	3,06

Таблица 10 Конечные результаты кластеризации и ранжирования образцов обучающей выборки данных

№	Учебные образы	Размер, %	Слой	Выборка
1	2; 4; 3; 5; 6	5	1	Исходная
2	9; 8; 7; 10; 13; 11	6	1	Исходная
3	18; 15; 22; 19; 16; 21; 14; 12; 20; 25; 17	11	2	Исходная
4	26; 29; 27; 32; 24; 28; 37; 23; 30	9	2	Исходная
5	38; 36; 67; 51; 31; 34; 46; 50; 54	9	3	Преобразованная
6	48; 33; 49; 39; 40; 64; 43; 47; 75; 69; 42; 41; 68; 53; 35; 62; 52; 65; 71	19	3	Преобразованная
7	55; 59; 66; 57; 58; 77; 78; 60; 61; 63; 81; 96; 70; 86; 80; 73	13	3	Преобразованная
8	83; 76; 56; 90; 82; 95; 91; 84; 44; 85; 79; 94; 88; 87; 74; 97; 100; 98; 99; 89; 93; 45; 92; 72	27	3	Преобразованная

в порядке убывания рейтинга, последовательно формируют соседствующие кластеры. Рейтинг убывает в разрезе кластеров, что говорит об однородности их формирования: кластерное отличие в случае непрерывных вещественных величин заключается в увеличении/уменьшении всех средних значений финансовых показателей от кластера к кластеру.

Важно, что рейтингование позволило выделить в составе кластеров как стабильные банки, так и банки, находящиеся в пограничном состоянии, давая возможность спрогнозировать будущее в зависимости от положения банка в конкретном кластере. Так, если банк по рейтингу находится на границе с более сильным кластером, то можно прогнозировать его будущее развитие в сторону усиления финансового положения на рынке. И наоборот: если банк по рейтингу находится на границе с кластером, представленным банками-банкротами, то такой банк также близок к банкротству.

6 Заключение

Описаны результаты применения нового метода кластеризации и ранжирования образцов НС Кохонена с учетом пространственно-корреляционных свойств исходной многомерной выборки данных, который позволил получить однозначные результаты, не зависящие от начальных настроек алгоритма обучения сети. Стабильные результаты кластеризации данных были достигнуты за счет повышения корреляции данных методом линейного преобразования, применение которого эффективно для любого алгоритма обучения сети Кохонена, так как оно выполняется на этапе предобработки данных.

Применение линейного преобразования и многослойной кластеризации позволило получить хорошее согласие результатов кластеризации как клеточной НС Кохонена в Excel, так и классической НС Кохонена в Deductor.

Продемонстрирован метод ранжирования как самостоятельный способ анализа пространственно неразделимых данных, так и используемый для проверки результатов кластеризации.

Практическое применение методики иллюстрируется примером работы с выборкой банков по ряду их финансовых показателей. В результате кластеризации получены обоснованные и стабильные результаты в виде объективной картины состояния анализируемой банковской сферы в разрезе ее участников.

Литература

1. *Кохонен Т.* Самоорганизующиеся карты. — М.: БИНОМ. Лаборатория знаний, 2008. 655 с.
2. *Estivill-Castro V.* Why so many clustering algorithms // ACM SIGKDD Explorations Newsletter, 2002. Vol. 4. No. 1. P. 65–75.
3. *Roy S., Bhattacharyya D. K.* An approach to find embedded clusters using density based techniques // Distributed computing and Internet technology / Ed.

- G. Chakraborty. — Lecture notes in computer science ser. — Springer Verlag, 2005. Vol. 3816. P. 523–535.
4. Kriegel H., Kröger P., Sander J., Zimek A. Density-based clustering // WIREs Data Mining Knowledge Discovery, 2011. Vol. 1. No. 3. P. 231–240.
 5. Witten I. H., Eibe F., Hall M. Data mining: Practical machine learning tools and techniques. — 3rd ed. — Elsevier, 2011. 700 p.
 6. Аникин В. И., Карманова А. А. Обучение искусственной нейронной сети Кохонена клеточным автоматом // Информационные технологии, 2014. № 11. С. 73–80.
 7. Аникин В. И., Карманова А. А. Моделирование и исследование клеточной нейронной сети Кохонена в электронных таблицах // Нейроинформатика-2015: Мат-лы XVII Всеросс. науч.-технич. конф. — М.: НИЯУ МИФИ, 2015. Ч. 2. С. 118–127.
 8. Deductor: Продвинутая аналитика без программирования // Технологии анализа данных: Аналитическая платформа. www.basegroup.ru.
 9. Аникин В. И., Аникина О. В. Визуальное табличное моделирование клеточных автоматов в Microsoft Excel. — Тольятти: ПВГУС, 2013. 324 с.
 10. Ultsch A. U*-matrix: A tool to visualize clusters in high dimensional data. — Department of Computer Science, University of Marburg, 2003. Technical Report No. 36. 12 p. <http://www.uni-marburg.de/fb12/datenbionik/pdf/pubs/2003/ultsch03ustar>.
 11. Андреев А. М., Березкин Д. В., Морозов В. В., Симаков К. В. Метод кластеризации документов текстовых коллекций и синтеза аннотаций кластеров // Электронные библиотеки: перспективные методы и технологии, электронные коллекции — RCDL'2008: Тр. X Всеросс. конф. — Дубна: ОИЯИ, 2008. С. 220–229. http://rcdl2008.jinr.ru/pdf/220_229_paper26.pdf.

Поступила в редакцию 10.07.15

THE EXPERIMENTAL ANALYSIS OF THE METHOD OF CLUSTERING AND RANKING OF MULTIDIMENSIONAL DATA USING THE KOHONEN NEURAL NETWORK

V. I. Anikin¹, O. V. Anikina², and A. A. Karmanova³

¹Volga State Service University, 4 Gagarina Str., Togliatti 445017, Russian Federation

²Togliatti State University, 14 Belorusskaya Str., Togliatti 445667, Russian Federation

³LLC “NetCraker,” 4b Frunze Str., Togliatti 445037, Russian Federation

Abstract: The paper proposes a methodology of clustering and ranking data using the Kohonen neural network based on space-correlation properties of a training sample regardless of the network learning algorithm. The possibility of applying the promising method of linear transformation of training samples coordinates for clustering weakly correlated spatially inseparable data is shown experimentally. The paper demonstrates the usage of ranking to highlight the border instances and define the level of closeness to neighborhood cluster, which makes it possible to solve the problem of finding cluster boundaries in spatially inseparable data. The

necessity of the multilayer clustering is justified in the case of uneven spatial data distribution. The method of clustering and ranking is illustrated by the example of analysis of financial statements empirical data. The technique is applicable to samples of small and medium size.

Keywords: multidimensional clustering; ranking; Kohonen neural network; cellular automaton; linear transformation; correlation matrix

DOI: 10.14357/08696527160104

References

1. Kokhonen, T., eds. 2008. *Samoorganizuyushchiesya karty* [Self-organizing maps]. Moscow: BINOM. Laboratoriya znaniy. 655 p.
2. Estivill-Castro, V. 2002. Why so many clustering algorithms. *ACM SIGKDD Explorations Newsletter* 4(1):65–75.
3. Roy, S., and D. K. Bhattacharyya. 2005. An approach to find embedded clusters using density based techniques. *Distributed computing and Internet technology*. Ed. G. Chakraborty. Lecture notes in computer science ser. Springer Verlag. 3816:523–535.
4. Kriegel, H., H. Kriegel, P. Kröger, J. Sander, and A. Zimek. 2011. Density-based clustering. *WIREs Data Mining Knowledge Discovery* 1(3):231–240.
5. Witten, I. H., F. Eibe, and M. A. Hall. 2011. *Data mining: Practical machine learning tools and techniques*. 3rd ed. Elsevier. 700 p.
6. Anikin, V. I., and A. A. Karmanova. 2014. Obuchenie iskusstvennoy neyronnoy seti Kokhonen na kletochnykh avtomatov [Training of artificial Kohonen neural network by cellular automaton]. *Informatsionnye Tekhnologii* [Information Technology] 11:73–80.
7. Anikin, V. I., and A. A. Karmanova. 2015. Modelirovanie i issledovanie kletochnoy neyronnoy seti Kokhonen v elektronnykh tablitsakh [Simulation and research of Kohonen cellular neural network using spreadsheet]. *XVII Vseross. Nauch.-Tekhnich. Konf. "Neiroinformatika-2015"* [17th All-Russian Conference "Neuroinformatics-2015" Proceedings]. Moscow: MIFI. 2:118–127.
8. Deductor: Analiticheskaya platforma [Deductor Studio Analytical Platform]. Available at: www.basegroup.ru (accessed March 2, 2016).
9. Anikin, V. I., and O. V. Anikina. 2013. *Vizual'noe tablitsnoe modelirovanie kletochnykh avtomatov v Microsoft Excel* [Visual spreadsheet simulation of cellular automata in Microsoft Excel]. Togliatti: PVGUS. 324 p.
10. Ultsch, A. 2003. U*-matrix: A tool to visualize clusters in high dimensional data. Department of Computer Science, University of Marburg. Technical Report No. 36. 12 p. Available at: <http://www.uni-marburg.de/fb12/datenbionik/pdf/pubs/2003/ultsch03ustar> (accessed March 2, 2016).
11. Andreev, A. M., D. V. Berezkin, V. V. Morozov, and K. V. Simakov. 2008. Metod klasterizatsii dokumentov tekstovyykh kolleksiy i sinteza annotatsiy klasterov [Method of the clustering of document collections and clusters annotations synthesis]. *X Vseross. Nauch. Konf. RCDL* [10th All-Russian Conference on RCDL Proceedings]. Dubna: OIYI. 220–229. Available at http://rcdl2008.jinr.ru/pdf/220.229_paper26.pdf (accessed on March 2, 2016).

Received July 10, 2015

Contributors

Anikin Valery I. (b. 1949) — Doctor of Science in technology, professor, Volga State Service University, 4 Gagarina Str., Togliatti 445017, Russian Federation; anikin_vi@mail.ru.

Anikina Oksana V. (b. 1978) — Candidate of Science (PhD) in technology, associate professor, Togliatti State University, 14 Belorusskaya Str., Togliatti 445667, Russian Federation; blue-waterfall@yandex.ru

Karmanova Aleksandra A. (b. 1986) — software engineer, LLC “NetCraker,” 4b Frunze Str., Togliatti 445037, Russian Federation; turaeva.alexandra@gmail.com

ОБ ОДНОЙ РЕАЛИЗАЦИИ АВТОМАТИЗИРОВАННОГО СРЕДСТВА АТТЕСТАЦИИ

А. К. Горшенин¹, А. С. Кучин²

Аннотация: Описаны ключевые аспекты реализации электронной многопользовательской системы тестирования с централизованным хранением данных и использованием ряда мер защиты информации от внешних угроз. Разработанный сервис предназначен для оценки уровня знаний и успешности прохождения различных этапов обучения. Возможность формирования баз тестов в ручном режиме позволяет создавать разнообразные проверочные работы, прежде всего в интересах проведения аттестации в рамках различных курсов в высших учебных заведениях.

Ключевые слова: автоматизация образования; дистанционное обучение; тестирование; архитектура «клиент-сервер»; MySQL; Java

DOI: 10.14357/08696527160105

1 Введение

Активное развитие информационных технологий и их использование в различных областях привело к необходимости применения электронных средств в образовательном процессе. Одним из наиболее ярких примеров привлечения современных технологий служат различные дистанционные курсы, внедренные в учебные планы практически во всех ведущих мировых университетах. При этом образовательный процесс в таких программах зачастую ориентирован на работу с десятками и сотнями тысяч слушателей, что неизбежно приводит к необходимости автоматизации образовательного процесса. Однако подобные системы можно вполне эффективно использовать и для обучения меньшего числа студентов — группы (25–30 чел.), потока (200–250 слушателей) и т. п.

Образовательные методики, используемые для обучения слушателей, могут иметь серьезные отличия на концептуальном уровне. Но для каждого курса независимо от числа обучающихся возникает необходимость в проведении промежуточной (например, контрольные работы) или итоговой (зачеты, экзамены) аттестаций. Автоматизация проверки качества знаний значительно ускоряет и упрощает проведение аттестационных мероприятий, предоставляет профессорско-преподавательскому составу возможность уделять больше времени индивидуальной работе со студентами, а также повысить эффективность обучения.

¹Институт проблем информатики ФИЦ ИУ РАН; Московский технологический университет (МИРЭА), agorshenin@frscs.com

²Московский технологический университет (МИРЭА), kuchin.russia@gmail.com

Такая форма, очевидно, способствует и уменьшению влияния субъективных факторов при оценке знаний.

Все чаще в современных условиях для проведения аттестации используются электронные ресурсы. В частности, российские школьники в достаточной степени привыкли к различным видам тестирования в виде ГИА (государственной итоговой аттестации) и ЕГЭ (Единого государственного экзамена), поэтому им не потребуется значительного времени, чтобы привыкнуть к сдаче тестов в высших учебных заведениях.

В работе [1] представлены основные требования, которые предъявляют современные образовательные стандарты к разработке программного обеспечения для проверки знаний обучающихся (на примере курсов направления «Программирование»). В данной статье рассматривается разработанное в соответствии с этими принципами автоматизированное средство аттестации, созданное на базовой кафедре проблем информатики Института информационных технологий Московского технологического университета.

2 Обзор готовых решений и постановка задачи

Разработка любого программного продукта для дальнейшего внедрения в образовательный процесс должна начинаться с поиска возможных доступных аналогов. Рассмотрим возможности нескольких типовых решений для проведения различных форм контроля, представленных на рынке.

Система тестирования INDIGO (<http://indigotech.ru/>) представляет собой профессиональный инструмент для тестирования и контроля знаний учащихся, проведения опросов, организации олимпиад и конкурсов. Приложение построено согласно принципам архитектуры «клиент-сервер», допускается одновременное подключение множества клиентов (их число ограничено лицензией). Доступны просмотр результатов, статистики, формирование отчетов, создание и редактирование тестов. Данная система ориентирована на работу только с операционными системами (ОС) семейства Windows. Годовая академическая лицензия для тестирования 200 студентов обойдется не менее чем в 75 000 руб. Из особенностей реализации стоит отметить отсутствие прямого доступа к базам данных (БД) сервера даже для администраторов, все возможности реализуются исключительно через клиентские приложения, что может породить определенные виды уязвимостей при подключении к серверу.

Компания SunRav Software (<http://sunrav.ru/>) предлагает целый комплекс программ как для офлайн-, так и для онлайн-тестирования, при этом общая лицензия на год обойдется не менее чем в 50 000 руб. Предлагаются возможности формирования и редактирования тестов, однако количество тем внутри каждого из них ограничено 256. «Локальная» версия реализована только для ОС Windows, тесты и результаты хранятся в зашифрованных файлах, но их передача возложена на пользователя (например, с помощью электронной почты), что также не гарантирует достаточную степень защищенности данных. Редактирование

тестов для онлайн-решения доступно только для ОС Windows, хотя тестирование и администрирование можно осуществлять в браузере, запущенном в любой ОС.

Существуют онлайн-площадки, которые предлагают организацию процесса тестирования, хранения баз, их редактирования исключительно в рамках веб-интерфейса, однако в данной ситуации не могут быть гарантированы доступность сервиса в нужное время (например, из-за технических проблем), безопасность данных (пользователю не предоставляется информации о реализации политик разграничения прав доступа) и т. д.

Таким образом, предлагаемые рыночные решения, несмотря на их гибкость и функциональную наполненность в соответствии с основными требованиями работы [1], не лишены ряда существенных недостатков как с технической (например, отсутствие кроссплатформенности), так и с экономической (высокая стоимость пользования, зависящая от числа одновременно тестируемых студентов) точек зрения. Указанные соображения ведут к необходимости разработки собственного средства аттестации, которое может быть задействовано в учебном процессе, при этом его функциональные возможности должны быть относительно легко расширяемы, в том числе при увеличении числа студентов, проходящих ту или иную форму рубежного контроля.

Коротко сформулируем основные идеи, положенные в основу разрабатываемого средства аттестации обучающихся.

- Сервис должен быть построен в рамках архитектуры «клиент-сервер» с реализацией обмена данными по протоколу TCP/IP для эффективной организации одновременного доступа к системе тестирования большого числа пользователей и использования в рамках существующих сетей (локальных, глобальных) без каких-либо дополнительных требований. Сервер должен использоваться для хранения тестовых заданий, ответов к ним, а также информации об обучающихся для повышения надежности системы.
- Необходимо использовать бесплатные решения в качестве основы для сервера с целью снижения затрат на проектирование архитектуры хранения данных и реализации взаимодействия по сетевым протоколам.
- Средство аттестации должно быть кроссплатформенным как для серверной, так и клиентской частей.
- Для каждого студента, проходящего тестирование, должна быть создана личная учетная запись для персонализированного доступа в систему и просмотра результатов.
- Необходимо обеспечить достаточный уровень защиты информации на всех этапах. Для данных, хранимых на сервере, необходимо обеспечить шифрование личной информации пользователей. При передаче сведений по незащищенным сетям необходимо использование шифрования при обмене данными. При предоставлении информации клиентскому приложению необходимо разработать групповые политики безопасности для доступа к каждому

из ключевых интерфейсов приложения и совершению значимых действий, в том числе для исключения возможности несанкционированной передачи клиенту информации о правильных ответах (проверка заданий должна осуществляться после завершения теста на сервере).

- Система тестирования должна обладать значительной гибкостью в вопросе формирования тестовых заданий для учета различных форм аттестации, особенностей авторских курсов и т. п. Тесты могут формироваться из заданной тематики ручным или случайным выбором фиксированного количества вопросов из общей совокупности (в рамках темы, раздела, курса), при этом ответы на вопросы могут быть основаны на выборе одного или нескольких вариантов из предложенных либо самостоятельного заполнения текстового поля.
- После завершения теста обучающийся должен иметь возможность доступа к сервису для просмотра результатов.

3 Реализация электронного средства аттестации

Рассмотрим практическую реализацию принципов, сформулированных в предыдущем разделе, в рамках разработки автоматизированного средства аттестации.

Для реализации работы серверного модуля была использована система управления БД (СУБД) MySQL [2], в частности из соображений отсутствия ограничений на использование в проектах с открытым исходным кодом. Разработка приложений проводилась в IDE Eclipse на языке программирования Java. Для взаимодействия Java-приложений с СУБД MySQL использован стандарт JDBC (Java Database Connectivity).

3.1 Обеспечение безопасности

Система аттестации подвержена следующим основным видам угроз:

- угроза целостности;
- угроза доступности;
- угроза конфиденциальности.

Критически важной для системы аттестации является угроза раскрытия конфиденциальных сведений о пользователях, содержании тестовых материалов, ответов к ним и т. д.

Для разграничения прав доступа в системе реализованы механизмы авторизации и аутентификации пользователей. Для авторизации пользователей предлагается использование стандартной пары «логин–пароль», при этом логин автоматически формируется из кириллического имени пользователя в соответствии

со стандартами ГОСТ Р 52535.1-2006. Без авторизации в системе пользователю недоступны какие-либо действия.

Так как клиентское приложение хранит адрес, логин и пароль для БД в доступном для дешифровки виде, необходимо использование дополнительных мер защиты для противодействия противоправной авторизации в БД с использованием обычного клиента. Для этого использован механизм двухэтапной авторизации.

Первоначальное подключение к БД дает возможность вызвать только хранящую процедуру авторизации. В случае успешной аутентификации процедура вернет адрес, логин и пароль для рабочей области, после чего клиент получит возможность подключиться и авторизоваться повторно уже в ней.

Во время обмена информацией возможен перехват данных (в том числе паролей) с помощью MitM-атак (от *англ.* Man in the Middle) [3], заключающихся в перехвате данных, передаваемых по сети между клиентом и сервером. Чтобы избежать утечки информации, устанавливается шифрованное соединение между сервером и клиентом с помощью возможностей JDBC и MySQL (обмен информацией происходит по протоколу SSL (secure sockets layer)).

Пользовательские пароли в БД хранятся с использованием хеширования. Этот тип защиты может быть атакован с помощью радужных таблиц [4] — предварительно сгенерированных пар вида «фраза–хеш», выполненных перебором или по словарю. В качестве меры противодействия данному типу уязвимости на вход хеш-функции SHA-256 подается строка из пароля и логина пользователя (стратегия так называемого «добавления соли»).

3.2 Проектирование серверной части

Табличная структура серверной части может быть разделена на несколько основных логических зон:

- зона пользовательских данных;
- зона обеспечения безопасности;
- зона тестов;
- вспомогательная зона.

Зона пользовательских данных отвечает за хранение всех аккаунтов пользователей и связанной с ними личной информации. Выделяются следующие сущности, необходимые для хранения пользовательской информации: «пользователь», «группа пользователей», «имя», «фамилия», «отчество», «логин», «пароль», «электронный адрес», «учебная группа», «университет». Из данного списка можно выделить следующие сущности, которым потребуется создание каталогов, а следовательно, и собственных таблиц:

- «пользователь»;
- «группа пользователей»;
- «университет»;
- «учебная группа».

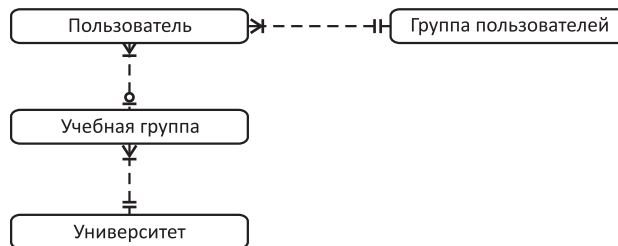


Рис. 1 Модель отношений зоны пользовательских данных

Модель отношений зоны пользовательских данных представлена на рис. 1.

Зона обеспечения безопасности используется для организации доступа с учетом ограничений, установленных для определенных категорий пользователей, а также для общей организации безопасности внутри приложения. Для реализации групповых политик безопасности вводятся понятия «объект безопасности» и «правило безопасности». Каждый объект безопасности логически связан с совершением критических действий в системе, например манипулированием данными посредством вызовов соответствующих API (application programming interface), открытием определенных интерфейсов и т. д. Правило безопасности связывает группу пользователей и объект безопасности, определяя соответствующий уровень доступа. Данная концепция позволяет управлять доступом к ресурсам системы для групп пользователей посредством связывания объекта безопасности с этим ресурсом и проверки его уровня доступа. Таким образом, в зону безопасности входят следующие сущности:

- «группа пользователей»;
- «правило безопасности»;
- «объект безопасности».

Модель отношений зоны безопасности представлена на рис. 2.

Зона тестов предназначена для организации хранения в БД заданий, а также связанной с ними информации.

В системе должны храниться не только результаты выполнения тестов в виде числа правильных ответов к общему числу вопросов, но и каждый вопрос и вари-

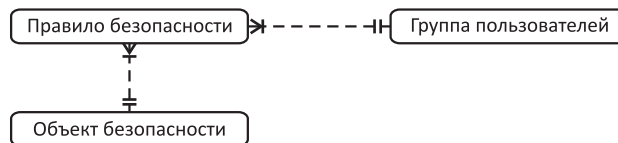


Рис. 2 Модель отношений зоны безопасности

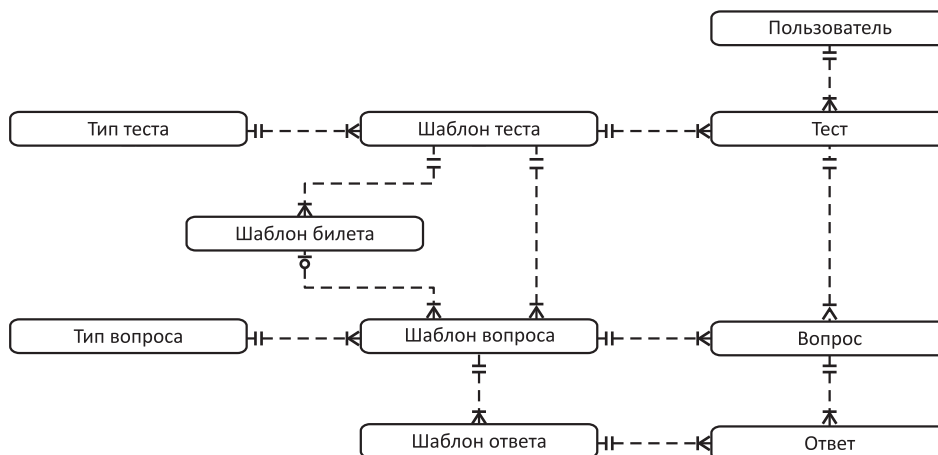


Рис. 3 Модель отношений зоны тестов

ант ответа, выбранные тестируемым. Другими словами, в БД должна храниться информация, позволяющая в любой момент времени после завершения тестирования выгрузить для каждого студента список вопросов и соответствующих ответов. Это позволит составлять аналитические отчеты, вести статистику и объяснять тестируемым их ошибки для повышения успешности образовательного процесса.

Задания подразделяются на шаблоны и тесты. Шаблоны содержат настройки и типы тестов, тексты и типы вопросов и ответов, флаги, определяющие, является ли ответ правильным. Сами же тесты (экземпляры тестов, формируемые для каждого пользователя в момент начала их прохождения) хранят основную информацию о запуске (дата, время, пользователь) и множество ссылок на шаблоны тестов, организованные с помощью внешних ключей. Возможные сущности в зоне тестов и соответствующая модель отношений представлены на рис. 3.

Вспомогательная зона объединяет в себе все таблицы, которые нельзя отнести ни к одной из предыдущих категорий; такие объекты несут дополняющий и служебный характер.

Разделение структуры БД на отдельные зоны позволяет вести независимо проектирование и разработку объектов внутри каждого из них без снижения общей эффективности. Общая итоговая структура БД представлена на рис. 4. Краткие описания полей приведены в таблице на с. 70.

3.3 Примеры интерфейсов сервера и клиентов

В данном разделе рассмотрим примеры интерфейсов для ряда характерных случаев как для серверной, так и для клиентской частей средства аттестации.

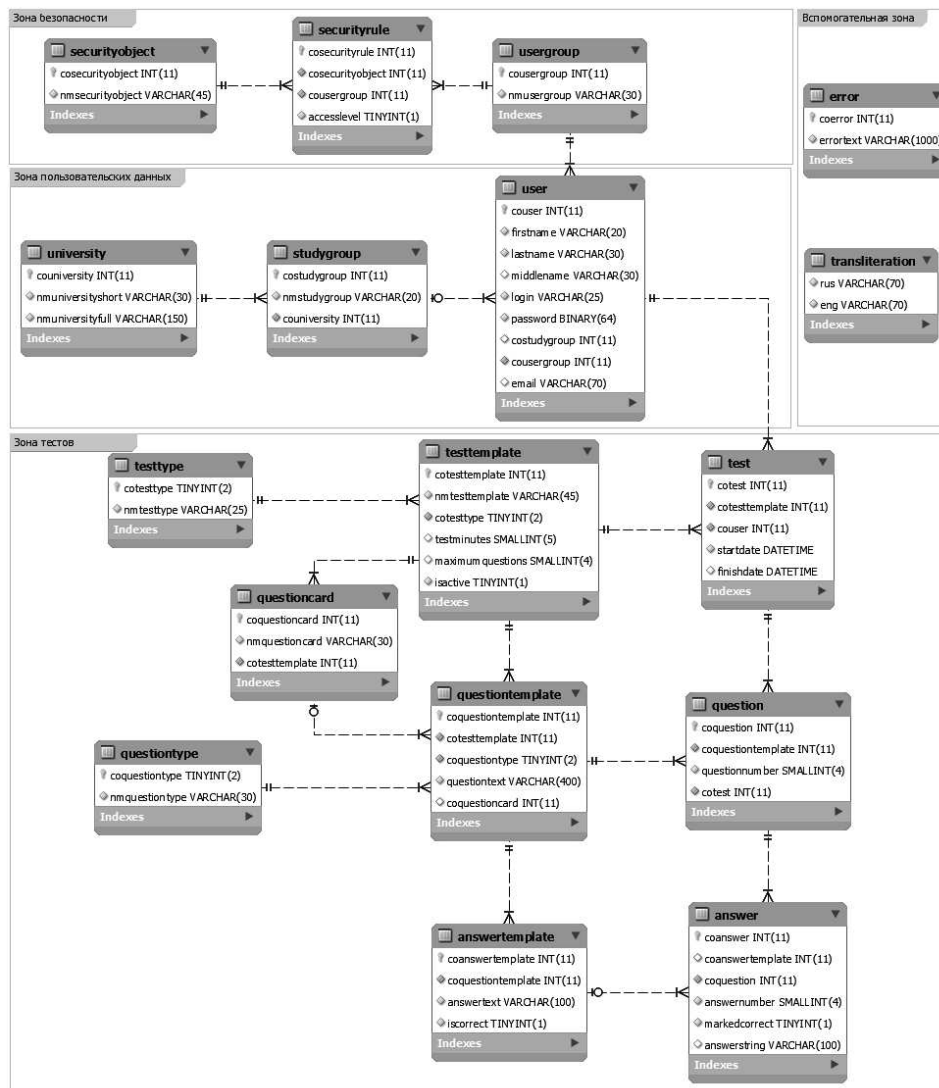


Рис. 4 Общая структура БД

На рис. 5 представлен интерфейс серверной части, позволяющий редактировать базу тестов, при этом ответ должен быть введен в текстовое поле в виде строки. В данной ситуации необходимо предусмотреть все возможные варианты ответов для корректной обработки пользовательского ввода. Для каждого из вопросов вид ответа может быть видоизменен с помощью поля «Тип вопроса».

Описания полей в итоговой структуре БД

Поле	Описание
accesslevel	уровень доступа
answernumber	номер ответа
answerstring	ответ-строка
answertext	текст ответа
coanswer	код ответа
coanswertemplate	код шаблона ответа
coerror	код ошибки
coquestion	код вопроса
coquestioncard	код шаблона билета
coquestiontemplate	код шаблона вопроса
coquestiontype	код типа вопроса
cosecurityobject	код объекта безопасности
cosecurityrule	код правила безопасности
costudygroup	код учебной группы
cotest	код теста
cotesttemplate	код шаблона теста
cotesttype	код типа теста
couniversity	код университета
couser	код пользователя
cousergroup	код группы пользователей
email	электронный адрес
eng	латиница
errortext	описание ошибки
finishdate	дата завершения теста
firstname	имя
isactive	активный
iscorrect	правильный
lastname	фамилия
login	логин
markedcorrect	ответ
maximumquestions	количество вопросов
middlename	отчество
nmquestioncard	имя шаблона билета
nmquestiontype	имя типа вопроса
nmsecurityobject	имя объекта безопасности
nmstudygroup	имя учебной группы
nmtesttemplate	имя шаблона теста
nmtesttype	название типа теста
nmuniversityfull	полное имя вуза
nmuniversityshort	краткое имя вуза
nmusergroup	имя группы пользователей
password	пароль
questionnumber	номер вопроса
questiontext	текст вопроса
rus	кириллица
startdate	начало теста
testminutes	время на тест

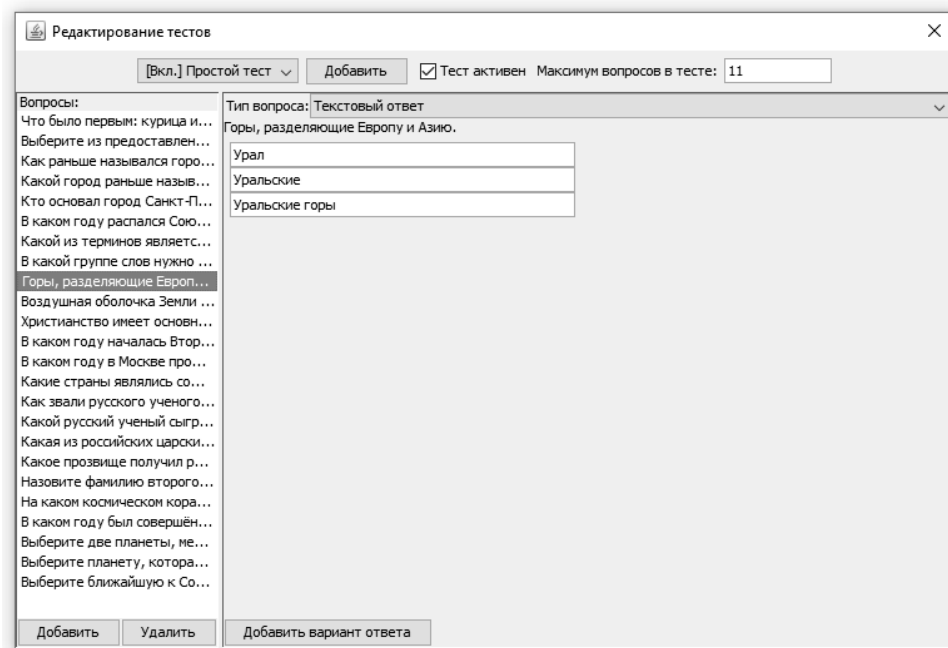


Рис. 5 Пример интерфейса для редактирования тестов (текстовое поле для ответа)

В нижней части окна слева расположены кнопки «Добавить» и «Удалить» для изменения набора заданий.

На рис. 6 представлен интерфейс серверной части, позволяющий редактировать базу тестов, при этом можно отметить сразу несколько вариантов. Всего в базу занесено 24 вопроса, тесты из этой базы формируются с помощью случайной компоновки 11 заданий из общей совокупности (данный параметр указан в поле «Максимум вопросов в тесте» и может быть изменен для каждой конкретной ситуации).

На рис. 7 представлено окно регистрации нового пользователя для обучающегося (приложение-клиент). Предусмотрено заполнение обязательных полей «Фамилия», «Имя» и «E-mail», можно выбрать университет, свою группу из списка предложенных (в примере рассматриваются учебные группы МИРЭА), а затем ввести собственный пароль.

Пример на рис. 8 демонстрирует возможность выбора одного варианта ответа с помощью элемента интерфейса типа «переключатель». Обучающийся может отметить одну из позиций, нажать на кнопку «Ответить» и перейти к следующему вопросу. В нижней части окна расположен индикатор выполнения, показывающий количество пройденных заданий в процентах. В левой части окна

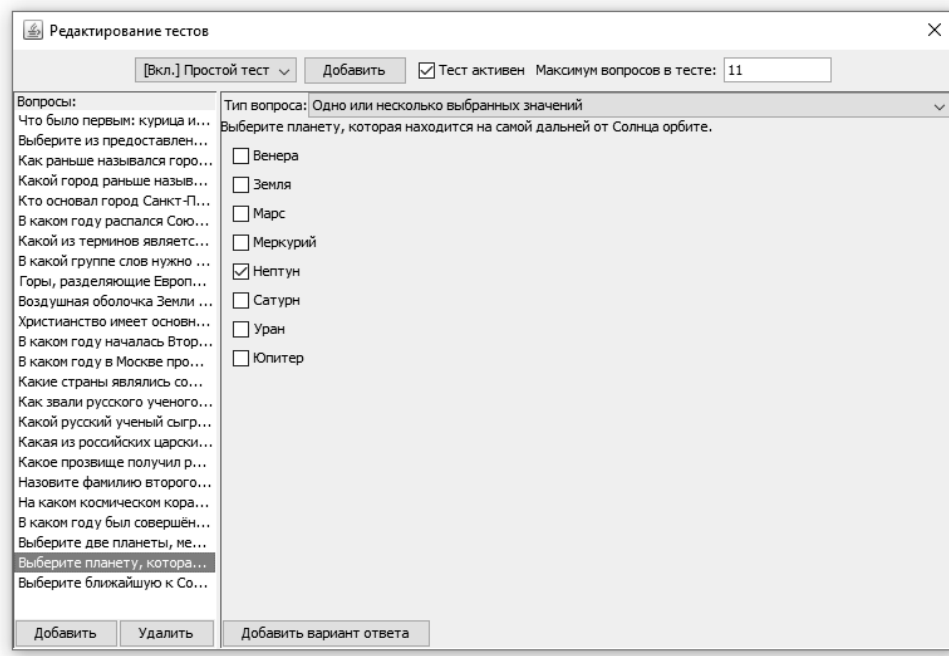


Рис. 6 Пример интерфейса для редактирования тестов (возможность выбора нескольких вариантов ответа)

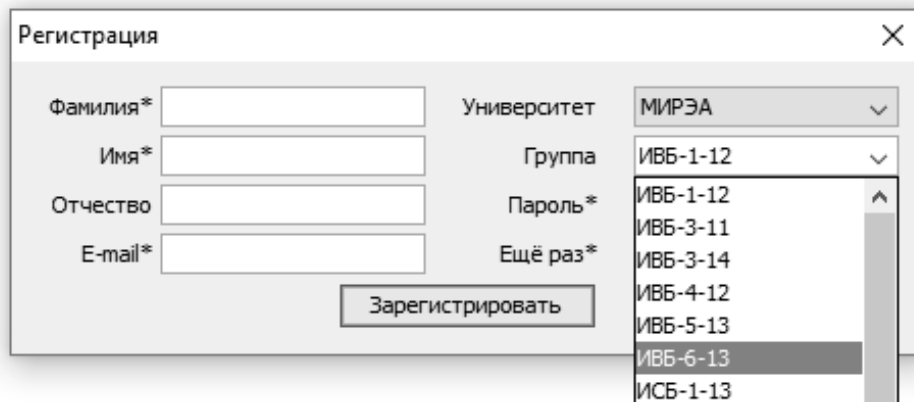


Рис. 7 Пример окна регистрации нового пользователя

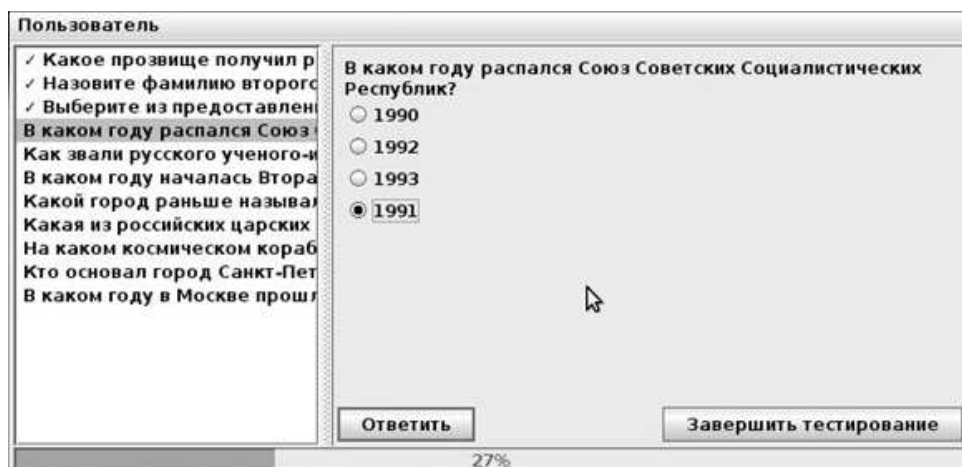


Рис. 8 Пример прохождения теста пользователем

отображаются все вопросы теста, при этом те задания, для которых был получен ответ, отмечены символом «✓». На каждом этапе предусмотрена возможность завершения тестирования. При этом, если в тесте остались невыполненные задания, появится диалоговое окно, которое предупредит пользователя о данной ситуации.

3.4 Тестирование

В систему с помощью API серверной части были введены 24 вопроса по общей тематике и соответствующие им корректные варианты ответов. Тестовое задание составлялось из 11 случайным образом выбранных вопросов. В пробном тестировании приняли участие 96 человек, из них 64 при регистрации указали, что являются студентами МИРЭА. Были сгенерированы и запущены 125 экземпляров тестов, полностью пройдены 113 из них. Собранная статистика показала, что 43% участников использовали персональные компьютеры с установленной ОС Windows 8, 40% — Windows 7, 9% — с ОС на базе Linux и 8% — с MacOS X. Неполадок в работе системы, нарушений политик безопасности и других проблем выявлено не было.

4 Заключение

В статье подробно разобрана реализация автоматизированного средства аттестации для студентов. Безусловно, применение данного сервиса не ограничено сугубо образовательными целями в рамках высшей школы, возможно внедрение процессов аттестации и для сотрудников организаций, компаний и т. д. Для это-

го достаточно обеспечить формирование специальных баз знаний (тестов), что, безусловно, представляет собой отдельную и достаточно трудоемкую задачу.

Предлагаемая система в рамках основного функционала находится на уровне рыночных предложений, при этом является кроссплатформенной, расширяемой и требует меньших затрат на свое внедрение для различных организаций, прежде всего высших учебных заведений. При обеспечении необходимого уровня серверного оборудования данная система может послужить базой для развития дистанционных курсов и со значительным числом слушателей.

В разработанной системе использованы классические средства идентификации, основанные на паре «логин–пароль». Однако дополнительный уровень контроля для исключения прохождения аттестационных мероприятий третьими лицами может обеспечить использование веб-камер и иных технических приспособлений. При этом стоит отметить, что это повысит не только степень контроля, но и увеличит общую стоимость настольного решения, рассмотренного в статье. На преодоление данной проблемы будут направлены дальнейшие исследования.

Литература

1. *Горшенин А. К.* О принципах разработки электронных средств аттестации учащихся по курсам направления «Программирование» // Информатизация инженерного образования: Тр. Междунар. науч.-методич. конф. ИНФОРИНО-2014. — М.: МЭИ, 2014. С. 529–530.
2. *DuBois P.* MySQL. — 5th ed. — Upper Saddle River, NJ, USA: Addison-Wesley Professional, 2013. 1176 p.
3. Encyclopedia of cryptography and security / Eds. H. C. A. van Tilborg, S. Jajodia. — New York, NY, USA: Springer US, 2011. 1416 p.
4. *Oechslin P.* Making a faster cryptanalytic time-memory trade-off // Advances in cryptology — CRYPTO 2003 / Ed. D. Boneh. — Lecture notes in computer science ser. — Springer, 2003. Vol. 2729. P. 617–630.

Поступила в редакцию 11.01.16

ON A REALIZATION OF AN AUTOMATED TESTING SERVICE

A. K. Gorshenin^{1,2} and A. S. Kuchin²

¹Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation

²Moscow Technological University (MIREA), 78 Vernadskogo Ave., Moscow 119454, Russian Federation

Abstract: The paper describes key issues of the development of the automated multiuser testing system with a centralized data storage and a number of information security means to protect data against some external threats.

The service is intended for evaluation of the knowledge and learning success at different educational stages. Different qualifying works can be created manually by lecturers. It is very useful option, primarily for the certification in higher education.

Keywords: education automation; distance learning; testing; client-server model; MySQL; Java

DOI: 10.14357/08696527160105

References

1. Gorshenin, A. K. 2014. O printsipakh razrabotki elektronnykh sredstv attestatsii uchashchikhsya po kursam napravleniya “Programmirovanie” [On principles of development of the electronic certifications for students in programming courses]. *INFORINO-2014 Proceedings*. Moscow. 529–530.
2. DuBois, P. 2013. *MySQL*. 5th ed. Upper Saddle River, NJ: Addison-Wesley Professional. 1176 p.
3. Van Tilborg, H. C. A., and S. Jajodia, eds. 2011. *Encyclopedia of cryptography and security*. New York, NY: Springer US. 1416 p.
4. Oechslin, P. 2003. Making a faster cryptanalytic time-memory trade-off. *Advances in cryptology — CRYPTO 2003*. Ed. D. Boneh. Lecture notes in computer science ser. Springer. 2729:617–630.

Received January 11, 2016

Contributors

Gorshenin Andrey K. (b. 1986) — Candidate of Science (PhD) in physics and mathematics, senior scientist, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilova Str., Moscow 119333, Russian Federation; associate professor, Moscow Technological University (MIREA), 78 Vernadskogo Ave., Moscow 119454, Russian Federation; agorshenin@frcscs.com

Kuchin Alexander S. (b. 1993) — student, Moscow Technological University (MIREA), 78 Vernadskogo Ave., Moscow 119454, Russian Federation; kuchin.russia@gmail.com

ПОСТРОЕНИЕ КЛИЕНТСКИХ .NET-ПРИЛОЖЕНИЙ В РАСПРЕДЕЛЕННЫХ СХЕМОТЕХНИЧЕСКИХ СИСТЕМАХ АВТОМАТИЗИРОВАННОГО ПРОЕКТИРОВАНИЯ*

В. Н. Гридин¹, В. И. Анисимов², А. Д. Ахмад³

Аннотация: Рассматриваются общие принципы и методы построения клиентских .NET-приложений в распределенных схмотехнических САПР. Показывается, что при построении таких приложений необходимо организовать передачу и получение большого числа аргументов с простыми и сложными типами данных. Отмечается, что следует учитывать возможную организацию клиент-серверных взаимодействий в гетерогенных средах, когда клиентские и серверные приложения написаны на разных языках. Приводится методика преобразования данных к единообразной форме, что обеспечивает надежность функционирования клиент-серверных взаимодействий в гетерогенных средах. Предлагается включить в программное обеспечение дополнительный информационный метод веб-службы, назначением которого является расширение WSDL (Web Service Description Language) документа, в котором отсутствует информация о выбранных при построении веб-службы способах преобразования многомерных массивов к унифицированной форме. Рассматривается содержание дополнительного информационного метода веб-службы, необходимого для построения клиентского приложения. Отмечаются особенности построения клиентских Windows-приложений и веб-приложений.

Ключевые слова: системы автоматизированного проектирования; автоматизация схмотехнического проектирования; клиент-серверные взаимодействия; веб-службы; гетерогенные среды; распределенные системы

DOI: 10.14357/08696527160106

1 Введение

Построение систем автоматизированного проектирования (САПР) на основе интернет-технологий требует разработки методов реализации этих технологий с учетом специфики предметной области проектирования и особенностей функционирования веб-служб в конкретных средах. Несмотря на различие физической природы предметных областей, все переменные в компонентных уравнениях

*Работа выполнена при финансовой поддержке РФФИ (проект 14-07-00112 а).

¹Центр информационных технологий в проектировании Российской академии наук (ЦИТП РАН), info@ditc.ras.ru

²Центр информационных технологий в проектировании Российской академии наук (ЦИТП РАН); Санкт-Петербургский государственный электротехнический университет «ЛЭТИ», vianisimov@inbox.ru

³Санкт-Петербургский государственный электротехнический университет «ЛЭТИ», juneidavid@hotmail.com

моделируемой схемы могут быть разделены на две группы — параллельные (продольные) и последовательные (поперечные) переменные, при этом для каждой пары полюсов компонента можно выделить фундаментальные параллельную и последовательную переменные, произведение которых определяет мощность, подводимую к этой паре полюсов [1]. Такая общность описания переменных в различных системах позволяет реализовать универсальные веб-службы, которые при условии унификации передаваемых им аргументов предоставляют возможность обслуживать системы с различной физической природой. Вместе с тем характеристики предметной области существенно влияют на структуру клиентских приложений и требуют учета особенностей каждой предметной области при реализации клиентского приложения. Поэтому актуальной задачей является рассмотрение общих принципов построения клиентских приложений на основе имеющегося унифицированного описания информационных массивов, через которые осуществляется клиент-серверное взаимодействие.

При построении распределенных САПР возможно использование как Windows-приложений, так и веб-приложений [2–4]. Windows-приложения позволяют наиболее полно реализовать оконную графику и организовать диалоговое взаимодействие пользователя с системой. Веб-приложения предоставляют возможность разместить все программное обеспечение клиента в сети Интернет. Достоинством такого приложения является возможность открытого доступа к использованию распределенной САПР через стандартный браузер, недостатком веб-приложения является повышение временных затрат на ввод описания компонентов моделируемой системы. Для клиентского приложения любого типа вызов методов веб-службы выполняется одинаково и возможно использование любых способов реализации клиентских приложений на различных языках. При необходимости клиентские приложения могут быть легко модифицированы применительно к изменяющимся условиям проектирования. Клиентское приложение любого типа может содержать базу данных, информационные ресурсы которой содержат описание применяемых при проектировании компонентов.

Для распределенных систем схемотехнического проектирования характерна работа с большим разнообразием компонентов, в число которых входят резисторы, конденсаторы, индуктивности, управляемые частотно-зависимые источники различных типов (источники тока и напряжения, управляемые напряжением и током — ИТУН, ИНУН, ИТУТ, ИНУТ), трансформаторы, биполярные и униполярные транзисторы, операционные усилители, идеальные трансформаторы и операционные усилители и др. [5–7]. При организации клиент-серверных взаимодействий для таких систем необходимо организовать эффективную передачу и получение большого числа аргументов с простыми и сложными типами данных. При этом необходимо учитывать возможную работу веб-службы и клиентских приложений в гетерогенных средах, когда клиентские и серверные приложения написаны на разных языках, например на C# и Java [8, 9].

В настоящей статье рассматривается предлагаемая авторами новая методика построения клиентских приложений в среде .NET, дающая возможность, в от-

личие от имеющихся методик, обеспечить надежность функционирования клиент-серверных взаимодействий в гетерогенных средах, примененная в реальных задачах построения клиентских .NET-приложений распределенных схмотехнических САПР.

2 Информационные методы веб-служб и унификация данных

Основные принципы разработки распределенных САПР основаны на предположении, что построение веб-службы и клиентских приложений осуществляется независимо друг от друга разными коллективами программистов. При этом каждый коллектив разработчиков, как правило, не имеет полной информации обо всех деталях программного обеспечения приложения, разрабатываемого другим коллективом.

Хотя при построении веб-службы формируется WSDL-документ с ее описанием и доступ к этому документу возможен до построения клиентского приложения, задача независимого построения клиентского приложения остается весьма сложной. Это объясняется тем, что WSDL-документ, с одной стороны, обеспечивает требуемую информацию для возможности клиент-серверного взаимодействия, но, с другой стороны, не содержит достаточной информации для независимого построения клиентского приложения. Так, в частности, в WSDL-документе отсутствует информация о выбранных при построении веб-службы способах преобразования многомерных массивов к единообразной форме, что требуется для обеспечения надежного функционирования клиентских и серверных приложений в гетерогенных средах, когда приходится отказываться от передачи информации между такими приложениями через многомерные массивы. Поскольку для описания всех компонентов моделируемой схемы используются двумерные массивы, то необходима предварительная упаковка двумерных массивов в одномерные с последующей передачей упакованных массивов в качестве аргументов методам веб-службы. Процедуру упаковки многомерных массивов следует выполнять также и перед передачей результатов работы веб-службы клиентскому приложению. Поэтому каждая веб-служба распределенной системы автоматизации схмотехнического проектирования должна обеспечивать разработчиков клиентских приложений дополнительной информацией, необходимой для независимого от разработчика веб-службы построения этих приложений. Такая задача может быть решена путем построения в веб-службе дополнительного информационного метода, например `Info()`, не имеющего аргументов и возвращающего простую строковую переменную. Содержание этой переменной должно обязательно включать в себя:

- алгоритм преобразования многомерных массивов описания компонентов схемы к унифицированной форме `In_comp`, `Z_comp`, обеспечивающей гарантированное функционирование распределенной системы в гетерогенных средах;
- описание возвращаемых результатов работы методов веб-службы в унифицированной форме;

- описание преобразования унифицированной формы возвращаемых результатов методов веб-службы к виду, требуемому для отображения результатов работы веб-службы в клиентских приложениях.

Если при построении веб-службы ввести содержание этой переменной в аргумент **Description** атрибута **WebServices** для метода **Info**, то содержание этой переменной отобразится в браузере при вводе URL (Uniform Resource Locator) адреса веб-службы, а также в WSDL-документе. При этом нет необходимости вызова самого метода **Info**, что дает возможность выполнить разработку клиентского приложения независимо от разработчика веб-службы.

Таким образом, если для описания компонентов какого-либо вида открыт двумерный массив некоторого типа **T** с именем **arr_comp**, число строк которого (число компонентов данного вида) имеет значение **ncomp**, а число столбцов массива имеет значение **ncol**, т. е.

$$Tarr_comp = \text{new } T[ncomp + 1, ncol],$$

где строка с нулевым индексом не используется, то такой массив после заполнения его в клиентском приложении должен быть упакован по столбцам в одномерный массив с именем **Arr_comp**, т. е.

$$T Arr_comp = \text{new } T[(ncol - 1) * ncomp + 1],$$

при этом элемент с нулевым индексом не используется.

Аналогичным образом в методах веб-служб упаковываются многомерные массивы с результатами работы этих методов, при этом алгоритм преобразования многомерных массивов описывается в информационном методе веб-службы.

3 Реализация клиентских Windows-приложений .NET

Клиентские Windows-приложения позволяют наиболее эффективно реализовать системные средства оконной графики и организовать диалоговое взаимодействие пользователя с системой. Вне зависимости от вида конкретного клиентского приложения его структура должна быть построена в соответствии с блок-схемой, приведенной на рис. 1.

В первом блоке осуществляется ввод описания компонентов моделируемой схемы. Для каждого компонента, к которым относятся двухполюсники типа *R*, *C*, *L*, частотно-зависимые управляемые источники, трансформаторы, биполярные транзисторы, униполярные транзисторы, операционные усилители, задающие источники тока и напряжения, здесь должны быть сформированы массив включения компонента **in_comp** и массив значений его параметров **z_comp**, и оба массива должны быть двумерными. Для реализации ввода описания с клавиатуры целесообразно создание совокупности диалоговых окон, хотя возможно описание компонентов непосредственно в текстовом файле в распространенном Spice-формате с последующим преобразованием его в массивы описания компонентов. При построении первого блока необходимо также предусмотреть возможность записи

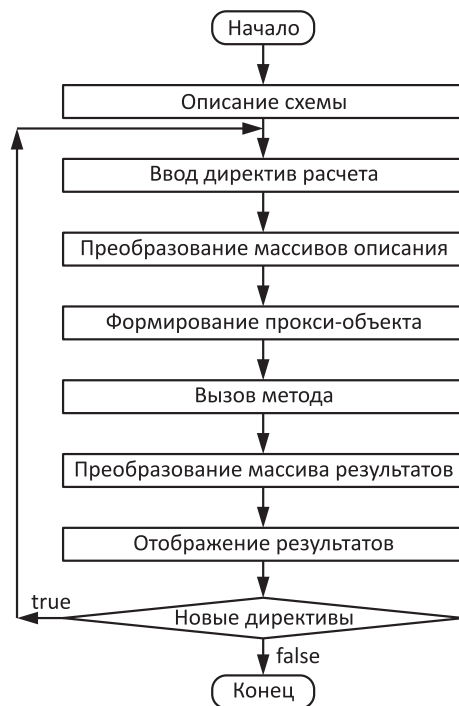


Рис. 1 Структура клиентского приложения

ционирование распределенной системы в гетерогенных средах. Следующий блок содержит оператор формирования прокси-объекта для взаимодействия веб-службы с клиентскими приложениями (например, `MFD ws = new MFD()`, где `MFD` — имя веб-службы, заданное значением аргумента `Name` при создании веб-службы). При затруднениях с записью этого оператора, а также оператора для вызова основного метода `Calc` веб-службы их можно извлечь из `WSDL`-документа веб-службы, где для информационного метода `Info` отображается аргумент `Description` атрибута `WebServices` (рис. 2).

В `WSDL`-документе для информационного метода `Info` также описывается алгоритм распаковки возвращаемого массива `Out`. Результаты распаковки этого массива отображаются в клиентском приложении.

Для расширения возможностей `Windows`-приложения можно включить в него дополнительные функции:

- ввод описания компонентов из базы данных;
- вывод описания схемы в базу данных;
- вывод результатов расчета в требуемой табличной форме;
- ввод описания схемы при помощи графического редактора.

введенного с клавиатуры описания схемы в файл, а также реализовать повторный ввод описания схемы непосредственно из файла. Кроме того, следует расширить этот блок программной реализацией системы редактирования отдельных компонентов для обеспечения возможности реализации многовариантных расчетов моделируемой схемы.

Блок ввода директив расчета реализуется в соответствии с общей целевой функцией веб-службы. Так, при построении веб-службы для моделирования схем в частотной области к директивам расчета относится вид частотной характеристики и задание входных и выходных узлов.

В следующем блоке реализуется алгоритм преобразования многомерных массивов описания компонентов схемы `in_comp` и `z_comp` к унифицированной форме одномерных массивов `In_comp` и `Z_comp`, что обеспечивает гарантированное функци-

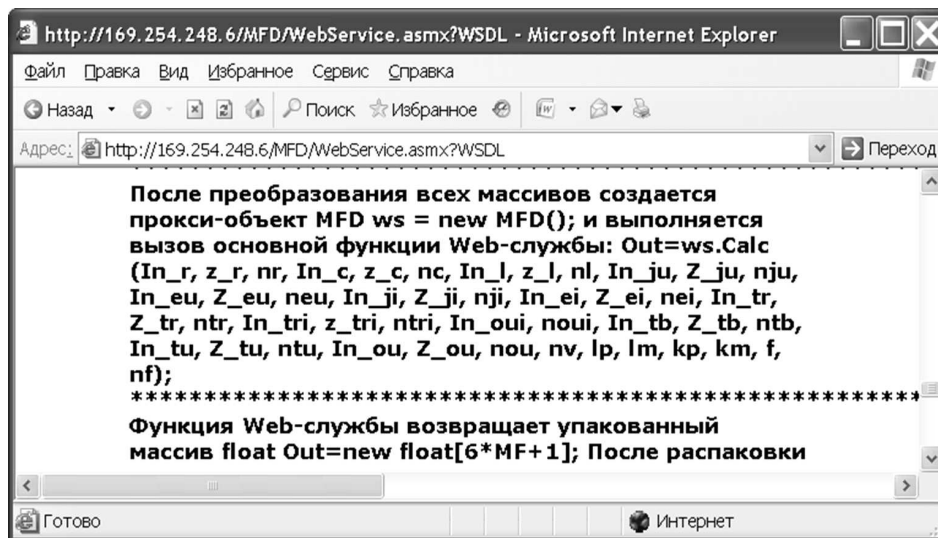


Рис. 2 Фрагмент WSDL-документа веб-службы MFD

Кроме того, может оказаться целесообразным создание встроенного в клиентское приложение веб-браузера для обращения к удаленным базам данных.

4 Реализация клиентских веб-приложений .NET

Клиентские веб-приложения предоставляют возможность расположить как клиентскую, так и серверную часть распределенной системы в сети Интернет. Достоинством такой организации клиент-серверного взаимодействия является открытый доступ к использованию системы через существующее клиентское веб-приложение без необходимости построения клиентского приложения на компьютере пользователя. Известным недостатком такого подхода является увеличение времени, требуемого для ввода описания данных проектируемой системы, что связано с необходимостью ожидания реакции системы при вводе данных.

Возможны различные способы размещения в сети клиентского веб-приложения. Первый способ предполагает размещение клиентского веб-приложения в том же узле сети, где расположена веб-служба. Второй способ заключается в расположении клиентского веб-приложения на компьютере пользователя. И наконец, третий способ предусматривает выделение для клиентского веб-приложения отдельного узла, где будут расположены все клиентские веб-приложения распределенной системы.

Вне зависимости от выбранного способа размещения в сети клиентского веб-приложения его вызов осуществляется непосредственно через браузер пользовательского компьютера путем ввода в него соответствующего URL-адреса.

Для практической реализации построения клиентского веб-приложения в среде .NET проще всего воспользоваться шаблоном ASP.NET Empty Web Site, что приводит к созданию приложения с единственным файлом web.config весьма простой структуры. Выбором пункта меню WebSite | Add New Item с последующей активизацией шаблона Web Form можно построить требуемые страницы ASP.NET и создать необходимые файлы с расширением .aspx.

Все создаваемые страницы ASP.NET должны выполнять следующие задачи:

- выбор некоторого значения и переход к другой странице. Страницы этого типа позволяют организовать функционирование главного меню клиентского веб-приложения;
- реализация ввода некоторой совокупности переменных с переходом к начальной странице. Страницы этого типа позволяют осуществить ввод из текстовых окон в программу некоторой совокупности значений переменных с переходом после завершения ввода к начальной странице Default.aspx;
- одноразовый ввод некоторой совокупности переменных с переходом к следующей странице. Страницы этого типа позволяют осуществить одноразовый ввод из текстовых окон в программу некоторой совокупности значений переменных с переходом после завершения ввода к следующей странице, имя которой определяется в соответствии со значениями этих переменных;
- циклический ввод некоторой совокупности переменных с переходом к следующей странице. Страницы этого типа позволяют осуществить циклический ввод из диалоговых текстовых окон в программу некоторой совокупности значений переменных с переходом после завершения цикла к следующей странице, имя которой определяется согласно значениям переменных, определяющих размерность схемы;
- вызов метода веб-службы и вывода результатов его работы. Страницы этого типа, как правило, не имеют элементов управления для ввода данных и содержат только скриплеты, в которых приводится код для вызова веб-сервиса и отображения результатов его работы. Здесь также размещается код для упаковки передаваемых веб-службе данных в одномерные массивы.

Поскольку клиентские веб-приложения для распределенных систем схемотехнического проектирования могут содержать значительное число ASP.NET-страниц, то использование шаблонов таких страниц повышает эффективность разработки клиентских веб-приложений, так как становится возможным многократное использование типовых структур с незначительными изменениями их программного кода.

5 Заключение

Характеристики предметной области существенно влияют на структуру клиентских приложений и требуют учета особенностей каждой предметной области при реализации программного обеспечения клиентского приложения. Автора-

ми предложена методика построения дополнительного информационного метода, назначением которого является расширение WSDL-документа, отличающаяся от известных ранее методик наличием информации о выбранных при построении веб-службы способах преобразования многомерных массивов к унифицированной форме. Использование предлагаемой методики на основе общих принципов построения клиентских приложений для распределенных систем автоматизации схмотехнического проектирования позволяет существенно повысить эффективность функционирования таких систем. Достоинством рассмотренного подхода является возможность организации клиент-серверных взаимодействий в случаях, когда клиентские и серверные приложения написаны на разных языках. Приведенная методика преобразования данных к единообразной форме обеспечивает надежность функционирования клиент-серверных взаимодействий в гетерогенных средах.

Литература

1. Ильинский Н. Ф., Цаценкин В. К. Приложение теории графов к задачам электромеханики. — М.: Энергия, 1968. 200 с.
2. Гридин В. Н., Анисимов В. И. Методы построения систем автоматизированного проектирования на основе интернет-технологий и компактной обработки разреженных матриц // Информационные технологии в проектировании и производстве, 2009. № 1. С. 3–7.
3. Гридин В. Н., Дмитриевич Г. Д., Анисимов Д. А. Построение систем автоматизированного проектирования на основе Web-сервисов // Автоматизация в промышленности, 2011. № 1. С. 9–12.
4. Гридин В. Н., Дмитриевич Г. Д., Анисимов Д. А. Построение систем автоматизированного проектирования на основе Web-технологий // Информационные технологии, 2011. № 5. С. 23–27.
5. Гридин В. Н., Дмитриевич Г. Д., Анисимов Д. А. Построение веб-сервисов систем автоматизации схмотехнического проектирования // Информационные технологии и вычислительные системы, 2012. № 4. С. 79–84.
6. Алмаасали С. А., Анисимов В. И. Построение распределенных систем автоматизированного проектирования на основе методов диакоптики // Известия СПбГЭТУ, 2014. № 1. С. 15–19.
7. Гридин В. Н., Анисимов В. И., Шабани М. А. Методы построения высокопроизводительных распределенных систем автоматизации схмотехнического проектирования // Информационные технологии, 2014. № 8. С. 59–63.
8. Морган С., Райан Б., Хорн Ш., Бломсма М. Разработка распределенных приложений на платформе Microsoft.Net / Пер. с англ. — СПб.: Питер, 2008. 608 с. (Morgan S., Ryan B., Horn Sh., Blomsma M. Microsoft .NET FRAMEWORK 2.0 distributed application development. — Microsoft Press, 2007. 608 p.)
9. Дей Н., Мандел Л., Райман А. Eclipse: Платформа Web-инструментов / Пер. с англ. — М.: КУДИЦ-ПРЕСС, 2008. 688 с. (Dai N., Mandel L., Ryman A. Eclipse Web tools platform. — 1st ed. — New York, NY, USA: Addison-Wesley, 2008. 688 p.)

Поступила в редакцию 07.09.15

BUILDING .NET-CLIENT APPLICATIONS IN DISTRIBUTED CIRCUIT COMPUTER-AIDED DESIGN

V. N. Gridin¹, V. I. Anisimov^{1,2}, and A. D. Ahmad²

¹Design Information Technologies Center, Russian Academy of Sciences, 7a Marshal Biryuzov Str., Odintsovo, Moscow Region 143000, Russian Federation

²Saint Petersburg Electrotechnical University “LETI,” 5 Professora Popova Str., St. Petersburg 197376, Russian Federation

Abstract: The paper deals with the general principles and methods of building client .NET-applications in distributed circuit computer-aided design. It is shown that the construction of such applications requires the transfer and reception of a large number of arguments with simple and complex data types. It is noted that the organization should consider the client–server interactions in heterogeneous environments, when client and server applications are written in different languages. The paper describes the technique of converting data to a uniform manner, which ensures reliable operation of client–server interactions in heterogeneous environments. It is proposed to include the additional information software Web service method, the purpose of which is to expand the WSDL-document in which there is no information about selected when building a Web service methods converting multidimensional arrays to a unified form. The paper describes the content of the additional information of the Web service method necessary to build the client application. The paper describes the particular features of client build Windows-based applications and Web applications.

Keywords: computer-aided design; circuit design automation; client-server interaction; Web services; heterogeneous environments; distributed systems

DOI: 10.14357/08696527160106

Acknowledgments

The work was financially supported by the Russian Foundation for Basic Research (project 14-07-00112 a).

References

1. Il'inskiy, N. F., and V. K. Tsatsenkin. 1968. *Prilozhenie teorii grafov k zadacham elektromekhaniki* [The application of graph theory to problems in electrical engineering]. Moscow: Energiya. 200 p.
2. Gridin, V. N., and V. I. Anisimov. 2009. Metody postroeniya sistem avtomatizirovannogo proektirovaniya na osnove internet-tekhnologiy i kompaktnoy obrabotki razrezhenykh matrits [Construction methods of CAD systems based on internet technology and compact processing of sparse matrices]. *Informatsionnye Tekhnologii v Proektirovanii i Proizvodstve* [Information Technology in Design and Manufacture] 1:3–7.

3. Gridin, V. N., G. D. Dmitrevich, and D. A. Anisimov. 2011. Postroenie sistem avtomatizirovannogo proektirovaniya na osnove Web-servisov [Construction of computer-aided design based on Web-services]. *Avtomatizatsiya v Promyshlennosti* [Automation in Industry] 1:9–12.
4. Gridin, V. N., G. D. Dmitrevich, and D. A. Anisimov. 2011. Postroenie sistem avtomatizirovannogo proektirovaniya na osnove Web-tekhnologiy [Construction of computer-aided design Web-based technologies]. *Informatsionnye Tekhnologii* [Information Technology] 5:23–27.
5. Gridin, V. N., G. D. Dmitrevich, and D. A. Anisimov. 2012. Postroenie veb-servisov sistem avtomatizatsii skhemotekhnicheskogo proektirovaniya [Building of Web services for automation circuit design]. *Informatsionnye Tekhnologii i Vychislitel'nye Sistemy* [Information Technology and Computer Systems] 4:79–84.
6. Almaasali, S. A., and V. I. Anisimov. 2014. Postroenie raspredelennykh sistem avtomatizirovannogo proektirovaniya na osnove metodov diakoptiki [Distributed CAD systems based on diakoptiks methods]. *Izvestiya SPbGETU* [News of ETU] 1:15–19.
7. Gridin, V. N., V. I. Anisimov, and M. A. Shabani. 2014. Metody postroeniya vysokoproizvoditel'nykh raspredelennykh sistem avtomatizatsii skhemotekhnicheskogo proektirovaniya [Methods of construction of high-performance distributed systems of automation of circuit design]. *Informatsionnye Tekhnologii* [Information Technology] 8:59–63.
8. Morgan, S., B. Ryan, Sh. Horn, and M. Blomsma. 2007. *Microsoft .NET FRAMEWORK 2.0 distributed application development*. Microsoft Press. 608 p.
9. Dai, N., L. Mandel, and A. Ryman. 2008. *Eclipse Web tools platform*. 1st ed. New York, NY: Addison-Wesley. 688 p.

Received September 7, 2015

Contributor

Gridin Vladimir N. (b. 1944) — Doctor of Science in technology, professor, scientific leader, Design Information Technologies Center, Russian Academy of Sciences, 7a Marshal Biryuzov Str., Odintsovo, Moscow Region 143000, Russian Federation; info@ditc.ras.ru

Anisimov Vladimir I. (b. 1926) — Doctor of Science in technology, professor, Saint Petersburg Electrotechnical University “LETI,” 5 Professora Popova Str., St. Petersburg 197376, Russian Federation; principal scientist, Design Information Technologies Center, Russian Academy of Sciences, 7a Marshal Biryuzov Str., Odintsovo, Moscow Region 143000, Russian Federation; vianisimov@inbox.ru

Ahmad Altaib Davod (b. 1982) — PhD student, Saint Petersburg Electrotechnical University “LETI,” 5 Professora Popova Str., St. Petersburg 197376, Russian Federation; juneidavid@hotmail.com

ПРОБЛЕМЫ ОБЕСПЕЧЕНИЯ СЕМАНТИЧЕСКОЙ ГЕОИНТЕРОПЕРАБЕЛЬНОСТИ И СОГЛАСОВАНИЯ ПОНИМАНИЯ СЕМАНТИКИ ГЕОДАНЫХ*

С. К. Дулин¹, Н. Г. Дулина², Д. А. Никишин³

Аннотация: Основной задачей работы ставилась выработка целостного и по возможности всестороннего взгляда на проблему интероперабельности (ИО, *англ.* Interoperability) взаимодействующих информационных систем (ИС) вообще и семантического аспекта этой проблемы в частности. Рассматриваются основные причины возникновения проблемы межсистемной ИО, ее различные определения и подходы к ее структуризации, на основании чего сделана попытка структурировать проблему ИО, выявляются неоднозначные моменты и предлагаются возможные варианты их устранения. Анализируется состояние дел в области стандартизации метаданных, на основании чего сделан вывод о необходимости стандартизации форм и способов выполнения поисковых запросов и межсистемных функциональных вызовов. Рассматриваются основные моменты проблемы понимания смысла данных, существующие подходы к ее структурированию и решению.

Ключевые слова: геоданные; семантическая геоинтероперабельность; онтологии; Semantic Web

DOI: 10.14357/08696527160107

1 Введение

В настоящее время, как никогда ранее, актуальны и остры проблемы информационного взаимодействия бесчисленного множества ИС — как на уровне различных государственных и негосударственных организаций, ведомств, компаний, так и на международном уровне — внешнеполитических союзов, блоков стран и т. п. Так, в работе [1] отмечается, что из-за растущей взаимосвязанности информационных процессов и глобализации систем поддержки принятия решений увеличивается цена каждой ошибки при принятии решений, а реакция на принимаемые решения приходит быстрее, сокрушительнее. И одним из краеугольных камней проблемы информационного взаимодействия, наряду

*Работа выполнена при поддержке РФФИ (проект 14-07-00040).

¹Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, s.dulin@ccas.ru.

²Вычислительный центр им. А. А. Дородницына Федерального исследовательского центра «Информатика и управление» Российской академии наук, ngdulina@mail.ru

³Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, dmnikishin@mail.ru

с проблемами обеспечения конфиденциальности, коммерческой, государственной тайны, различий законодательных и юридических подходов и т. п., является проблема информационной ИО.

Целью данной статьи является анализ исключительно технического аспекта проблемы ИО, организационные и прочие аспекты этой проблемы не затрагиваются. В работе рассматриваются основные причины возникновения проблемы ИО, ее различные определения и структура. Как будет показано ниже, проблема ИО на техническом уровне в общем случае сводится к обеспечению семантической ИО (*англ.* Semantic Interoperability, SI). В связи с этим также затрагивается проблема понимания вообще и понимания смысла данных в частности, рассматриваются ее определения и подходы к структуризации этой проблемы. Анализ прикладных вопросов посвящен последнему разделу, где рассматривается примерный подход к построению и функционированию интероперабельной гиперсистемы (инфраструктуры).

2 Определение понятия «интероперабельность»

Основные причины возникновения проблемы ИО заключаются в том, что ИС создавались (а зачастую создаются и по сей день) стихийно для решения конкретных, насущных задач в соответствии с разными требованиями, различными группами разработчиков и по различным технологиям, в разное время и являются, по существу, черными ящиками [1]. Сохранение существующего положения диктуется не только технологическими проблемами, не только индивидуальными, корпоративными, государственными интересами (сделанными колоссальными вложениями, необходимостью сохранения авторства, коммерческой и государственной тайны, монопольного влияния и т. п.), но и отсутствием достаточно конкретной, детальной, универсальной и конвенциональной модели обеспечения информационного обмена, которая потенциально обеспечила бы решение всех вероятных задач обеспечения взаимодействия, которые могут возникнуть перед создаваемой ИС по крайней мере на основном этапе ее жизненного цикла.

Проведем анализ определения понятия «интероперабельность», опираясь на отечественный ГОСТ Р 55062-2012 «Системы промышленной автоматизации и их интеграция. Интероперабельность. Основные положения» [2]: «Интероперабельность, наряду со свойствами «переносимость» и «масштабируемость», представляет важнейшее свойство открытых систем»; «Интероперабельность — способность двух или более информационных систем или компонентов к обмену информацией и к использованию информации, полученной в результате обмена».

Согласно определению стандарта ISO/IEC 24765 «Systems and Software Engineering Vocabulary», «интероперабельность — способность двух и более систем или элементов обмениваться информацией и использовать эту информацию». Однако некоторые авторы, занимающиеся проблемой ИО, приходят к более развернутому определению ИО как «обеспечения согласованного взаимо-

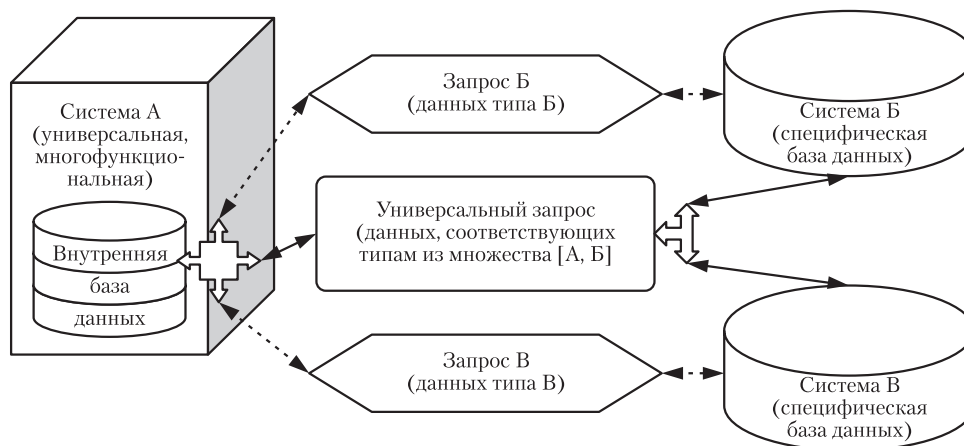


Рис. 1 Иллюстрация асимметричной ИО при взаимодействии ИС с различной функциональностью

действия участников того или иного информационного процесса, для чего должно достигаться согласованное понимание целей и методов взаимодействия» [1]. При этом подразумевается, что в современных условиях должна обеспечиваться возможность полной автоматизации процесса интерпретации участвующими во взаимодействии ИС смысла информации, принимаемой извне. Здесь можно отметить три момента:

- (1) две взаимодействующие системы не обязательно должны быть симметричны по степени обеспечения ИО. Так, на рис. 1 система А может быть в полной мере интероперабельна по отношению к системе В, а вот В может и не быть полностью интероперабельна с А: к примеру, система В может оказаться специфическим информационным ресурсом (например, базой данных), умеющим работать только со своими «родными» данными (типа В или со сторонними данными, но тоже соответствующими идеологии В и никакой другой), а система А является универсальной системой, для которой данные типа В являются только частным случаем;
- (2) на сегодняшний день отсутствуют общие методы количественной оценки ИО любых систем, использующих произвольный способ взаимодействия. Этому вопросу посвящен ряд публикаций (см., например, [3]). Основным препятствием для такой оценки является отсутствие количественных характеристик степени обеспечения ИО, поэтому используется субъективный, экспертный подход;
- (3) что касается «обмена» и «использования» информации, то здесь напрашивается следующий вопрос: если информацию не предполагается использовать, то какой тогда смысл ею вообще обмениваться?

Таким образом, приведенные в стандартах определения не раскрывают сути ИО и сводятся к ее пониманию лишь как способности к взаимодействию. В этом плане более ценным выглядит расширенное определение, где обозначены проблемы согласованного понимания и организации согласованного взаимодействия участников взаимодействия.

3 Структура интероперабельности: уровни интероперабельности

При структуризации проблемы ИО в плане того, какие механизмы задействованы для ее достижения, по аналогии с моделью OSI выделяются несколько уровней ИО, причем различные авторы выделяют от трех до девяти (и даже десяти) таких уровней [4, 5]. В ГОСТ 7498-1-99 [6] была принята 7-уровневая эталонная модель взаимосвязи открытых систем. В ГОСТ 55062-2012 была введена 3-уровневая эталонная модель ИО [2].

Анализ публикаций, касающихся классификации уровней ИО, показал различия в количестве выделяемых уровней и определениях их содержания, а также наличие различных подходов к их взаимной структуризации. Наглядное сопоставление этих уровней представлено в табл. 1.

Так, в ГОСТе 55062-2012 [2] отмечается, что «технический уровень описывает синтаксис или форматы передаваемой информации»; «семантический уровень описывает ... содержательную сторону обмениваемой информации»; «организационный уровень акцентирует внимание на прагматических аспектах взаимодействия (деловых или политических)».

В публикации [4] авторов упомянутого ГОСТа перечень уровней ИО полнее, а их содержание раскрыто более развернуто:

- «техническая ИО — способность к обмену цифровыми сигналами, поддержка согласованных (стандартизованных) интерфейсов, протоколов и механизмов

Таблица 1 Сопоставление уровней ИО в различных источниках

ГОСТ 55062 [2]	Работа [4]	Работа [3]	Обобщенные уровни ИО (за исключением отмеченного серым цветом, соответствуют классификации в работе [7])	
Техническая	Техническая	Физическая	Техническая	Физическая
	Синтаксическая	Синтаксическая		Синтаксическая
Семантическая	Семантическая	Семантическая		Семантическая
	Прагматическая			Прагматическая
Организа- ционная	Динамическая	Организа- ционная	Организа- ционная	Консолидационная
	Организационная			Координационная
	Концептуальная			
	Обменная			
	Интеграционная			

доступа к информационным ресурсам» — т. е. здесь уже речь идет только о протоколах передачи данных типа ТСР/ІР и т. п., а форматы данных перенесены в синтаксическую ІО. В таком контексте техническую ІО можно понимать как способность воспринимать потоки информации без знания того, как они устроены и что они собой представляют;

- «синтаксическая ІО — способность к обмену данными» — здесь, по всей видимости, имеется в виду способность ІС «понимать» форматы данных (в частности, файлов), т. е. способность анализировать и разбирать информационные потоки на составляющие их информационные сущности: значащие поля, разделители и прочие элементы, а также восстанавливать их взаимную структуру (иерархическую или др.);
- «семантическая ІО — способность к обмену информацией» — здесь, по видимому, идет речь об «осмыслении» отдельных структурных элементов данных и сопоставлении им тех или иных смыслов (в виде описаний или образов объектов), т. е. способность ІС определять смысл выделенных на предыдущем этапе информационных сущностей (отождествлять их с заложенными в данной ІС тезаурусами), чтобы в дальнейшем применять их в процессе своей работы в соответствующем качестве.

Далее в [4] рассматриваются еще несколько видов ІО (хотя в [5] они рассматриваются как подвиды семантической), в том числе «прагматическая — способность к совместному использованию информации в контексте решаемых задач», т. е. способность разнотипных, но не разнородных данных использоваться совместно в процессе решения той или иной конкретной задачи (отработке функции); определения остальных видов ІО, приведенные в [4] («динамическая», «организационная», «концептуальная», «на основе обмена», «интеграционная»), относятся скорее к организационному уровню проблемы ІО и в рамках задачи данного исследования интереса не представляют.

В работе [3] приводится такой перечень «уровней» ІО:

- физическая — «способность ... ІС ... к обмену сигналами ... на основе поддержки согласованных интерфейсов, коммуникационных протоколов и механизмов доступа к хранилищам данных»;
- синтаксическая — «способность ... ІС ... к обмену данными... на основе согласования кодов, форматов и типов данных»;
- семантическая — «способность ... ІС ... к согласованному функционированию на основе единой, недвусмысленной, адекватной интерпретации информации, полученной в результате обмена»;
- организационная — «способность ... бизнес-объектов ... использующих, возможно, различную информационную инфраструктуру, к согласованному функционированию» — т. е. данный уровень относится уже не к самим ІС, а к использующим их бизнес-субъектам.

Здесь следует еще раз оговориться, что данная работа ограничивается только техническим аспектом проблемы ИО, куда можно отнести физический, синтаксический и семантический уровни, а также прагматический (в понимании [4]). Такие уровни, как организационный, интеграционный и т. п., по всей видимости, следует отнести к организационным аспектам ИО. В работе [7] представлено схожее разделение модели для целей развития средств ИО в распределенных геоинформационных системах (ГИС): техническая (куда входят физическая, синтаксическая, семантическая) и организационная (куда включены консолидационная и координационная).

Теперь настало время сосредоточиться на понятии семантической ИО. Проанализируем ее определение из ГОСТа 55062-2012 [2]: «Семантическая интероперабельность (Semantic Interoperability): Способность любых взаимодействующих в процессе коммуникации информационных систем одинаковым образом понимать смысл информации, которой они обмениваются».

Здесь можно отметить, что если системы не могут одинаково (а точнее сказать, адекватно) понимать смысл (семантику) обмениваемой информации, то тогда о какой обработке данных может идти речь и, соответственно, какой смысл такого обмена информацией, что полезного получится в результате такого взаимодействия? Подмена выражения «одинаковым образом» на «адекватно» призвана подчеркнуть то обстоятельство, что система должна понимать смысл информации лишь в необходимом и достаточном для решения своих задач объеме, требовать от нее абсолютно одинакового и полного понимания всех тонкостей и особенностей обрабатываемых ею данных не только нецелесообразно, но, наверное, и невозможно.

К примеру, маршрутизатор передает пакеты, не «задумываясь» над тем, какой контент они несут (текст, картинку или что-либо другое), но ведь для процесса передачи он должен понимать, что это именно пакеты данных, причем определенного стандарта или протокола (например, TCP или UDP), и в процессе их передачи ему требуется выделять, анализировать и обрабатывать определенные поля пакетов, т. е. необходимый уровень семантической ИО здесь все же должен обеспечиваться¹.

Таким образом, ИО в практическом плане должна обеспечиваться на всех уровнях, вплоть до верхнего — семантического уровня. Подтверждение этого тезиса имеет место и в работе [1], где говорится, что «интероперабельность функциональных систем может быть только семантической». Для обеспечения уровня семантической ИО в обеих взаимодействующих системах должна быть заложена единая система (словарь, тезаурус) наименований тегов, атрибутов, ключевых слов и образуемых ими структур (классов или объектов), каждому из которых сопоставляется тот или иной смысл и, соответственно, возможности, правила и ограничения их обработки.

¹Хотя в то же время распознавание типа контента маршрутизатором желательно для расстановки приоритетов в зависимости от типа контента — что-то нужно передавать «в реальном времени», а что-то может «скачиваться» по остаточному принципу, в периоды простоя или низкой нагрузки.

На основе этого можно сделать заключение, что так называемые уровни ИО скорее следует понимать как неотделимые друг от друга аспекты (стороны) обеспечения ИО, так как без обеспечения какого-либо из них о практическом обеспечении ИО говорить не приходится.

Здесь же следует рассмотреть еще один аспект технической ИО — возможность совместного использования разнотипных и/или разноточных данных. Дело в том, что в процессе совместной обработки должно обеспечиваться согласование данных не только по смыслу, но и по формату (точности или подробности) их представления в той или иной ИС. Так, в разных ИС идентичные по смыслу атрибуты могут быть представлены в вещественных форматах `float` и `double`, а следовательно, для совместной обработки их необходимо преобразовывать к тому или иному типу, но в любом случае точность получаемого после обработки результата будет на уровне наименее точного типа. Другим моментом является необходимость совместно использовать данные, полученные различными способами и вследствие этого имеющие различную реальную точность (при этом оба вида данных могут использовать одинаковый тип, например `double`). Поэтому техническая ИО должна также обеспечивать возможность: (а) оперативной конвертации исходных данных к определенному типу (обеспечения однотипного представления данных) и (б) их согласования (выравнивания) по наименьшей реальной точности. Этот момент, по сути, подходит под определение прагматического уровня ИО, упомянутого в [4].

4 Структура интероперабельности: способы взаимодействия

Среди различных классификаций подходов [1, 8] к решению проблемы ИО вообще и семантической ИО в частности с точки зрения функционирования ИС можно выделить два основных способа организации взаимодействия.

1. Интероперабельность обмена данными заключается в обеспечении возможности обмена информационными объектами (файлами, документами, сообщениями и т. п.) между различными взаимодействующими агентами: ИС в целом, их подсистемами, модулями, выполняющимися функциями и т. п. Другими словами, как показано на рис. 2, система А использует систему Б лишь в качестве источника исходных данных, а саму функциональную задачу решает самостоятельно посредством собственных аппаратно-программных средств. Интероперабельность на уровне информационного обмена (или ИО межсистемных сообщений [1]) обеспечивается согласованием форматов, структуры, синтаксиса, семантики информационных объектов, циркулирующих между взаимодействующими ИС (представляемыми как черные ящики). Этот подход является достаточно ограниченным по сравнению со следующим.
2. Интероперабельность на уровне функциональных процессов обеспечивает возможность делегирования отдельных задач (связанных, например, с вы-

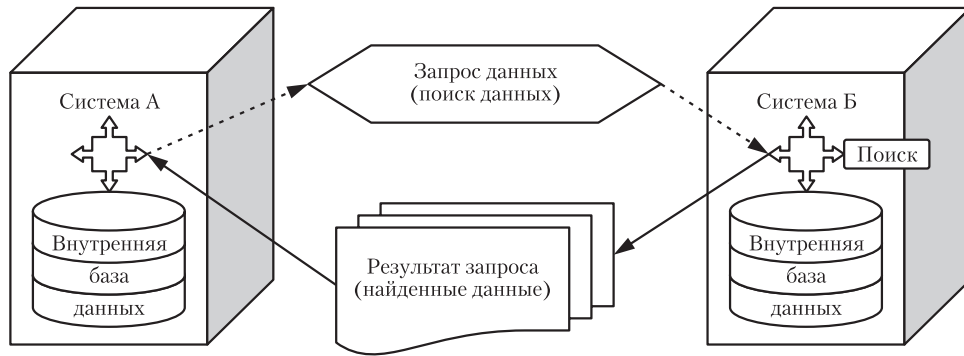


Рис. 2 Схема взаимодействия ИС посредством обмена данными

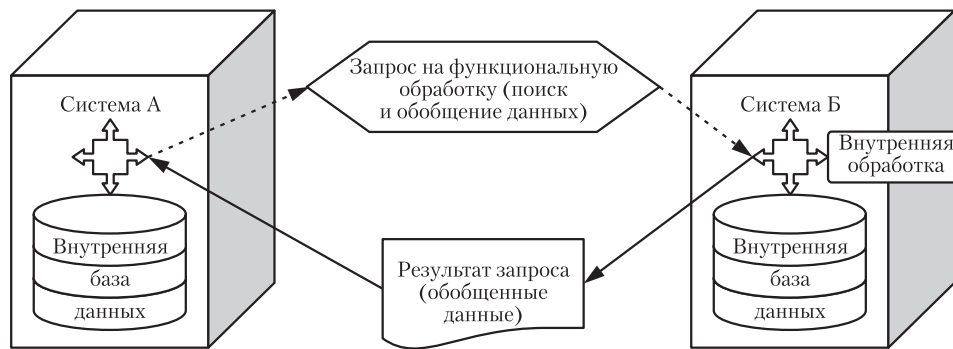


Рис. 3 Схема взаимодействия ИС посредством функциональных вызовов (запросов)

числительной обработкой, анализом или обобщением данных и т. п.) от одной системы к другой. Реализация такого вида ИО возможна посредством сервисно-ориентированной архитектуры (Service-Oriented Architecture, SOA) интероперабельной среды, которая обеспечивает межсистемное взаимодействие входящих в нее систем не только на уровне предоставления данных одной системой по запросу другой, но и обеспечивает выполнение определенных функций одной ИС для решения задач другой ИС. Так, на рис. 3 система А может использовать в процессе решения своих задач не только данные, но и определенную функциональность системы Б: например, система А отправляет системе Б запрос на решение той или иной задачи, система Б анализирует запрос, осуществляет подбор исходных данных и выполняет их обработку, возвращает конечный результат системе А. В это же самое время система Б может аналогичным образом использовать функциональность А для решения своих задач.

Положительным моментом ИО функциональных процессов является то, что делегирование функции позволяет выполнять ее в той же системе, в которой находятся и исходные данные для ее работы. Это исключает возможные потери при преобразовании исходных данных (как это может происходить в случае обмена только данными), а сама обработка происходит с использованием «родных» (оптимизированных для данной ИС) представлений данных и реализаций внутренних функций для их обработки, преобразованию подвергается только конечный результат работы функции, передаваемый системе-«заказчику»¹. Следствием этого является повышение эффективности, уменьшение межсистемного трафика, а также исключается межсистемный обмен исходными («внутренними») данными, которые могут быть конфиденциальны. Следует отметить, что это достоинство может не проявляться в случае, когда необходимо использовать исходные данные из нескольких источников.

В завершение отметим еще несколько неясных моментов: согласно расхожему определению [9], SOA представляет собой «набор физически независимых, слабо связанных, заменяемых приложений». В связи с этим можно отметить:

- если приложения «физически независимы», то о какой их связи может идти речь? Скорее всего, речь идет о том, что они создаются и эксплуатируются независимыми коллективами по разным технологиям и с разными целями, однако все они должны создаваться в рамках идеологии SOA, т. е. и в этом случае определенная степень зависимости имеет место;
- «слабо связанных» — «слабо» не является критерием, скорее здесь нужно понимать, что взаимодействие осуществляется через, насколько это возможно, простые интерфейсы (использующие минимум необходимых полей, предельно простой структуры), так как чем проще интерфейс, тем он более универсален и прост в обеспечении ИО. Вместе с тем нужно иметь в виду, что ИО как средство универсализации наряду с положительным моментом — возможностью обеспечить обмен разнородной информацией между различными системами для решения произвольных задач — имеет и обратную сторону — усложнение механизма их работы и, соответственно, общее снижение эффективности выполнения этих процессов;
- «заменяемость» — скорее всего, основным аргументом для ее обеспечения следует считать не столько взаимозаменяемость альтернативных по функциональности блоков (хотя и это может иметь место), сколько возможность переконфигурировать систему под новые задачи или технологические схемы,

¹Парадигма функциональной ИО справедлива и для взаимодействия организаций: так, зачастую проще заказать выполнение тех или иных работ организации-владельце исходных данных (у которой к тому же может быть отлажен и поставлен «на поток» соответствующий технологический процесс и которая во всяком случае лучше, чем кто-либо другой, ориентируется в «собственных» данных), чем приобретать только исходные данные и самостоятельно осваивать все тонкости процесса их обработки.

например в соответствии с меняющимися требованиями бизнеса или при ее интеграции с другими системами.

5 Информационные объекты в проблеме интероперабельности

Попробуем теперь определиться со спецификой объектов ИО. Выше шла речь об ИО как при обмене данными между ИС, так и при функциональном взаимодействии ИС. В первом случае объектами ИО были только данные, а во втором объектами выступали также функциональные подсистемы и отдельные функции ИС.

Вообще говоря, употребление термина «данные» в контексте ИО не совсем корректно, здесь скорее нужно говорить об ИО типов данных по форме представления и по семантике:


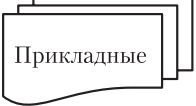
- типизация формы представления данных включает следующие моменты: общее понимание типа данных (например, вещественное, целочисленное, текстовое), конкретный способ компьютерного представления (формат) данных («тип данных» в понимании программиста: double, float, int32, int64 и т. п.), а также типы и форматы более сложных структур: файлов и баз данных;
- типизация семантики данных, определяющая их прикладной смысл, например: личные данные («пол», «дата рождения»), геоданные («высота», «превышение», «ширина», «численность населения») и другие данные, которые могут иметь различную форму представления в различных ИС.

Структурируя множество формальных и семантических типов данных в плане степени обеспечения их ИО, можно выделить (табл. 2):

- типы данных, носящих достаточно абстрактный характер и поэтому являющихся универсальными (и ставших де-факто стандартными) для многих предметных областей, например: текстовые файлы (и форматы txt, xml), изображения (и связанные с ним форматы bmp, tiff, jpg и др.), таблицы, записи и поля баз данных и т. п.;
- прикладные типы данных, специфичных для конкретной предметной области (библиографические описания, формы финансовой отчетности, пространственные данные и т. п.). Эта специфичность предполагает более глубокую структуризацию и вследствие этого множество вариантов реализации представления данных для одной и той же предметной сущности, обусловленных различными подходами, адаптацией под определенные функциональные процессы, оптимизацией под ту или иную аппаратно-программную специфику.

Таким образом, складывается парадоксальная ситуация: несмотря на то что абстрактный класс имеет дело с более широким кругом пользователей, чем каждый из прикладных, в плане обеспечения ИО абстрактные типы данных гораздо более благополучны.

Таблица 2 Классификация объектов ИО

Тип данных	Данные	Функциональное ПО
 Абстрактные	Тексты (txt, xml), Изображения (jpg, bmp, tif)	Интернет-браузеры, текстовые и графические редакторы
	Таблицы, записи и поля баз данных	СУБД (запросы SQL)
 Прикладные	Структурированные библиографические описания (MARK-21)	Информационно-аналитиче- ские системы научного цитиро- вания (Web of Science, Scopus, РИНЦ и т. п.)
	Электронные формы бухгалтерской отчетности	Бухгалтерские системы (типа 1С)
	Электронные карты и другие пространственные данные (SXF, SHP, MIF/MID, KML)	ГИС (Панорама, ArcInfo, MapInfo, Google Earth, GoogleMaps и т. п.)

Аналогичная ситуация складывается и для функционального взаимодействия: программы (программные продукты), функционирующие в составе ИС, также можно разделить на два класса [1]:

- программные продукты, для которых характерна абстрактность обрабатываемой ими информации, например: текстовый процессор, интернет-браузер, система управления базами данных (СУБД) и т. п. Они в плане ИО также достаточно благополучны; так, «интероперабельность текстового процессора и т. п. настолько логична, бесшовна, глубока и адекватна, что миллионы людей пользуются ими, не задумываясь»;
- программные продукты, обрабатывающие информацию на основе анализа ее семантики, формы представления, очередности ее поступления, что предполагает заложенное в эти продукты глубокое знание предметной области и специфических функциональных процессов. Характерным примером таких продуктов являются ГИС, где заложены знания из областей географии (в частности, системы классификации геообъектов), геометрии (пространственные формы и топологические сети, методы их анализа и обработки), а также процессы картосоставления и способы машинной графики. Соответственно, для таких программ реализация ИО связана с сильной зависимостью между данными (их структурой, семантикой) и алгоритмами их обработки, а вследствие этого — с большим разнообразием интерфейсов функциональных вызовов. Все это подкрепляет тезис о том, что ИО функциональных систем может быть только семантической.

Таким образом, в функциональном аспекте также просматривается деление на абстрактный и множество прикладных классов, хотя граница между теми

и другими достаточно размыта. При этом, говоря об обеспечении функциональной ИО, следует иметь в виду, что она включает в себя обеспечение ИО и на уровне данных, т. е. является информационно-функциональной.

6 Проблемы обеспечения семантической интероперабельности

Подводя итоги, можно сказать, что основными типичными проблемами в достижении семантической ИО являются:

- (1) концептуальная несовместимость, нецелостность, противоречивость, неконвенциональность описания предметной области и представления данных в рамках разных систем; отсутствие комплексной безопасности фрагментарных систем и межсистемного информационного взаимодействия;
- (2) проблемы ИО разноплатформенных данных и интерфейсов унаследованных и вновь разрабатываемых программных комплексов;
- (3) несовместимость (и затрудненность ее устранения), возникшая вследствие допущений и ошибок, возникших на разных этапах жизненного цикла ИС — при ее проектировании, реализации и особенно при ее модификациях под вновь возникающие задачи (зачастую спонтанных).

К числу сопутствующих проблем можно отнести: избыточное описание предметов и процессов предметной области (наличие дублирующих или зависимых полей в рамках одной ИС), что вызывает необходимость их согласования; наличие в системах особых механизмов доступа к данным.

Если говорить о семантическом аспекте задачи обеспечения ИО, то ее сущность заключается в создании методологии и механизмов согласования идентичных или близких по смыслу полей и/или структур данных в разных системах, как это показано на рис. 4. В общем случае смыслы данных могут не иметь строгого соответствия, что потенциально вызывает потерю части смысловой нагрузки или искажение смысла. Это может быть несущественно при решении одних задач, но неожиданно стать «камнем преткновения» при решении других. Таким образом, возникает необходимость определения понимания семантической близости данных и ее меры. Кроме того, необходимо иметь в виду ситуации (см. рис. 4), когда одиночному полю системы А (например, неструктурированному описанию) в системе Б может соответствовать структура полей (структурированное описание), а в системе В — структура другого типа, не в полной мере сопоставимая с Б; другими словами, необходим механизм сопоставления данных по структурному и формальному признакам.

Практическая реализация интероперабельной среды как посредством SOA, так и обменом только данными имеет определенные проблемы, перечисленные, например, в [1]. Среди них можно отметить следующие:

- такая интеграция зачастую заключается в использовании функций существующих систем, а не в создании новых сервисов. В этом случае даже при

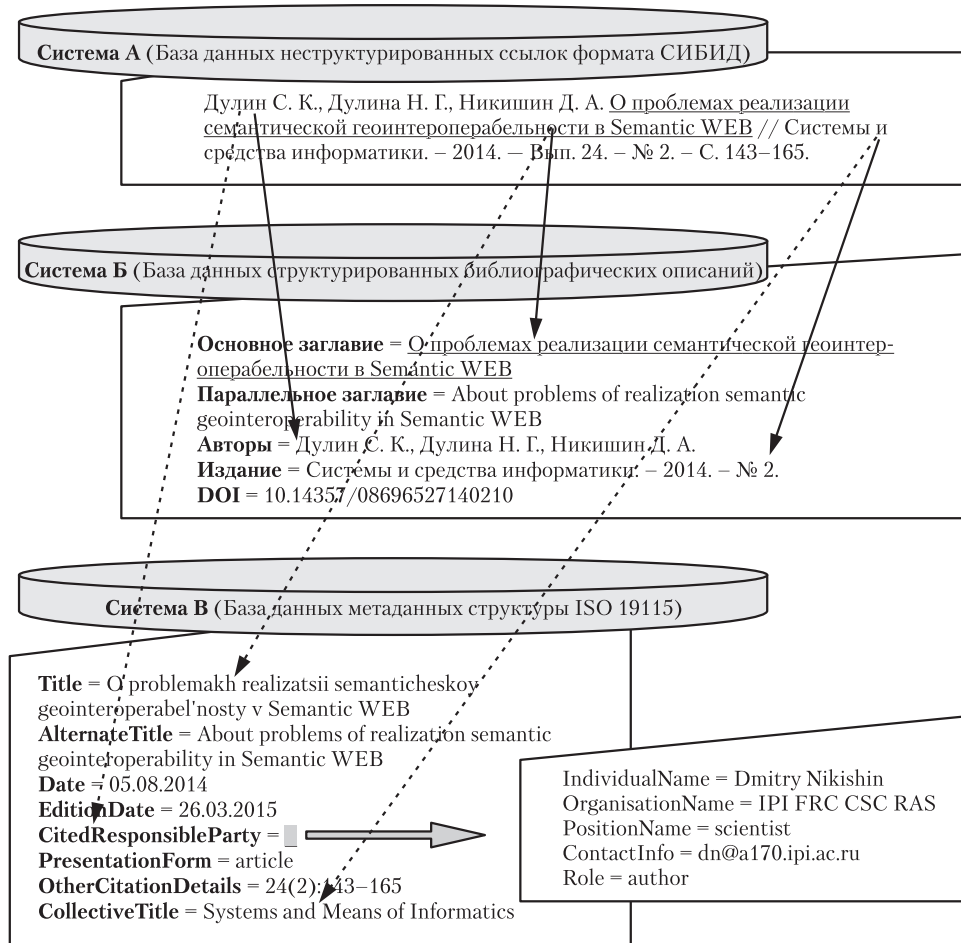


Рис. 4 Схема идентичности полей и элементов данных в различных ИС

использовании SOA программное обеспечение (ПО) ее подсистем будет обладать традиционным жизненным циклом, что оставляет нерешенной проблему преемственности от унаследованных ИС, неизбежные различия их реализации разными производителями, а также не учитывает динамику изменения требований к объединенным в SOA информационным системам;

- необходимость обеспечения попарной трансляции данных и функциональных запросов между каждой парой взаимодействующих систем требует создания должного количества трансляторов [10] и их интеграции в каждую из задействованных систем. Поэтому ИО часто пытаются реализовать посредством

дополнительных интегрирующих «надстроек»¹, которые могут включать адаптеры, хабы, среды передачи сообщений, единое хранилище общих данных, бизнес-моделлеры, онтологии баз данных/знаний и т. п. [1].

Все это вызывает значительные материальные и временные затраты на создание, эксплуатацию и модернизацию такой системы, превышающие суммарную стоимость отдельных компонентов; сложность системы неизбежно сказывается на ее надежности.

В связи с этим можно отметить, что:

- построение интероперабельной среды по централизованному принципу неизбежно вступает в конфликт с требованием гибкости, так как при каждом добавлении в систему нового компонента требуется перестройка ее «ядра»;
- интероперабельная система должна быть в максимальной степени унифицирована, т. е. использовать минимум предельно простых интерфейсов, стандартов и т. п., что вызывает необходимость модернизации существующих систем.

Кроме того, проблема ИО усугубляется при масштабировании, т. е. при переходе на новый уровень сложности, например от корпоративных стандартов и технологий к стандартам отрасли, т. е. от стандартов, предназначенных для решения насущных задач корпорации (и вследствие этого достаточно однородных), создаваемых и применяемых единоначально, к стандартам, ориентированным на решение всевозможных задач, полиморфных, для создания и внедрения которых необходимо согласие между всеми взаимодействующими сторонами.

7 Проблемы стандартизации в области интероперабельности

Согласно ГОСТу 55062-2012 [2] «основным способом решения проблемы интероперабельности или «прозрачности» гетерогенной среды выступает последовательное применение принципов открытых систем и методологии функциональной стандартизации». Основными объектами стандартизации выступают так называемые «ключевые интерфейсы», которые можно разделить на *горизонтальные* — между системами и компонентами одного ранга — и *вертикальные* — между системами разных рангов [4].

В этой связи следует отметить, что предлагаемые на данный момент стандарты на метаданные пространственных данных (международный ISO 19115-2003 [11] и отечественный ГОСТ 52573-2006 [12]) весьма объемны (в ГОСТе содержится почти 200 полей, а в ISO их более 400), а предусмотренные ими универсальные

¹Иллюстрацией такой надстройки на рис. 1 служит блок «универсальный запрос», который, вообще говоря, должен решить, в какой последовательности запросить системы Б и В (при этом если они используют различные интерфейсы, то адаптировать универсальный запрос для каждой из них), затем получить ответы от них и синтезировать из них результат, который будет возвращен системе А. Эту функциональность символизирует Т-образная стрелка справа от него.

подходы к описанию пространственных данных любой природы обуславливают их сложную структуру, что делает стандартизацию по ним весьма трудоемкой. Проблема сложности использования универсального стандарта метаданных косвенно поднимается, например, в [13]: «В стандартах на метаданные (в частности, ГОСТ 52573), предусматриваются обязательные поля, информация для заполнения которых на практике зачастую отсутствует. Так, при практической реализации инфраструктуры пространственных данных ... возникла проблема с отсутствием в реальных данных информации для заполнения ряда полей, являющихся обязательными согласно стандарту ГОСТ 52573, поэтому ... такие поля были переклассифицированы как рекомендованные к заполнению».

Еще одной проблемой, упоминаемой в [13], является то, что существует ряд ведомственных систем описания геоданных, например кадастровые, адресные, авто/железнодорожные и т. п. Они нацелены на стоящие перед ними специфические требования и задачи и в большинстве своем несовместимы друг с другом. Здесь также играют роль ведомственные интересы.

На основе этого в качестве обоснований для совершенствования системы стандартизации метаданных можно упомянуть следующее:

- вместо единого стандартного набора метаданных на все возможные случаи жизни представляется целесообразным использовать систему локальных стандартов для специфических структур метаданных, определяемых, в свою очередь, спецификой данных (их моделью, точностью, предназначением и т. п.) и предназначенных для соответствующих типов поисковых запросов;
- семантику данных и метаданных можно подразделить на прикладную, относящуюся к предметной области данных, и на модельную, относящуюся к способу организации данных [14], поэтому при стандартизации семантики следует выделять прикладную и модельную составляющие, которые должны обладать возможностью фигурировать независимо друг от друга: например, прикладная составляющая географического смысла данных должна быть независима от модели их метрического представления;
- стандартизация должна подразумевать не только состав и форму представления данных/метаданных, но и их достоверность, точность и полноту [15], для чего требуется сформулировать современные требования к геоданным;
- стандартизация должна распространяться не только на состав данных или метаданных, но и на форму и способы выполнения поисковых запросов, т. е. запрос должен определенным образом обрабатывать наличие/отсутствие определенных признаков поиска у соответствующих данных. Развивая тему, поисковый запрос можно даже обязать искать альтернативные пути поиска, например при отсутствии у некоторого геообъекта «мост» поля «длина» может быть сделана попытка определить его длину по координатам концов моста;
- стандартизация функционального взаимодействия должна распространяться не только на форму представления данных, но и на процессы их преобразова-

ния, например, с целью обеспечения соответствующей разрядности и точности вычислений при численных преобразованиях.

Важным требованием является обеспечение доступности этой системы стандартов для их конечных пользователей (операторов обрабатывающих комплексов), т. е. содержание стандартов должно излагаться на языке пользователя, в доступной для технического специалиста и сжатой форме.

Внедрение стандартов также затрудняет существующая в настоящее время практика «продажи» полнотекстовых описаний стандартов, что наряду со значительными материальными затратами организаций на их приобретение требует непроизводительных временных и трудовых затрат на анализ закупленного и совершение новых закупок, поскольку специалист зачастую не может заранее определить, какие именно стандарты ему будут необходимы и достаточны для его работы. Это приводит к неполному (ограниченному) или даже к некорректному внедрению стандартов, что в дальнейшем порождает проблемы взаимодействия организаций и их ИС.

8 Анализ проблемы понимания семантики геоданных

Сделаем попытку рассмотреть проблему понимания вообще, а также понимание семантики геоданных, опираясь на материал [16]. Там приводится ряд определений понятия понимания, обобщая которые для задач обеспечения ИО ИС можно индусировать следующее определение:

Понимание (усвоение) — соотнесение, сравнение вновь поступающей информации (текста, образа, речи, поведения, явления и т. п.) с уже известной, накопленной и структурированной информацией (т. е. сутью знанием) посредством когнитивных структур, схем поведения, понятий и категорий понятий; и ее оценка с определенных позиций, на основе некоторого образца, стандарта, нормы, принципа и т. п., например накопленного опыта или обстоятельств, в которых происходит процесс восприятия и понимания, особенностей когнитивной системы/когнитивного стиля человека.

Непонимание чаще всего обусловлено необоснованностью, неочевидностью или отсутствием такого стандарта, а также может быть вызвано возникающим когнитивным диссонансом (в терминологии ИС — нарушением целостности существующего знания), а также, в случае осуществления процесса понимания человеком, субъективными факторами восприятия (неблагоприятными условиями для восприятия и понимания/усвоения информации). В последнем случае понимание также может происходить не сразу, а с задержкой, например:

- по прошествии определенного времени — было время «подумать» или создались благоприятные условия для обдумывания;
- при изменении условий — была получена дополнительная информация, скажем, произошли определенные события, что дополнило, «протолкнуло» недопонятую ранее информацию, «поставило все на свои места».

Понимание (а точнее сказать, *знание*) дает возможность осуществлять действия с новыми предметами на основе уже сложившихся умений и навыков, а также рефлексировать о понятой/усвоенной информации (*знании*), выводить из нее разные следствия.

Касательно процесса понимания (интерпретации) можно сформулировать следующее: автор (генератор) выражает свои субъективные образы или суждения в формализованном виде (текст, речь), а интерпретатор восстанавливает этот образ или суждение в собственном субъективном контексте. Этот субъективный контекст определяется предыдущим опытом субъекта и текущими обстоятельствами, в которых происходит интерпретация. Таким образом, язык выступает «универсальной онтологией», которая обеспечивает ИО субъектов — носителей языка (точнее говоря, пользователей языка).

Для генератора не существует (скорее она для него незаметна) проблемы омонимии слов/терминов (неоднозначности используемого слова/термина), поскольку он знает подразумеваемое под ним значение. А вот интерпретатор сталкивается с такой проблемой и вынужден ее решать или посредством опоры на контекст сообщения, а если это не помогает, то вынужден использовать вероятностные (частотные) или нечеткие оценки (типа «автор, по всей видимости, имел в виду следующее» или «автор мог иметь в виду либо то, либо это»). Поиск неявно/нечетко выраженного содержания происходит в результате обращения к знаниям, имеющим отношение к данному тексту.

Семиотическая коммуникация предполагает участие отправителя и получателя информационного сообщения, причем в процессе общения каждый из них выступает в роли то генератора, то интерпретатора сообщения.

В практическом плане проблема понимания возникает, когда взаимодействующие агенты сталкиваются с «ситуациями неопределенности ключевых понятий, с нечеткими постановками задач, с лексической полисемией и другими препятствиями к однозначной и эксплицитной трактовке терминов, их смысловых связей и их восприятия специалистами из разных областей» [17].

9 Заключение

Основной задачей данной работы ставилась выработка целостного и по возможности всестороннего взгляда на проблему ИО. Подводя итог вышеизложенному, представляется важным отметить следующее.

Рассматривая структуру ИО, можно выделить две обобщенных группы ее аспектов: техническую ИО, включающую физический, синтаксический, семантический и прагматический аспекты, и организационную ИО, включающую вопросы организации межсистемного взаимодействия (например, бизнес-процессы, корпоративные правила и т. п.). Важным представляется тезис о том, что перечисленные моменты являются именно аспектами, а не уровнями ИО, так как без обеспечения какого-либо из них говорить о практическом обеспечении ИО не

приходится. А в качестве уровней обеспечения ИО выступают способы организации взаимодействия: информационный (только обмен данными) и информационно-функциональный (делегирование функциональных задач информационным системам-партнерам).

Определение «семантическая интероперабельность» можно понимать как способность взаимодействующих ИС адекватно понимать смысл как самих данных, так и вызовов функциональных процессов, фигурирующих в межсистемном взаимодействии.

Важным моментом является выделение прагматической ИО как возможности совместного использования в одном функциональном процессе данных, идентичных по смыслу, но различных по форме представления (подробности, разрешающей способности) или по степени точности. Основными типичными проблемами в достижении ИО, в частности семантической, являются:

- концептуальная несовместимость, нецелостность, противоречивость, неконвенциональность и тому подобные проблемы описания предметных областей;
- проблемы, возникающие вследствие разноплатформенных представлений данных и интерфейсов в унаследованных и вновь разрабатываемых программных комплексах;
- проблемы, возникшие вследствие допущений и ошибок, возникших при проектировании, реализации и модификации той или иной ИС.

Проблема ИО усугубляется при масштабировании, например при переходе от корпоративных стандартов и технологий, разрабатываемых и внедряемых единовременно, к стандартам отрасли (особенно в международном масштабе), для создания и внедрения которых необходимо согласие между всеми взаимодействующими сторонами.

Анализ проблемы стандартизации метаданных в плане обеспечения ИО наталкивает на следующие выводы:

- рациональным подходом представляется использование системы локальных стандартов для специфических структур метаданных и данных, при этом важным требованием является обеспечение доступности этой системы для их конечных пользователей (операторов обрабатывающих комплексов);
- стандартизация должна определять не только состав и форму представления данных/метаданных, но и их достоверность, точность и полноту, для чего требуется сформулировать современные требования к геоданным;
- стандартизация должна распространяться не только на состав и способы описания данных/метаданных, но и на форму и способы выполнения поисковых запросов, в том числе правила обработки отсутствующих или приведения к «общему знаменателю» альтернативных или семантически-зависимых атрибутов поиска, а также на форму межсистемных функциональных вызовов (для систем типа SOA).

Анализ проблемы понимания приводит к осознанию того, что для обеспечения уровня семантической ИО во всех взаимодействующих системах должна быть заложена единая система (словарь, тезаурус) наименований тегов, атрибутов, ключевых слов и образуемых ими структур (классов или объектов), каждому из которых сопоставляется тот или иной смысл, и, соответственно, возможности, правила и ограничения их обработки.

На основании анализа и структурирования проблемы ИО можно сделать предположение, что интероперабельное функционирование множества разнородных ИС возможно в рамках объединяющей их гиперсистемы. Функциями последней должны стать обеспечение и контроль общих требований и стандартов, предъявляемых к каждой из систем-участниц, и администрирование тех функциональных процессов, которые относятся к гиперсистеме в целом (т. е. когда имеет место межподсистемное взаимодействие).

Каждую из подсистем можно рассматривать как черный ящик, детали реализации и функционирования которой мало интересуют гиперсистему и остаются на совести организаций-операторов этих подсистем. Для гиперсистемы важно лишь то, что ее подсистемы как черный ящик будут адекватно реагировать на определенное множество внешних запросов, сформулированных в общей идеологии и терминологии интероперабельной гиперсистемы. Однако построение интероперабельной системы по централизованному принципу неизбежно вступает в конфликт с требованием гибкости, так как при каждом добавлении в систему нового компонента требуется перестройка ее «ядра».

Изложенный материал предполагается использовать как теоретическое обоснование для разрабатываемой авторами модели обеспечения согласованного понимания пользователями семантики геоданных.

Литература

1. *Бородакий Ю. В., Лободинский Ю. Г.* К проблеме обеспечения интероперабельности // Информационные технологии и вычислительные системы, 2009. № 5. С. 16–24.
2. ГОСТ Р 55062-2012. Системы промышленной автоматизации и их интеграция. Интероперабельность. Основные положения. — М.: Стандартинформ, 2012. 20 с.
3. *Батоврин В. К., Королев А. С.* Способ количественной оценки интероперабельности // Открытые системы. Интероперабельность, 2009. № 5. С. 91–96.
4. *Гуляев Ю. В., Журавлев Е. Е., Олейников А. Я.* Методология стандартизации для обеспечения интероперабельности информационных систем широкого класса // Журнал радиоэлектроники, 2012. № 3. <http://jre.cplire.ru/jre/mar12/2/text.html>.
5. *Журавлев Е. Е., Иванов С. В., Каменщиков А. А., Олейников А. Я., Разинкин Е. И., Рубан К. А.* Интероперабельность в облачных вычислениях // Ж. радиоэлектроники, 2013. № 9. <http://jre.cplire.ru/jre/sep13/4/text.html>.
6. ГОСТ Р ИСО/МЭК 7498-1-99. Информационная технология. Взаимосвязь открытых систем. Базовая эталонная модель. Часть 1. Базовая модель. — М.: Стандартинформ, 1999. 62 с.

7. *Гришенцев А. Ю., Коробейников А. Г.* Средства интероперабельности в распределенных геоинформационных системах // Ж. радиоэлектроники, 2015. № 3. <http://jre.cplire.ru/jre/mar15/7/text.html>.
8. *Франгулова Е. В.* Классификация подходов к интеграции и интероперабельности информационных систем // Вестник АГТУ. Сер. Управление, вычислительная техника и информатика, 2010. № 2. С. 176–180.
9. *Лис К.* Мета модель сервисов для интероперабельности систем, поддерживающих сервисно-ориентированную архитектуру. <http://tvvlibrary.narod.ru/papers/2010/27.pdf>.
10. *Дулин С. К., Дулина Н. Г., Никишин Д. А.* Особенности моделей геоданных и методов их обработки в аспекте обеспечения семантической геоинтероперабельности // Информационные технологии, 2015. Т. 21. № 3. С. 224–235.
11. ISO 19115:2003. Geographic information — Metadata. — Oslo, Norway: Norwegian Technology Centre, 2002. 147 p.
12. ГОСТ Р 52573-2006. Географическая информация. Метаданные (утв. Приказом Ростехрегулирования от 28.09.2006 N 216-ст). — М.: Стандартинформ, 2006. 71 с.
13. *Анисимова О. Л., Пестов И. Д., Серебряков С. В. и др.* Информационное и программное обеспечение управления метаданными на базе регионального центра УрФО // Пространственные данные, 2009. № 2. С. 26–34.
14. *Калиниченко Л. А.* Синтез канонических моделей, предназначенных для достижения семантической интероперабельности неоднородных источников информации // Системы и средства информатики. Спец. вып.: Формальные методы и модели в композиционных инфраструктурах распределенных информационных систем. — М.: ИПИ РАН, 2005. С. 11–39.
15. *Комков А. М., Николаев С. А., Шилов Н. И.* Составление и редактирование карт. — М.: ВИА, 1958. 248 с.
16. *Тарасов В. Б.* Проблема понимания: настоящее и будущее искусственного интеллекта // Открытые семантические технологии проектирования интеллектуальных систем = Open Semantic Technologies for Intelligent Systems (OSTIS-2015): Мат-лы V Междунар. науч.-технич. конф. — Минск: БГУИР, 2015. С. 25–42.
17. *Кожунова О. С.* Когнитивная интероперабельность экспертного взаимодействия в задаче обработки русско-французских параллельных текстов: лингвокогнитивные аспекты // Информатика и её применения, 2013. Т. 7. Вып. 3. С. 81–93.

Поступила в редакцию 19.01.16

PROBLEMS OF MAINTENANCE OF SEMANTIC GEOINTEROPERABILITY AND COORDINATION OF UNDERSTANDING OF GEODATA SEMANTICS

S. K. Dulin¹, N. G. Dulina², and D. A. Nikishin¹

¹Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation

²A. A. Dorodnicyn Computing Center, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 40 Vavilov Str., Moscow 119333, Russian Federation

Abstract: The primary goal of the work was to develop a complete and, whenever possible, all-round point of view on the problem of interoperability of cooperating information systems in general and the semantic aspect of this problem in particular. The paper considers principal causes of occurrence of the problem of intersystem interoperability, its various definitions, and approaches to its structurization. An attempt to structure the problem of interoperability is made, the ambiguous moments are discovered and possible variants of their elimination are suggested. The state-of-the-art is analyzed in the field of standardization of metadata. The conclusion about necessity of standardization of forms and ways of performance of search inquiries and intersystem functional calls is made. The basic moments of the problem of understanding of sense of the data, existing approaches to its structurization, and solving are considered.

Keywords: geodata; semantic geointeroperability; ontologies; Semantic Web

DOI: 10.14357/08696527160107

Acknowledgments

The work was supported by the Russian Foundation for Basic Research (project 14-07-00040).

References

1. Borodakiy, Yu. V., and Yu. G. Lobodinskiy. 2009. K probleme obespecheniya inter-operabel'nosti [To a problem of ensuring interoperability]. *Informatsionnye Tekhnologii i Vychislitel'nye Sistemy* [Information Technologies and Computing Systems] 5:16–24.
2. GOST R 55062-2012. 2012. Sistemy promyshlennoy avtomatizatsii i ikh integratsiya. Interoperabel'nost'. Osnovnye polozheniya [Systems of industrial automation and their integration. Interoperability. Basic statements]. Moscow: Standartinform Publ. 20 p.

3. Batovrin, V. K., and A. S. Korolev. 2009. Sposob kolichestvennoy otsenki interoperabel'nosti [Sposob of a quantitative assessment of interoperability]. *Informatsionnye Tekhnologii i Vychislitel'nye Sistemy* [Information Technologies and Computing Systems] 5:91–96.
4. Gulyaev, Yu. V., E. E. Zhuravlev, and A. Ya. Oleynikov. 2012. Metodologiya standartizatsii dlya obespecheniya interoperabel'nosti informatsionnykh sistem shirokogo klassa [Methodology of standardization for ensuring interoperability of information systems of a wide class]. *Zh. Radioelektroniki* [J. Radioelectronics] 3. Available at: <http://jre.cplire.ru/jre/mar12/2/text.html> (accessed January 19, 2016).
5. Zhuravlev, E. E., S. V. Ivanov, A. A. Kamenshchikov, A. Ya. Oleynikov, E. I. Razinkin, and K. A. Ruban. 2013. Interoperabel'nost' v oblachnykh vychisleniyakh [Interoperability in cloud computing]. *Zh. Radioelektroniki* [J. Radioelectronics] 9. Available at: <http://jre.cplire.ru/jre/sep13/4/text.html> (accessed January 19, 2016).
6. GOST R ISO/MEK 7498-1-99. 1999. Informatsionnaya tekhnologiya. Vzaimosvyaz' otkrytykh sistem. Bazovaya etalonnaya model'. Chast' 1. Bazovaya model' [Information technology. Interrelation of open systems. Basic reference model. Part 1. Basic model]. Moscow: Standartinform Publ. 62 p.
7. Grishentsev, A. Yu., and A. G. Korobeynikov. 2015. Sredstva interoperabel'nosti v raspredelennykh geoinformatsionnykh sistemakh [Means of interoperability in the distributed geographic information systems]. *Zh. Radioelektroniki* [J. Radioelectronics] 3. Available at: <http://jre.cplire.ru/jre/mar15/7/text.html> (accessed January 19, 2016).
8. Frangulova, E. V. 2010. Klassifikatsiya podkhodov k integratsii i interoperabel'nosti informatsionnykh sistem [Classification of approaches to integration and interoperability of information systems]. *Vestnik AGTU. Ser. Upravlenie, vychislitel'naya tekhnika i informatika* [AGTU Bulletin. Ser. Management, computer facilities, and informatics] 2:176–180.
9. Lis, K. 2010. Metamodel' servisov dlya interoperabel'nosti sistem, podderzhivayushchikh servisno-orientirovannuyu arkhitekturu [Metamodel of services for interoperability of the systems supporting service-oriented architecture]. Available at: <http://tvvlibrary.narod.ru/papers/2010/27.pdf> (accessed January 19, 2016).
10. Dulin, S. K., N. G. Dulina, and D. A. Nikishin. 2015. Osobennosti modeley geodannykh i metodov ikh obrabotki v aspekte obespecheniya semanticheskoy geointeroperabel'nosti [Features of models of geodata and methods of their processing in aspect of ensuring semantic geointeroperability]. *Informatsionnye Tekhnologii* [Information Technologies] 21(3):224–235.
11. ISO/FDIS 19115:2002(E). 2003. Geographic information — Metadata. Norway, Majorstuen: Norwegian Technology Centre. 147 p.
12. GOST R 52573-2006. 2006. *Geograficheskaya informatsiya. Metadannye* [Geographical information. Metadata]. Moscow: Standartinform Publ. 71 p.
13. Anisimova, O. L., I. D. Pestov, and S. V. Serebryakov. 2009. Informatsionnoe i programmnoe obespechenie upravleniya metadannymi na baze regional'nogo tsentra UrFO [Information and program support of management of metadata on the basis of the UrFO regional center]. *Prostranstvennye Dannye* [Spatial Data] 2:26–34.
14. Kalinichenko, L. A. 2005. Sintez kanonicheskikh modeley, prednaznachennykh dlya dostizheniya semanticheskoy interoperabel'nosti neodnorodnykh istochnikov informatsii [Synthesis of the initial models intended for achievement of semantic interoperability]

- of nonuniform sources of information]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 15:11–39.
15. Komkov, A. M., S. A. Nikolaev, and N. I. Shilov. 1958. *Sostavlenie i redaktirovanie kart* [Drawing up and editing maps]. Moscow: VIA Publ. 248 p.
 16. Tarasov, V. B. 2015. Problema ponimaniya: Nastoyashchee i budushchee iskusstvennogo intellekta [Problem of understanding: Present and future of artificial intelligence]. *5th Conference (International) “Open Semantic Technologies for Intelligent Systems” Proceedings*. Minsk. 25–42.
 17. Kozhunova, O. S. 2013. Kognitivnaya interoperabel’nost’ ekspertnogo vzaimodeystviya v zadache obrabotki russko-frantsuzskikh parallel’nykh tekstov: Lingvokognitivnye aspekty [Cognitive interoperability of expert interaction in a problem of processing of the Russian-French parallel texts: Linguistic aspects]. *Informatika i ee Primeneniya — Inform. Appl.* 7(3):81–93.

Received January 19, 2016

Contributors

Dulin Sergey K. (b. 1950) — Doctor of Science in technology, professor; leading scientist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; s.dulin@ccas.ru

Dulina Natalia G. (b. 1947) — Candidate of Science (PhD) in technology, senior scientist, A. A. Dorodnicyn Computing Center, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 40 Vavilov Str., Moscow 119333, Russian Federation; ngdulina@mail.ru

Nikishin Dmitry A. (b. 1976) — Candidate of Science (PhD) in technology, Head of Laboratory, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; dmnikishin@mail.ru

НЕКОТОРЫЕ ВОПРОСЫ ПРАКТИЧЕСКОЙ РЕАЛИЗАЦИИ КОНЦЕПЦИИ SDN

*В. Б. Егоров*¹

Аннотация: Подвергается разностороннему анализу широко обсуждаемая в последнее время концепция создания программно-конфигурируемых сетей SDN (software-defined networking). Показана неочевидность ряда распространенных суждений о сетях, построенных на принципах SDN, и безусловность таких широко рекламируемых их преимуществ, как ускорение и удешевление создания, упрощение масштабирования и конфигурирования, повышение эффективности функционирования и управления при сокращении эксплуатационных затрат. Отмечены такие существенные, но обычно не акцентируемые особенности сетей SDN, как тяга к территориальной компактности, необходимость индивидуального прикладного программирования каждой конкретной сети, желательность выделенных доверенных каналов управления устройствами сетевой инфраструктуры, замедленная коммуникация между плоскостями данных и управления, сложность обеспечения надежности и доступности сети из-за централизации управления, объективные трудности сопряжения с существующими традиционными сетями. Сделана попытка объяснения причин нынешней популярности концепции SDN, а также дана оценка перспектив и условий практической реализации концепции SDN в будущем.

Ключевые слова: NFV; RUNOS; SDN; виртуализация сетевых функций; «оркестровка»; «оркестровщик»; ПКС; программно-конфигурируемые сети; программно-определяемые сети; сетевая операционная система

DOI: 10.14357/08696527160108

1 Введение

В последнее время вошли в моду и активно дискутируются так называемые программно-конфигурируемые сети. Иногда их называют, формально ближе к оригинальному английскому термину, программно-определяемыми сетями. На самом деле оба названия не только не являются точными переводами, но и оба не слишком удачны, поскольку не отражают существа дела. Оригинальный термин software-defined networking правильнее переводить, в соответствии с точным значением слова networking, как программно-определяемые (задаваемые, описываемые) **создание** и **организация** сетей. Вследствие такой неадекватности находящихся в обращении русскоязычных терминов в дальнейшем тексте будет применяться оригинальный акроним SDN в его исходном значении концепции

¹Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, vegorov@ipiran.ru

создания, организации и функционирования сетей. Соответственно под сетями SDN будут пониматься сети, создаваемые на концептуальных принципах SDN, а само словосочетание «сеть SDN» не будет скрытой тавтологией.

Принципы SDN стали предметом самых разных публикаций — от серьезных научных статей до спекулятивной рекламы. Реклама и восторженные дифирамбы настойчиво создают впечатление, что SDN — чуть ли не панацея, решение едва ли не всех проблем, возникающих в современных пакетных сетях. Если верить им, то внедрение принципов SDN способно ускорить и удешевить создание сетей, упростить их масштабирование и конфигурирование, радикально повысить эффективность функционирования при уменьшении эксплуатационных затрат. Но можно ли поверить в изобретение некоего универсального средства от всех болезней?

Всеобщее непонимание даже «специалистами» сути SDN наглядно отражает определение Википедии: «...сеть передачи данных, в которой уровень управления сетью отделен от устройств передачи данных и реализуется программно, одна из форм виртуализации вычислительных ресурсов». В этом определении ошибочно все от начала до конца. Во-первых, как подчеркнуто выше, SDN — не сеть, а концепция создания и организации сетей. Во-вторых, отделение плоскости управления (control plane) от плоскости данных (data plane) давно стало общепринятой нормой разделения функций в традиционных сетях самого различного характера и назначения. Следовательно, оно не может служить отличительным признаком сетей SDN. В-третьих, плоскость управления в устройствах для традиционных сетей также повсеместно реализуется программно. Это объективно вызвано большим числом и сложностью протоколов управления (маршрутизация, резервирование полос пропускания, прокладка виртуальных трактов, распределение меток и пр.), а также необходимостью отслеживать их регулярное, зачастую радикальное, обновление и усовершенствование, что практически исключает для этой плоскости аппаратные решения. В-четвертых, сети SDN — никак не форма виртуализации вычислительных ресурсов, ибо вычислительные ресурсы вообще не являются компонентом сетей, в том числе сетей SDN, и виртуализируются иными механизмами.

После демонстрации неочевидности ряда распространенных суждений о сетях SDN далее также постараемся показать безусловность многих широко рекламируемых их преимуществ, объяснить причины нынешней популярности концепции SDN, а также оценить перспективы и условия ее практической реализации в будущем.

2 SDN под критическим углом зрения

Что же на самом деле, а не в рекламных роликах и кривом зеркале Википедии, представляют собой столь модные ныне сети SDN, и что отличает их от традиционных?

Первое. В сетях SDN плоскость управления не просто **логически отделена**, как в традиционных сетях, а **физически удалена** от плоскости данных. Кроме того, она централизована и сконцентрирована в логически единственном для всей сети SDN контроллере, хотя физически последний может быть реализован несколькими устройствами [1]. Физическое отдаление плоскостей друг от друга имеет свои положительные (например, объективно способствует обеспечению информационной безопасности в сети) и отрицательные (о них будет сказано ниже) стороны. Но в любом случае оно предполагает введение стандартизованных интерфейсов между ними, что, однако, так и не было сделано. Пока движение в этом направлении ограничилось простейшим протоколом взаимодействия контроллера сети с ее узлами OpenFlow, оставляющим без внимания многие существенные аспекты взаимодействия плоскостей, в частности и те же вопросы информационной безопасности.

Второе. Плоскость управления сетью реализуется не просто **программно**, а **прикладными программами**, работающими в контроллере сети, в качестве которого предполагается универсальный компьютер или даже целый многопроцессорный комплекс серверного класса. Именно возможность прикладного программирования функций управления на универсальном компьютере и точной их подгонки прикладными программами под особенности каждой конкретной сети лежит в основе всех потенциальных преимуществ SDN. При этом важно понимать, что все декларируемые достоинства SDN — это всего лишь «потенция» данного подхода. Каждое из этих достоинств надо реализовать соответствующим прикладным программным обеспечением (ПО), созданным для данной конкретной сети. При неудачном ПО SDN-подход не только не даст никаких преимуществ, но может привести к обратному эффекту.

Третье. Концепция SDN действительно создает предпосылки для виртуализации, причем, конечно же, не **вычислительных**, а **сетевых** ресурсов и, в частности, широко рекламируемой в контексте SDN виртуализации сетевых функций NFV (Network Functions Virtualization). Но виртуализация сети — не естественное свойство SDN, а лишь очередная потенциальная возможность, которая может быть в них реализована, лучше или хуже, дополнительным прикладным или промежуточным (middleware) ПО. Вообще говоря, NFV принципиально реализуема и в традиционных сетях. Но предполагается, что благодаря прикладному программированию в сетях SDN она эффективнее и достигается существенно проще, в том числе на уже существующей сетевой инфраструктуре [2], хотя это предположение так же не самоочевидно, как не очевидны и прочие декларируемые преимущества SDN. Не случайно в отношении другого столь же широко рекламируемого достоинства SDN-подхода — упрощения создания сетей — в англоязычном тексте Википедии очень аккуратно замечено, что такое упрощение всего лишь заявляется изобретателями подхода и продавцами соответствующей

аппаратуры («The inventors and vendors... claim that this [SDN] simplifies networking»).

Неудивительно, что из непонимания сути SDN под рекламным давлением «изобретателей подхода и продавцов соответствующей аппаратуры» рождается множество связанных с сетями SDN утверждений, которые кочуют из публикации в публикацию, но при внимательном рассмотрении оказываются, мягко говоря, не столь очевидными.

Неочевидность первая: будущее за сетями SDN, вскоре они повсеместно вытеснят традиционные сети. На самом деле сфера применения SDN-подхода может оказаться весьма ограниченной. Сеть SDN может быть любой сложности, но она имеет явную тенденцию к территориальной компактности по двум причинам:

- (1) вследствие необходимости обеспечивать прямую физическую связь по специальным доверенным и, как правило, высокоскоростным каналам между контроллером и каждым компонентом аппаратной сетевой инфраструктуры (коммутаторами, медиашлюзами и пр.). Использование для этой цели существующих каналов предполагает необходимость решения проблемы их доверенности, отнимает ресурсы у плоскости данных (тут речь, разумеется, не столько о полосах пропускания каналов, сколько о загруженности и пропускной способности устройств сетевой инфраструктуры) и снижает эффективность взаимодействия, а прокладка специальных выделенных каналов за сотни и тысячи километров сразу перечеркивает все преимущества SDN как перспективного решения для сетей масштаба MAN/WAN (metropolitan area network / wide area network);
- (2) из-за принципиальной желательности визуальной обозримости сети. Такая обозримость необходима как прикладным программистам при создании программ управления сетью, так и ее администратору, в чьи обязанности входит конфигурация сети, прокладка маршрутов через сеть потокам пакетов и корректировка их при отказах локального оборудования или реконфигурации сети. Правда, предполагается, что работа программистов и администратора может быть существенно упрощена за счет дополнительных возможностей специализированной операционной системы (ОС), а также так называемой «оркестровки» и виртуализации сетей SDN. Но ни то, ни другое, ни третье не падает с неба и требует, в свою очередь, дополнительных и весьма существенных усилий по предварительному программированию соответствующих далеко не тривиальных возможностей.

Вероятно вследствие обеих названных причин сегодня единичные реально развернутые сети SDN нашли применение только в качестве сетей центров обработки данных (ЦОД), и не видно причин ожидать массового покорения ими в будущем каких-то иных областей [3].

Неочевидность вторая: сети SDN создавать и масштабировать проще, чем традиционные. В реальности все может оказаться равным счетом наоборот. Для создания (масштабирования) традиционной сети достаточно купить (докупить) необходимое оборудование и соединить все узлы и компоненты сети подходящими каналами связи, после чего сеть сразу может начать функционировать. Пусть далеко не оптимально, но сразу. В дальнейшем при желании работу сети можно оптимизировать индивидуальной настройкой ее узлов, что, правда, трудоемко и не всегда обеспечивает желаемый результат. Но чтобы создать (промасштабировать) и заставить работать сеть SDN, мало купить (докупить) оборудование и проложить каналы связи. Надо еще написать (дописать или переписать) специфическое для данной сети прикладное ПО, управляющее ее работой, и, самое неприятное, неким до сих пор никем не объясненным способом отладить и верифицировать его в реальной сети. Вот теперь, если проблема создания и отладки ПО каким-то чудесным образом решена, можно его тьюнингом добиваться оптимального функционирования сети. Трудно судить априори, будет ли процесс глобального тьюнинга менее трудоемким, чем сепаратная настройка отдельных узлов традиционной сети, но и тут адепты SDN полагаются на свои универсальные палочки-выручалочки — специализированную ОС, всезнайку-«оркестровщика» и виртуализацию [2].

Неочевидность третья: внедрение SDN способно заметно сократить капитальные вложения при создании сетей за счет применения универсальных массово выпускаемых и потому относительно дешевых серверов вместо специализированного и, как правило, существенно более дорогого сетевого оборудования. Но при этом, как и в предыдущем случае, вновь почему-то замалчивается необходимость приобретения для этих «относительно дешевых серверов» специализированного ПО (сетевая ОС, «оркестровщик», инструментальные средства), а также написания и отладки «по месту» специального ПО (промежуточное ПО, прикладные программы управления сетью). А ведь программы давно стали на порядки более дорогим компонентом конечного продукта в сравнении с аппаратурой, на которой они работают. Между тем «существенно более дорогое сетевое оборудование» поставляется изготовителями с уже разработанным и отлаженным внутренним ПО, и оно немедленно готово к работе.

Неочевидность четвертая: оборудование для сетей SDN проще, чем для традиционных. Это утверждение отчасти справедливо в том смысле, что концептуально внутри сетей SDN вообще нет нужды в традиционных маршрутизаторах благодаря тому, что все задачи плоскости управления, включая маршрутизацию и сопряженные с нею функции, выполняет контроллер сети. (Хотя, следует заметить, граничные маршрутизаторы на периферии сети SDN остаются в своем традиционном назначении.) Зато коммутаторы в сетях SDN сложнее, чем в традиционных, так как обязаны оснащаться дополнительным специальным портом для связи с контроллером сети и обес-

печивать на нем поддержку протокола OpenFlow, который, на что также следует обратить внимание, быстро развивается, и регулярно появляющиеся версии будут быстро морально старить предназначенную для сетей SDN аппаратуру. При этом, разумеется, с коммутаторов в сетях SDN никто не снимает все традиционные функции плоскости данных: продвижение пакетов, управление трафиком, обеспечение качества обслуживания, гарантии сетевой безопасности и т. д.

Неочевидность пятая: унификация взаимодействия между плоскостями управления и данных, в частности протокол OpenFlow, устраняет зависимость создателей сетей от поставщиков сетевого оборудования. Как правило, из-за несовместимости коммутаторов различных поставщиков по набору и качеству реализуемых функций физически или морально устаревшее сетевое оборудование можно было заменять на новое только от того же самого производителя. Теоретически унификация протокола (но, отметим еще раз, не интерфейса!) между контроллером сети и компонентами сетевой инфраструктуры устраняет эту зависимость. Но, даже оставляя в стороне проблему унификации интерфейса, которая вообще пока никем не решается, кто на практике сможет гарантировать, что поголовно все поставщики реализуют в своем оборудовании все стандартизованные возможности того же OpenFlow и, главное, реализуют их абсолютно одинаковым образом? Если для простейших функций (устройств *type0* согласно стандарту OpenFlow) еще можно надеяться на относительную идентичность их исполнения, то для более сложных (устройств *type1* и выше), которые еще даже не специфицированы, говорить о полной совместимости и взаимозаменяемости по крайней мере преждевременно.

Из сказанного следует очевидный вывод: концепция SDN сама по себе — никакая не панацея, она отнюдь не решает всех проблем сетей пакетной коммутации и не лишена ряда причин для сомнений. Ниже перечислены наиболее очевидные из них:

- ограниченная применимость SDN-подхода только к территориально компактным сетям, особенно эксплуатирующим NFV, — на практике в основном к сетям ЦОД;
- необходимость индивидуального прикладного программирования каждой сети при ее создании и перепрограммирование при любом усовершенствовании и масштабировании с отладкой «в поле» сложных комплексов прикладного ПО;
- добавление в каждом устройстве порта для организации выделенного доверенного канала связи с контроллером сети, а также поддержка на нем дополнительного специального протокола, например OpenFlow;
- замедленная коммуникация между плоскостями данных и управления по каналу связи вместо простой передачи пакета для обработки в другую плоскость непосредственно внутри (в памяти) традиционных сетевых устройств;

- сложность обеспечения высокой надежности и доступности сети из-за централизации всех функций управления в одном контроллере;
- объективные трудности сопряжения с существующими традиционными сетями.

Почему же несмотря на внушительный перечень поводов для сомнений концепция SDN остается одним из ведущих трендов современности?

3 Причины популярности SDN

Не исключено, что главной движущей силой, буквально «проталкивающей» концепцию SDN в пакетные сети, являются не столько ее декларируемые преимущества, сколько те самые продавцы соответствующей аппаратуры, которые их декларируют. В частности, именно с SDN и NFV связывает надежды выхода на рынок сетевого оборудования такой тяжеловес, как компания Intel. До сих пор все попытки этого гранда микроэлектроники внедрить в сетевые устройства свои универсальные процессоры базовой архитектуры x86 регулярно терпели неудачу в конкуренции с более компактными и мало потребляющими процессорами архитектур ARM и MIPS [4]. Сегодня именно эти две архитектуры с примерно равными возможностями [5] продолжают между собой борьбу на быстро растущих рынках сетевых и мобильных устройств. Концепция SDN/NFV открывает дорогу в пакетные сети универсальным серверам на основе процессоров с фирменной маркировкой «Intel inside» в качестве контроллеров сети и платформ виртуализации. К усилиям Intel присоединяются крупные поставщики серверного ПО. В качестве примера можно привести компанию Juniper Networks, предлагающую готовые программные решения для сетей SDN, или VMware, специализирующуюся на ПО для ЦОД и «облачной» инфраструктуры. Но и старые игроки на рынке сетевого оборудования, такие как Cisco Systems или Alcatel-Lucent, не намерены терять свой кусок пирога и тоже активно рекламируют свои частные решения для инфраструктуры SDN и NFV.

В России уровень интереса к SDN должен не уступать мировому и даже, возможно, превосходить его вследствие нашего хронического технологического отставания. Ведь SDN внешне выглядит магистральным путем к замене сетевой аппаратуры, которую мы делать не умеем, прикладными программами, которые вроде бы вполне в состоянии написать. Для российских разработчиков и системных интеграторов открывается весьма заманчивая перспектива сдать в металлолом маршрутизаторы от Cisco и Alcatel, надежды на разработку и выпуск в нашей стране аналогов которых давно потеряны, а все пакетные сети создавать на основе виртуальных коммутаторов и серийных серверов, благо недавно в последних даже стали появляться процессоры отечественной разработки [6].

Однако на этом заманчивом пути у нас все равно остается необходимость понимания того, насколько серьезны перечисленные выше сомнительные аспекты

SDN, не окажутся ли они настолько неодолимым препятствием, что похоронят саму столь притягательную, особенно в российских условиях, идею. Имеющийся ничтожный опыт буквально единичных сетей, которые объявляются (т. е. вновь всего лишь декларируются!) построенными на принципах SDN, ответа на этот принципиальный вопрос не дают. Так что пока можно говорить только об ожиданиях и пытаться оценить перспективы и риски.

Что касается применимости SDN-подхода только к территориально компактным сетям, то это ограничение представляется фундаментальным. И теоретические рассуждения, и опыт единичных реализаций не дают оснований распространять область реализуемости SDN за рамки ЦОД. Может быть, частично и ограниченно принципы SDN могут найти применение в некоторых провайдерских сетях. При этом степень частичности и ограниченности может оказаться любой из-за расплывчатости самих принципов SDN и допустимости любых сочетаний декларативности и реальности. Попутно можно отметить, что требованию территориальной компактности естественным образом соответствуют локальные сети, в том числе виртуальные, и они вполне реализуемы на принципах SDN. Но вряд ли такая реализация окажется экономически выгодной с учетом массового выпуска специализированного дешевого оборудования для LAN.

Все прочие перечисленные выше трудности SDN представляются преодолимыми, но остается без ответа вопрос цены их преодоления. Чтобы потенциальные преимущества SDN-подхода дали реальную выгоду в простоте создания и качестве управления работой конкретной сети, для этой сети необходимо создать уникальное адекватное прикладное ПО. В России, где достаточно много квалифицированных программистов, а их зарплаты относительно невелики, какие-то сети действительно могут оказаться экономически более выгодными, а функционально более эффективными, будучи созданы на принципах SDN прямым программированием «в лоб». Но если говорить более общо, то для обретения обещанного рекламой выигрыша в затратах на развертывание сетей SDN, упрощения их масштабирования и высокой эффективности управления надо обеспечить их создателей адекватным инструментарием и надежной стартовой базой. В качестве последней видятся эффективные специализированные ОС и готовые базовые (эталонные) решения по контроллерам сетей SDN.

Говоря об ОС для сети SDN, часто используют термин «сетевая ОС», что способно породить путаницу. Традиционно под этим термином понимается ОС с дополнительными свойствами, позволяющими компьютеру работать в сети. Эти дополнительные свойства включают, например, поддержку сетевых протоколов, в первую очередь стека TCP/IP. Однако контроллер сети SDN **должен управлять сетью, но не обязан работать в сетях**, и, соответственно, для SDN-ОС вовсе не принципиальна поддержка сетевых протоколов. Для управления же сетью ей нужен один единственный протокол — OpenFlow или ему эквивалентный [7]. Таким образом, SDN-ОС — это не просто сетевая ОС, а весьма специфический продукт, который призван служить средством сопряжения прикладного и промежуточного ПО контроллера с сетевой инфраструктурой.

В этом смысле иногда понятия SDN-ОС и SDN-контроллер рассматриваются как синонимы.

4 Заключение

Какие-то экспериментальные сети SDN уже существуют, по крайней мере на уровне деклараций. Также анонсируются и некие SDN-ОС и SDN-контроллеры — как экспериментальные вроде Onix [8], так и коммерческие: NOX/POX, BigSwitch, FloodLight и др. Особенно приятно, что этот перечень пополнился отечественным SDN-контроллером RUNOS (Russian network operating system) от Центра прикладных исследований компьютерных сетей [9]. Менее понятна ситуация с «оркестровщиками» и инструментарием для создания рабочих мест администраторов сети и прочего прикладного ПО. Проблема в том, что, несмотря на интенсивные теоретические исследования [1, 10, 11], на сегодняшний день «отсутствует еще программный слой, посредством которого требования бизнес-логики транслируются на уровень управления инфраструктурой» [12]. Пока нечто напоминающее такой «программный слой» в качестве коммерческого продукта предлагают только компании Juniper Networks (универсальное ПО для SDN/NFV Contrail) и BATM Advanced Communications («оркестровщик» EdgeGenie). Также уместно вспомнить об отечественном SDN-ОС-контроллере RUNOS, хотя его статус как коммерческого продукта пока не ясен. Между тем именно на этом поле, поле многофункциональной эффективной SDN-ОС, общедоступного интеллигентного «оркестровщика» и практичного промежуточного ПО виртуализации, похоже, будет решаться судьба и определяться будущее сетей SDN.

Литература

1. *Tootoonchian A., Ganjali Y.* HyperFlow: A distributed control plane for OpenFlow // 2010 Internet Network Management Workshop/Workshop on Research on Enterprise Networking (INM/WREN'10) Proceedings. San Jose, CA, USA. 6 p. https://www.usenix.org/legacy/events/inmwren10/tech/full_papers/Tootoonchian.pdf.
2. *Варламов К.* Обработка трафика в облаке. Кому нужна виртуализация сетевых функций? // Хабрахабр, 07.09.2015. <http://habrahabr.ru/post/266343>.
3. *Брокс У., Вайнгарт Й.* Оптимизация ЦОД для облака // Журнал сетевых решений/LAN, 2013. № 12. С. 50–52.
4. *Dilger D. E.* How Intel lost the mobile chip business to Apple's A8 Application Processors // Appleinsider, 19.01.2015. <http://appleinsider.com/articles/15/01/19/how-intel-lost-the-mobile-chip-business-to-apples-ax-arm-application-processors>.
5. *Егоров В. Б.* Обзор и сравнение процессорных ядер ARM Cortex и MIPS Aptive // Электронные компоненты, 2013. № 6. С. 64–69.
6. Начались продажи PC на базе процессоров «Эльбрус-4С» // Вести: Экономика, 06.05.2015. <http://www.vestifinance.ru/articles/56946>.

7. McKeown N., Anderson T., Balakrishnan H., Parulkar G., Peterson L., Rexford J., Shenker S., Turner J. Openflow: Enabling innovation in campus networks // SIGCOMM Computer Communication Rev., 2008. Vol. 38. No. 2. P. 69–74. <http://archive.openflow.org/documents/openflow-wp-latest.pdf>.
8. Koponen T., Casado M., Gude N., Stribling J., Poutievski L., Zhu M., Ramanathan R., Iwata Y., Inoue H., Hama T., Shenker S. Onix: A distributed control platform for large-scale production networks // 9th USENIX Symposium on Operating Systems Design and Implementation (OSDI 2010) Proceedings. — USENIX Association, 2010. P. 351–364. <http://yuba.stanford.edu/~casado/onix-osdi.pdf>.
9. Первый российский ПКС/SDN-контроллер RUNOS. Центр прикладных исследований компьютерных сетей, 2015. <http://arccn.ru/research/653>.
10. Sherwood R., Gibby G., Yary K.-K., Appenzellery G., Casado M., McKeown N., Parulkary G. FlowVisor: A network virtualization layer // OpenFlow Technical Reports, 2009. No. 2009-1. <http://archive.openflow.org/downloads/technicalreports/openflow-tr-2009-1-flowvisor.pdf>.
11. Yeganeh S. H., Ganjali Y. Kandoo: A framework for efficient and scalable offloading of control applications // 1st Workshop on Hot Topics in Software Defined Networks (HotSDN'12) Proceedings. — ACM, 2012. P. 19–24. <http://conferences.sigcomm.org/sigcomm/2012/paper/hotsdn/p19.pdf>.
12. Дубова Н. Программный дуализм // Открытые системы, 2014. № 6. С. 14–16. <http://www.osp.ru/os/2014/06/13042309/>.

Поступила в редакцию 26.01.16

SOME ISSUES OF THE SDN CONCEPT PRACTICAL IMPLEMENTATION

V. B. Egorov

Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilova Str., Moscow 119333, Russian Federation

Abstract: The software-defined networking (SDN) concept, widely discussed lately, is subject to versatile consideration. Regarding the SDN, both unobviousness of several prevalent judgments and conventionality of many widely advertised advantages (e. g., deployment speed-up, operating costs reduction, and control quality improvement) are exposed. The analysis reveals such peculiarities of the SDN approach as inclination to spatial limitation, burdens of individual applied programming, desirability of introducing special trusted control channels, slowed down communication between the control and data planes, problems with ensuring reliability and availability, and objective complexities of connection with existing traditional networks. An attempt is made to explain the recent SDN concept popularity motives with commercial interests

of some heavyweights on the microelectronics and networking software markets. Perspectives and conditions of the SDN concept practical implementation in the future are shown.

Keywords: network functions virtualization; network operating system; NFV; orchestration; orchestrator; RUNOS; SDN; software-defined networking

DOI: 10.14357/08696527160108

References

1. Tootoonchian, A., and Y. Ganjali. 2010. HyperFlow: A distributed control plane for OpenFlow. *INM/WREN Proceedings*. San Jose, CA. 6 p. Available at: <http://www.cse.iitd.ac.in/siy107537/csl374/a5/files/Tootoonchian.pdf> (accessed February 9, 2016).
2. Varlamov, K. 2015. Obrabotka trafika v oblake. Komu nuzhna virtualizatsiya setevykh funktsiy? [Traffic processing in a cloud. Who needs the network functions virtualization?]. Available at: <http://habrahabr.ru/post/266343/> (accessed February 9, 2016).
3. Broks, U., and J. Weingart. 2013. Optimizatsiya TsOD dlya oblaka [Cloud data-center optimization]. *Zhurnal Setevykh Resheniy / LAN* [Magazine of Network Solutions / LAN] 12:50–52. Available at: <http://www.osp.ru/lan/2013/12/13039006/> (accessed February 9, 2016).
4. Dilger, D. E. 19.01.2015. How Intel lost the mobile chip business to Apple's Ax Application Processors. *Appleinsider* Available at: <http://appleinsider.com/articles/15/01/19/how-intel-lost-the-mobile-chip-business-to-apples-ax-arm-application-processors> (accessed February 9, 2016).
5. Egorov, V. B. 2013. Obzor i sravnenie protsessornykh yader ARM Cortex i MIPS Aptive [A review and comparison of the processor cores ARM Cortex and MIPS Aptive]. *Elektronnye Komponenty* [Electronic Components] 6:64–69.
6. Nachalis' prodazhi PC na baze protsessorov "El'brus-4C" [PCs on the base of the El'brus-4C processor are for sale]. 06.05.2015. *Vesti Ekonomika* [News Economics], May. Available at: <http://www.vestifinance.ru/articles/56946> (accessed February 9, 2016).
7. McKeown, N., T. Anderson, H. Balakrishnan, G. Parulkar, L. Peterson, J. Rexford, S. Shenker, and J. Turner. 2008. Openflow: Enabling innovation in campus networks. *SIGCOMM Computer Communication Rev.* 38(2):69–74. Available at: <http://archive.openflow.org/documents/openflow-wp-latest.pdf> (accessed February 9, 2016).
8. Koponen, T., M. Casado, N. Gude, J. Stribling, L. Poutievski, M. Zhu, R. Ramanathan, Y. Iwata, H. Inoue, T. Hama, and S. Shenker. 2010. Onix: A distributed control platform for large-scale production networks. *OSDI Proceedings*. USENIX Association. 351–364. Available at: <http://yuba.stanford.edu/~casado/onix-osdi.pdf> (accessed February 9, 2016).
9. Pervyy rossiyskiy PKS / SDN-kontroller RUNOS [The first Russian SDN controller RUNOS]. 2015. Tsentri prikladnykh issledovaniy komp'yuternykh setey [Computer Networks Applied Researches Center]. Available at: <http://arccn.ru/research/653> (accessed February 9, 2016).

10. Sherwood, R., G. Gibby, K.-K. Yapy, G. Appenzellery, M. Casado, N. McKeowny, and G. Parulkary. 2009. FlowVisor: A network virtualization layer. *Openflow Technical Reports 2009-1*. Available at: <http://archive.openflow.org/downloads/technicalreports/openflow-tr-2009-1-flowvisor.pdf> (accessed February 9, 2016).
11. Yeganeh, S. H., and Y. Ganjali. 2013. Kandoo: A framework for efficient and scalable offloading of control applications. *1st Workshop on Hot Topics in Software Defined Networks (HotSDN'12) Proceedings*. ACM. 19–24. Available at: <http://conferences.sigcomm.org/sigcomm/2012/paper/hotsdn/p19.pdf> (accessed February 9, 2016).
12. Dubova, N. 2014. Programmnyy dualizm [The software dualism]. *Otkrytye Sistemy [Open Systems]* 6:14–16. Available at: <http://www.osp.ru/os/2014/06/13042309/> (accessed February 9, 2016).

Received January 26, 2016

Contributor

Egorov Vladimir B. (b. 1948)—Candidate of Science (PhD) in technology, leading scientist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119133, Russian Federation; VEgorov@ipiran.ru

ИНФОРМАЦИОННАЯ БЕЗОПАСНОСТЬ СИТУАЦИОННЫХ ЦЕНТРОВ*

А. А. Зацаринный¹, В. И. Королёв²

Аннотация: Рассмотрено назначение и характеристика ситуационных центров (СЦ) как информационно-технологических объектов, обеспечивающих реализацию ситуационного управления. Ситуационный центр представляется системой. Отмечается сложный состав и структура информационных ресурсов, обеспечивающих принятие решений. В качестве фактора проблемы информационной безопасности (ИБ) СЦ принимается различный уровень конфиденциальности обрабатываемой информации. Информационная безопасность представляется как защита информации. Предлагается подход построения системы защиты информации (СЗИ) для СЦ на основании архитектурного подхода. При этом архитектура любой управляемой организации представляется как совокупность: бизнес-архитектура + архитектура информационных технологий (ИТ), а понятие архитектуры системы соответствует стандарту ГОСТ Р ИСО/МЭК 15288-2008. Определены актуальные задачи ИБ СЦ.

Ключевые слова: ситуационный центр; ситуационное управление; информационные ресурсы; информационно-коммуникационные технологии; конфиденциальность; информационная безопасность; система защиты информации; архитектурный подход

DOI: 10.14357/08696527160109

1 Назначение и характеристика ситуационных центров

Ситуационный центр представляется как комплекс, который концентрирует информацию о подконтрольном пространстве от различных источников и обеспечивает ситуационное управление, принятие управленческих решений с широким использованием информационно-коммуникационных технологий (ИКТ), моделей и методов ситуационного анализа.

«Ситуационность» означает, что действия субъектов, реализация процессов и действий определяются контекстом, в котором они осуществляются. Это понятие лежит в основе ситуационной теории управления, изучающей зависимость

* Работа выполнена при поддержке РФФИ (проект 15-29-07981 офи-м).

¹Федеральный исследовательский центр «Информатика и управление» Российской академии наук (ФИЦ ИУ РАН); Институт проблем информатики ФИЦ ИУ РАН, AZatsarinny@ipiran.ru

²Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук; факультет кибернетики и информационной безопасности Научно-исследовательского ядерного университета МИФИ, vkorolev@ipiran.ru

эффективности методов управления от того, в каком положении находится объект управления [1].

Ситуационное управление как система становится эффективным, когда необходимо, чтобы «управляющие воздействия субъекта управления основывались на анализе вариантов принятия решения с учетом текущего состояния объекта управления, располагаемых вариантов действий и прогноза последствий принимаемых управленческих воздействий» [2].

Такая необходимость практически постоянно присутствует, когда объектами управления являются большие и сложные объекты, по своей природе являющиеся социотехническими системами. Ими могут быть управленческие, предпринимательские, производственные, финансово-операционные и прочие образования (организации, предприятия, фирмы, финансово-кредитные учреждения и т. д.), а также муниципальные, региональные и федеральные образования с соответствующими государственными органами, обеспечивающими жизнедеятельность и управление.

Объекты управления считаются большими, если они имеют распределенную организационную структуру, которая включает в себя компоненты многочисленных видовых групп деятельности с достаточно представительным множеством территориально распределенных однородных с точки зрения вида деятельности структурных подразделений.

Объекты управления считаются сложными, если для реализации их целевых функций требуется создание и функционирование сложного и многопланового инструментария: моделей, процессов и технологий; организационных, технико-технологических и инженерных инфраструктур; нормативно-правового и планово-финансового обеспечения; людских ресурсов с широкими группами различной специализации. Живучесть таких объектов в современных условиях зависит от конкурентной борьбы и различного вида противостояний и поддерживается средствами с широким использованием ИТ — признаком и сущностью информационного общества.

Именно для таких больших и сложных системных объектов существует большая вероятность возникновения под влиянием внутреннего и внешнего воздействия совокупности обстоятельств, которые нарушают заданное функционирование и требуют его перевода в новое состояние, чтобы исключить нарушение штатного режима или обеспечить наибольшую эффективность функционирования. А следовательно, становится необходимым и эффективным ситуационное управление.

Если исходить из этого общего представления, СЦ должен иметь комплексную функциональную направленность обеспечения процессов управления и поддержки принятия решений и включать в себя следующие компоненты информационно-телекоммуникационной системы [3]:

- комплекс аппаратно-программных средств (средства отображения информации, средства хранения информации, средства телекоммуникационные,

средства защиты информации, автоматизированные рабочие места, средства жизнеобеспечения, средства общего и общесистемного программного обеспечения (ПО));

- комплексы специального ПО, реализующие выполнение функциональных задач по назначению СЦ;
- обслуживающий персонал, специально подготовленный для эффективного применения и поддержания функционирования СЦ в заданных режимах с соблюдением мер по ИБ.

Организационно СЦ состоит из четырех сегментов [3, 4]:

- (1) руководства органа управления;
- (2) мониторинга состояния контролируемого информационного пространства;
- (3) ситуационного анализа и поддержки процесса принятия управленческих решений, включая ситуационный зал (или кабинет, комнату);
- (4) администрирования информационного обеспечения, аппаратно-программных средств и средств защиты информации.

Таким образом, СЦ — это базирующийся на организационно-техническом комплексе информационно-технологический объект в системе управления, обеспечивающий реализацию необходимых ИКТ, подготовку и предложение информационно-аналитических решений и поддержку процессов принятия решений на компетентном уровне (лица принятия решений — ЛПР) в режимах ситуационного управления. На его основе должны быть обеспечены мониторинг факторов влияния на развитие происходящих процессов, информационно-аналитическая поддержка процедур и процессов, позволяющих оперативно анализировать, моделировать, прогнозировать сценарии развития ситуаций и динамично вырабатывать эффективные решения.

2 Фактор информационной безопасности

Одним из определяющих факторов выбора регламентов и обеспечения управления функционированием СЦ, организации работы каждого из перечисленных компонентов, особенностей выбора технических и инженерных решений при создании и эксплуатации СЦ является характеристика конфиденциальности используемых информационных ресурсов. Как правило, собственные накапливаемые информационные ресурсы СЦ и привлекаемые внешние информационные ресурсы для обеспечения ситуационного управления и принятия решений независимо от вида объекта управления содержат в той или иной степени информацию, требующую контролируемого доступа. В соответствии с законодательством Российской Федерации это могут быть персональные данные, сведения, относящиеся к определенным видам тайн (государственная, коммерческая, банковская, служебная и др.). И даже в общем случае открытые данные, но относящиеся

к планируемым инновационным и инвестиционным проектам, к авторским правам, к оперативным текущим характеристикам, часто в определенной конкретной ситуации становятся служебной информацией и также требуют контролируемого доступа. Большую роль играет при рассмотрении тех или иных ситуаций проблема безопасного агрегирования различной информации из собственных информационных ресурсов и внешних источников.

Исходя из этого, в современных условиях при использовании для принятия решений больших объемов информации, различной по уровню конфиденциальности и признаку владения и распоряжения, к СЦ предъявляются повышенные требования по обеспечению ИБ. При этом под обеспечением ИБ в данной работе понимается решение системной задачи защиты информации от несанкционированного доступа к ней, несанкционированной утечки информации или ее модификации (нарушения целостности) при обработке и хранении в информационно-аналитической системе (ИАС) СЦ, передаче в телекоммуникационные системы (ТКС), при реализации технологий ситуационного управления и принятия решений. Хотя это только часть сложной и многоаспектной проблемы обеспечения ИБ автоматизированных информационных систем на объектах информатизации [5].

Следует отметить, что на сегодняшний день такой подход соответствует действующим нормативным и руководящим документам, в которых ИБ при использовании ИТ определяется как «защита конфиденциальности, целостности и доступности информации» [6].

Поскольку предметом защиты является информация, а СЦ выступает объектом реализации определенных ИТ, то системная задача защиты информации решается на всем технологическом пространстве СЦ, в его компонентах, при реализации самих процессов управления, в которых участвует СЦ. В свою очередь, СЦ является материально-техническим объектом, и в этой части на него распространяются положения ГОСТ Р 51275-99 [7], в котором в состав защищаемого объекта информатизации включаются средства жизнеобеспечения (здания, сооружения, помещения, обеспечивающие технические средства и инженерные системы), которые необходимы для установки и эксплуатации средств и систем обработки информации. Особо выделяются «помещения и объекты, предназначенные для ведения конфиденциальных переговоров».

С учетом данных предпосылок СЦ формируется как объект защиты, и одним из его системных компонентов должен быть компонент обеспечения защиты информации.

3 Общее представление ситуационного центра как системы

Ситуационный центр является сложной автоматизированной ИАС (АИАС). Необходимо исходить из того, что СЦ является частью системы управления объекта (организации как объекта), дополняет ее, представляя собой, по существу, информационно-аналитическую подсистему или подсистему поддержки приня-

тия решений в составе системы управления объекта. При декомпозиции СЦ необходимо учитывать системную организацию и функциональное наполнение управления объектом как надсистемы для СЦ, а также учитывать это в части взаимосвязи всех структурных компонентов СЦ с целью обеспечения эффективного выполнения задач СЦ.

Общая структурно-функциональная схема СЦ как системы представляется в виде взаимосвязанной совокупности трех базовых составляющих: информационно-аналитической, информационно-технологической и технической [4, 8].

Информационно-аналитическая составляющая включает в себя совокупность методов, методик, моделей, типовых ситуаций, возможных реакций на угрозы, сценариев принятия решений, ориентированных на решение задач ситуационного анализа в рамках заданной предметной области СЦ. Практически это сервисные средства ситуационного управления, которые должны обеспечивать деятельность экспертно-аналитического и оперативного состава органов управления (экспертно-аналитическая и дежурная службы), отвечающих за разработку и выдачу информационно-аналитических продуктов для ЛПР и за подготовку и сопровождение процессов ситуационного управления.

Инструментально информационно-аналитическая составляющая представляет собой приложения прикладного ПО и регламенты работы с ним, предназначенные для решения функциональных задач СЦ. Это средства ситуационного анализа, моделирования ситуаций, системы задания, распознавания и извлечения из информационных ресурсов необходимой информации, формирования информационных продуктов визуализации и решения информационно-расчетных задач, выбора и настройки сценариев и регламентов технологического процесса ситуационного управления.

Информационно-технологическая составляющая — это исполнительная информационно-технологическая платформа СЦ. Она должна обеспечить необходимыми информационными ресурсами и реализовывать базовые технологии (технологические сервисы, процедуры) функционирования СЦ как части системы управления объекта, использовать в соответствующих сценариях ситуационного управления информационно-аналитические продукты, настройки и регламенты, подготавливаемые для этих сценариев.

Одной из базовых функций информационно-технологической составляющей является формирование и ведение информационных ресурсов СЦ. Это, прежде всего, сбор из различных источников данных (формализованной и неформализованной информации), обработка, анализ, обобщение и отображение данных в базах данных применительно к циклу «событие–ситуация–угроза», интеграция и предоставление данных для решения функциональных задач, а также управление внутренними и внешними информационными потоками. Ведение информационных ресурсов СЦ требует ведения единого лингвистического обеспечения.

Инструментально информационно-технологическая составляющая содержит комплексы специального ПО, а также общее ПО, выбор которого обуславлива-

ется спецификой функционирования СЦ и требованиями по защите информации. Очевидно, что эта составляющая определяет целевое предназначение и эффективность системы СЦ в целом.

Между информационно-аналитической и информационно-технологической составляющими существует непрерывное технологическое и информационное взаимодействие, обмен ресурсами, и в этой части существует постоянная потребность уточнения их корреляции, которая, как правило, определяется и осуществляется в рамках ИАС СЦ.

Техническая составляющая является аппаратно-программной платформой СЦ, обеспечивающей вычислительную и телекоммуникационную технологическую среду для реализации задач СЦ и его функционирования в целом.

Инструментально техническая составляющая включает в себя вычислительный комплекс ИАС, ТКС с оборудованием визуализации информации и видеоконференцсвязи, оборудование информационного взаимодействия с внешними ИАС и базами данных, в том числе в режиме облачного взаимодействия, их общесистемное ПО и общесистемные программные сервисы. К общесистемным программным сервисам следует отнести и сервисы управления компонентами и процессами аппаратно-программной платформы СЦ.

Согласно ГОСТ Р 51275-99 [7] в состав технической составляющей СЦ включаются средства жизнеобеспечения (здания, сооружения, помещения, обеспечивающие технические средства и инженерные системы), которые необходимы для установки и эксплуатации средств и систем аппаратно-программной платформы СЦ, в том числе помещения, предназначенные для ведения конфиденциальных переговоров при функционировании СЦ.

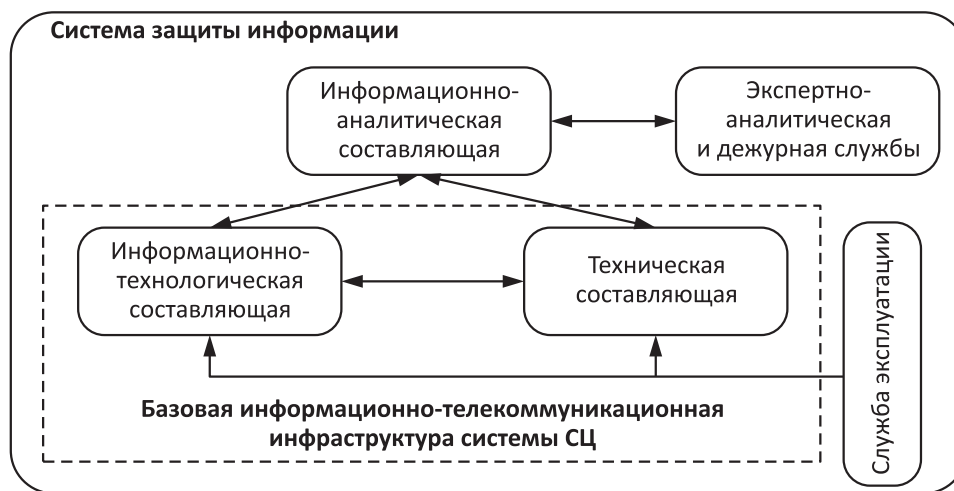


Рис. 1 Обобщенная структурно-функциональная схема системы СЦ

Информационно-технологическая и техническая составляющие формируют базовую информационно-телекоммуникационную инфраструктуру СЦ. Служба эксплуатации обеспечивает поддержание средств базовой информационно-телекоммуникационной инфраструктуры в требуемом состоянии и их штатное применение.

Информационная безопасность СЦ обеспечивается СЗИ, которая комплексно решает задачи защиты для СЦ в целом.

Обобщенная структурно-функциональная схема системы СЦ представлена на рис. 1.

4 Архитектурный подход к проектированию ситуационного центра и обеспечению информационной безопасности

Будучи частью системы управления организации в целом, СЦ представляет собой информационно-технологический объект. Его проектирование, с одной стороны, тесно связано с бизнес-процессами, деловыми и служебными процессами деятельности организации, с другой стороны, СЦ проектируется как системный компонент ИТ-архитектуры организации. Такая схема построения типична в современных условиях глубокого проникновения ИКТ в бизнес-процессы и управление и требует архитектурного подхода при проектировании, когда под архитектурой организации понимается совокупность: бизнес-архитектура + архитектура ИТ [9]. При этом будем придерживаться понятия, соответствующего стандарту ГОСТ Р ИСО/МЭК 15288-2008: архитектура системы — принципиальная организация системы, воплощенная в ее элементах, их взаимоотношениях друг с другом и со средой, а также принципы, направляющие ее проектирование и эволюцию.

В то же время следует учитывать и свойство рекурсивности в понятии архитектуры системы, часто используемое в области создания сложных программных продуктов: архитектура системы состоит из нескольких компонентов, внешних свойств и интерфейсов, связей и накладываемых ограничений, а также архитектуры этих внутренних компонентов.

Архитектурный подход позволяет проектировать обеспечение ИБ СЦ как частную архитектуру системной архитектуры организации, и ее построение должно соответствовать или, по крайней мере, не противоречить архитектурным, структурным и техническим решениям по автоматизации и информатизации организации [10].

5 Система защиты информации ситуационного центра

Реализация мер организационного, технологического и технического характера по защите информации в СЦ возлагается на СЗИ, которая при проектировании СЦ как системы является ее *обеспечивающей подсистемой* [4].

В рамках архитектурного подхода к созданию СЦ базовым положением проектирования СЗИ выступает следующее требование: структурные, технические, технологические и организационные решения по защите информации должны быть неразрывно связаны с проектными решениями по СЦ и обеспечивающей его функционирование инфраструктуре. Это требование служит предпосылкой для системной реализации ИБ и комплексного использования всех известных и соответствующих нормативам методов и средств защиты информации, т. е. создания комплексной СЗИ СЦ, которая осуществляет системную интеграцию разнородных задач и средств для достижения единой цели — обеспечения заданного уровня защиты.

Средства защиты, входящие в СЗИ СЦ, должны регулярно выполнять свои функции, не конфликтуя между собой и с другими компонентами, обеспечивающими функционирование СЦ. Они могут быть реализованы различными способами и методами в части моделей, алгоритмов и инструментария, могут быть программными или аппаратными средствами, организационными мерами и мероприятиями, а также представлять собой комплексные технические и организационные решения в виде отдельных функциональных изделий.

Системное решение обеспечения ИБ в виде СЗИ предполагает при функционировании СЦ обеспечение защиты информации [11]:

- при всех видах информационной деятельности на объекте управления, в которых используется СЦ;
- во всех структурных компонентах СЦ;
- при всех режимах функционирования СЦ;
- на всех этапах жизненного цикла объекта управления и, соответственно, СЦ;
- с учетом информационного взаимодействия с внешней средой.

В общем случае СЗИ СЦ разрабатывается как подсистема автоматизированной системы, так как ранее было отмечено, что СЦ — это сложная АИАС. Типовые подходы к проектированию систем защиты подобного назначения в настоящее время практически сложились [5], однако имеют место свои особенности и новые проблемы по ИБ, которые присущи СЦ как информационно-технологическому и функциональному компоненту в системе управления.

По системному представлению СЗИ СЦ является автоматизированной организационно-технической системой управления ИБ и включает в себя функциональную и обеспечивающую части.

Функциональная часть СЗИ СЦ — это совокупность функциональных задач по нейтрализации угроз информации, которые могут иметь различные источники и быть различной природы. *Функциональные задачи* ставятся и решаются в отношении актуальных для СЦ угроз. При этом обеспечивается:

- рассмотрение всех угроз информации, влияющих на состояние защиты информации в СЦ;

- выделение из них актуальных путем анализа порождаемых ими рисков и последующего возможного ущерба, если риск становится реальным событием;
- ранжирование актуальных угроз;
- введение в классификационную схему угроз информации всех выделенных актуальных угроз, учитываемых при создании СЗИ СЦ.

Функциональные задачи защиты образуют *функциональные подсистемы СЗИ СЦ*, объединяющие задачи по признакам их решения в определенных ИТ, компонентах СЦ или по принадлежности группы функциональных задач к выделенному направлению защиты информации.

Примерный состав функциональных подсистем СЦ, который, безусловно, должен быть конкретизирован и обоснован для каждого проектируемого СЦ, может быть следующим, а именно включать подсистемы:

- защиты информации в ТКС;
- защиты от несанкционированного доступа (НСД) к информации и ресурсам в ИАС, в том числе управления разграничением доступа к информации, ресурсам, режимам функционирования ИАС и режимам ситуационного управления;
- обеспечения ИБ операционного ситуационного помещения (зала/комнаты/кабинета), в том числе защиты информации визуализации и видеоконференцсвязи;
- контроля и разграничения доступа при внешнем и удаленном информационном взаимодействии, в том числе в регламенте облачного взаимодействия с внешними базами данных;
- секретного и конфиденциального электронного документооборота (ЭДО);
- обеспечения целостности информации и цифровой подписи, в том числе создания инфраструктуры открытых ключей для СЦ и взаимодействующих с ним субъектов и объектов;
- мониторинга и противодействия сетевым компьютерным атакам, в том числе защиты от вирусов;
- создания и поддержания доверенной программно-технической среды, в том числе предотвращения скрытного внедрения в программные и технические средства;
- защиты речевой информации в процессах ситуационного управления и принятия решений;
- защиты информации от утечки по техническим каналам, в том числе по техническим проводным каналам инженерных обеспечивающих систем;
- контроля защищаемого контура размещения служб и средств СЦ и другие возможные подсистемы.

Важным системным требованием является требование обеспечения ИБ на всем технологическом пространстве СЦ. При этом функция защиты определенного вида и назначения, относящаяся к выделенной функциональной подсистеме СЗИ, может выполняться в различных компонентах СЦ, технологических процессах и точках программно-технической среды (например, функция аутентификации).

Другим распространенным вариантом композиционного соотношения функций защиты является вариант, при котором одна и та же функция защиты должна выполняться в различных функциональных подсистемах СЗИ, оставаясь в то же время функцией выделенной функциональной подсистемы СЗИ (например, мониторинг и противодействие сетевым атакам).

Чтобы избежать такого структурного противоречия, при построении СЗИ СЦ продуктивно включать в функциональную часть *функциональные комплексы защиты информации*, которые становятся самостоятельными компонентами СЗИ СЦ. При этом интеграция однородных функциональных задач в рамках подсистемы СЗИ образует функциональный комплекс I вида, а интеграция однородных функциональных задач, реализуемых в отношении одной и той же функции назначения в различных подсистемах СЗИ, образует функциональный комплекс II вида.

Выделение функциональных комплексов при создании СЗИ и последующей ее эксплуатации — это очевидный путь к типовым решениям по применению или созданию средств защиты информации с учетом включенных в эти средства возможностей настройки на параметры политики безопасности и информационного взаимодействия в части сбора служебной контролирующей информации для реализации управления.

Схема формирования функциональных комплексов показана на рис. 2.

Обеспечивающая часть СЗИ СЦ — совокупность мер и средств, технических, организационных и правовых решений, обеспечивающих реализацию функциональных задач.

К обеспечивающей части предъявляются общесистемные требования [12]:

- должны быть разработаны и доведены до уровня регулярного использования все необходимые механизмы гарантированного обеспечения требуемого уровня защиты информации;
- механизмы требуемого уровня защиты информации должны существовать в практической реализации в виде средств и технологий;
- необходимо располагать сертифицированными средствами (или отвечающими требованиям обработки конфиденциальной информации соответствующего уровня) реализации всех необходимых технических решений по защите информации;
- должны быть разработаны способы, решения и условия оптимальной организации и обеспечения проведения всех мер и мероприятий по защите в процессе обработки информации;

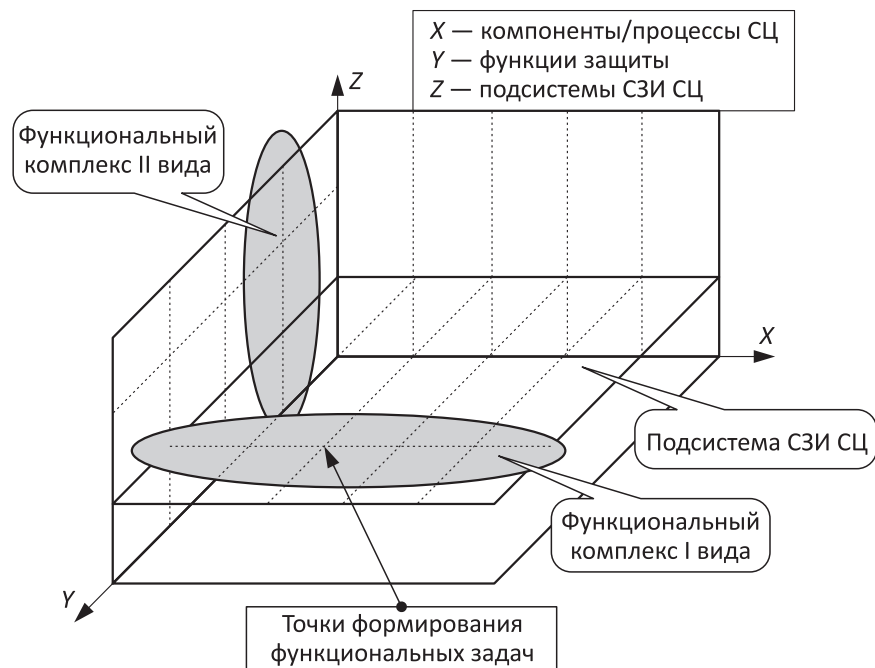


Рис. 2 Схема формирования функциональных комплексов СЗИ СЦ

- должно быть создано правовое поле для реализации функций защиты, механизмы и средства реализации правовых положений и требований при функционировании средств защиты и СЗИ СЦ в целом.

Системная интеграция функциональных задач и средств их реализации обеспечивается, наряду с *технологическим обеспечением СЗИ, управлением* как системным компонентом СЗИ. Основанием для принятия решений при управлении служат результаты мониторинга состояния ИБ. Как правило, в СЗИ СЦ выделяется *управление процессами защиты и менеджмент обеспечения ИБ в СЦ* в целом [10].

Технологическая интеграция и управление должны обеспечить достижение следующих основных целей [4]:

- предотвращение утечки информации по техническим каналам, в том числе речевой информации, а также утечки за счет возможно внедренных специальных устройств негласного получения информации;
- предотвращение несанкционированного уничтожения, искажения, копирования и блокирования информации за счет НСД к ресурсам и сетевых атак;
- соблюдение правового режима использования информационных ресурсов и программ обработки информации;

- обеспечение целостности информации при ее получении, обработке и хранении в СЦ;
- сохранение управляемости процессами обработки, хранения и использования информации в условиях несанкционированных воздействий на защищаемую информацию и среду ее обработки;
- обеспечение защищенного обмена информацией между объектами СЦ, в том числе удаленного и внешнего информационного обмена.

Проектирование и последующая эксплуатация СЗИ СЦ должна базироваться на концептуальных нормативных документах: *Модели угроз информационной безопасности*, *Модели вероятного нарушителя* и *Политики информационной безопасности СЦ*. Модели угроз и вероятного нарушителя являются исходными концептуальными документами, подготавливаемыми в основном по результатам

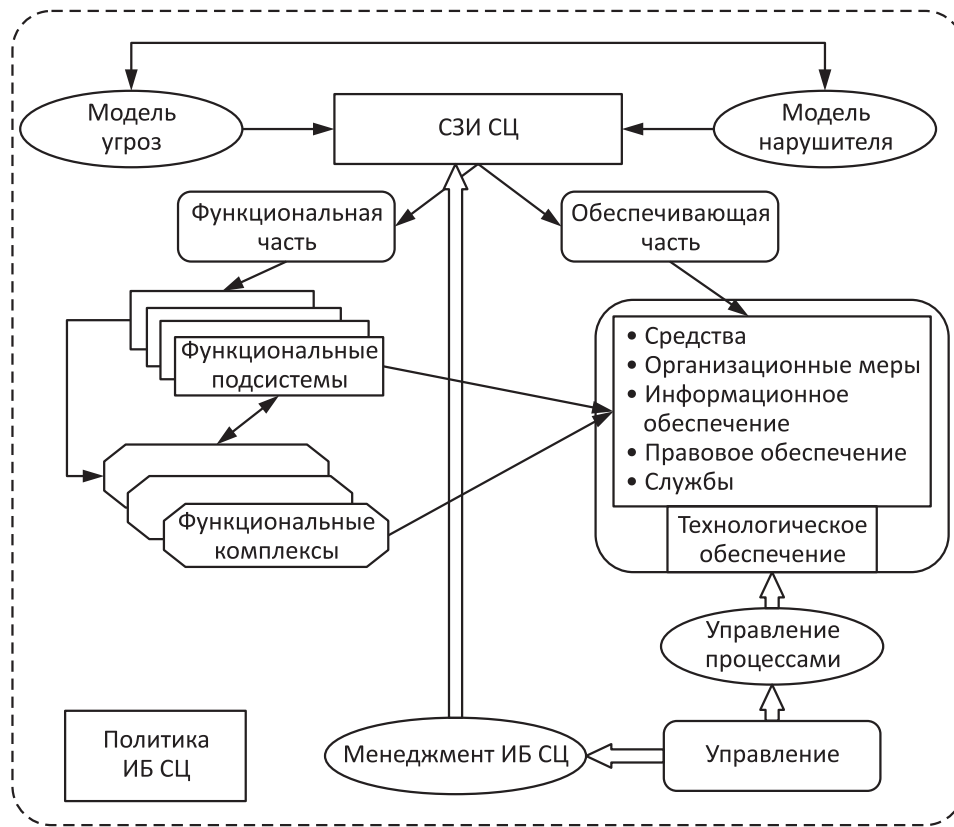


Рис. 3 Схема структурных отношений между компонентами СЗИ СЦ

предпроектного обследования. Они позволяют определить актуальные угрозы, уязвимости, должны ответить на вопросы — что защищать и от кого защищать, конкретизировать постановку задач по защите информации и тем самым обеспечить минимизацию затрат на ИБ.

Политика ИБ в своей основе рассчитана на время жизни СЦ. В виде нормативного системного документа она предъявляется на этапе приемо-сдаточных испытаний СЦ с возможной последующей корректировкой в ходе эксплуатации.

Схема структурных отношений между компонентами СЗИ СЦ представлена на рис. 3.

6 Актуальные задачи информационной безопасности ситуационного центра

Ситуационные центры являются объектами информационной деятельности с отчетливо выраженной спецификой. Они должны обеспечивать безопасное агрегирование получаемых из различных источников больших объемов различной информации и ее информационно-аналитическую обработку средствами собственной инфраструктуры. В СЦ стекается информация, разнородная не только по своей сути и форме (документы, потоковые данные, телеметрия и пр.), но и по уровню конфиденциальности — «грифу». В связи с этим общие требования по ИБ к СЦ формируются на основе требований к защищаемым ресурсам с наивысшим обрабатываемым «грифом».

В связи со спецификой СЦ, их усложнением и бурным развитием в настоящее время появляются новые актуальные задачи по обеспечению ИБ СЦ [13–15]. Практическое решение этих задач все более подтверждает тот факт, что функции защиты информации становятся неотъемлемой частью функциональных сервисов обработки информации и архитектурных решений при построении системных инфраструктурных компонентов. Это, безусловно, влияет на проектные решения СЗИ СЦ.

К таким актуальным задачам можно отнести:

- *сегментирование защищаемых ресурсов СЦ* (информационных, информационно-технологической инфраструктуры и программно-технической среды) — выделение контуров безопасности по «грифу». Это решение считается одним из доверенных механизмов защиты для накопительных информационных систем с большими объемами разнородной поступающей информации;
- *создание защищенной шины информационного взаимодействия*, способной объединить разнородные первичные ресурсы в случае их выраженной территориальной распределенности. Применение терминального доступа или использование клиент-серверных технологий в этом случае неэффективно, поскольку необходима реализация высоких требований по пропускной способности и качеству каналов связи, образующих единую сеть;

- *построение СЦ как частного облака*, когда возникает проблема виртуального пространства взаимодействия. Это приводит к усложнению СЗИ СЦ с учетом специфических угроз, присущих виртуальной среде исполнения;
- *организацию эргономичного рабочего места* с минимизацией многооконного интерфейса, не вполне удобного и понятного для ЛПР. Интеграция разноразмерных систем — «генераторов» — в рамках «одного окна» является сложнейшей проблемой не только с точки зрения обеспечения ИБ, но и с инженерной точки зрения;
- *безопасное хранение в ИАС СЦ поколений разнородной, в том числе графической и потоковой, информации*, что предъявляет жесткие требования по ИБ к системам управления базами данных (СУБД). Это требует применения сертифицированных СУБД или использования наложенных средств контроля доступа. Последнее решение существенно сужает возможности и самой СУБД, и средств углубленного анализа;
- *реализацию режима чрезвычайной ситуации*, когда все внимание СЦ направляется на формирование новостного информационного потока исключительно вокруг интересующего события. В этом случае уровень конфиденциальности СЦ может быть повышен до максимального уровня. Возможен разрыв стыков с открытыми контурами СЦ. Ситуационный центр может стать автономным объектом, взаимодействие с которым осуществляется либо по доверенным каналам связи, либо посредством подкачки данных с отчуждаемых носителей, физически доставляемых в СЦ;
- *формирование электронного указания* в виде приказа или распоряжения как способ управляющего воздействия. Это потребует электронного документооборота, обеспечивающего возможность статуса правового подтверждения, что, в свою очередь, потребует сопряжения СЦ с удостоверяющим центром и создания инфраструктуры открытых ключей для всего сетевого сообщества СЦ.

Следует отметить, что перечисленные и ряд других актуальных задач обеспечения ИБ СЦ, особенно связанных с планируемым созданием в России сети распределенных СЦ [8, 16], требуют своего достаточно быстрого и качественно-го решения, чтобы поддержать большую востребованность СЦ для решения задач управления в современном быстро меняющемся обществе.

Литература

1. *Маршев В. И.* История управленческой мысли. — М.: Инфра-М, 2005. 731 с.
2. Словарь терминов МЧС. — М.: Федеральное государственное унитарное авиационное предприятие МЧС России EdwART, 2010.
3. *Зацаринный А. А., Сучков А. П.* Системы ситуационных центров специального назначения. Основные определения, понятия и подходы к созданию // Межотраслевая информационная служба, 2015. № 4. С. 31–41.

4. *Зацаринный А. А.* Организационные и системотехнические подходы к построению современных ситуационных центров // Методы построения и технологии функционирования ситуационных центров: Сб. науч.-технич. статей / Под ред. А. А. Зацаринного. — М.: ИПИ РАН, 2011. С. 10–25.
5. *Королёв В. И.* Методология построения комплексной защиты информации на объектах информатизации // Системы высокой доступности, 2009. Т. 5. № 4. С. 4–24.
6. ГОСТ Р ИСО/МЭК 27002-2012. Информационная технология. Методы и средства обеспечения безопасности. Свод норм и правил менеджмента информационной безопасности.
7. ГОСТ Р 51275-99. Защита информации. Объект информатизации. Факторы, воздействующие на информацию.
8. *Зацаринный А. А.* Ситуационные центры как основа информационно-аналитической поддержки принятия решений в органах государственной власти // Аналитика развития и безопасности страны: реалии и перспективы: Сб. мат-лов 1-й Всеросс. конф. — М.: Столица, 2014. С. 277–295.
9. Подход системной инженерии к управлению жизненным циклом. Понятийный минимум. PraxOS. 1.0, 2008. http://techinvestlab.ru/files/495344/se_Is_minimum_praxos_1.doc.
10. *Королёв В. И.* Концептуальные аспекты построения информационной безопасности корпораций // Системы высокой доступности, 2015. Т. 11. № 2. С. 50–64.
11. *Герасименко В. А.* Защита информации в автоматизированных системах обработки данных. — М.: Энергоатомиздат, 1994. Кн. 1. 400 с.
12. *Королёв В. И., Сухотин И. Н., Королёв А. В.* Интегрированные системы безопасности и влияние информационных рисков на деятельность организации // Технологии безопасности: Официальный отчет X Международного форума: Сб. мат-лов / Под ред. И. К. Филоненко, Н. В. Александровой. — М.: ПРОЭКСПО, 2005. С. 284–291.
13. *Андреев В.* Защита информации в ситуационном центре — ключевые аспекты: Типовые и специальные задачи // Connect, 2012. № 5. <http://www.ситцентр.рф/pressa/andreev.pdf>.
14. *Шарай В. А., Малашихин А. К.* Логико-вероятностные модели угроз информационной безопасности ситуационных центров // Научные труды КубГТУ, 2014. № 6. <http://ntk.kubstu.ru/file/306>.
15. *Никифоров Д.* Об информационной безопасности будущей системы ситуационных центров в России // Connect WIT, 2015. № 7-8. <http://www.connect-wit.ru/category/magazine/connect-wit-2015-7-8>.
16. Стратегия национальной безопасности Российской Федерации. Утверждена Указом Президента РФ от 31.12.2015 № 683. Ст. 112.

Поступила в редакцию 09.02.16

INFORMATION SECURITY SITUATION CENTER

A. A. Zatsarinny¹ and V. I. Korolev^{1,2}

¹Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilova Str., Moscow 119333, Russian Federation

²Department “Cybernetics and Information Security,” National Research Nuclear University “MEPhI,” 31 Kashirskoye Sh., Moscow 115409, Russian Federation

Abstract: The paper considers the purpose and characteristics of situational centers as information technology facilities that ensure the implementation of situational management. A situational center is a system. The paper indicates complicated composition and structure of information resources used to make decisions. The variety of levels of confidentiality related to the processed information is a factor of situational centers information security. Information security is performed as information protection. The paper presents an approach to building the systems of information security and protection for situational centers based on the architectural approach. Herewith, this architecture of any run organization is of a set of business architecture + IT architecture and the concept of system architecture complies with ISO/IEC 15288-2008. The paper formulates the actual problems of information security situational centers.

Keywords: situational center; situational management; informational resources; information and communication technologies; confidentiality; information security; system of information protection; architectural approach

DOI: 10.14357/08696527160109

Acknowledgments

The work was supported by the Russian Foundation for Basic Research (project 15-29-07981 ifi-m).

References

1. Marshev, V. I. 2005. *Istoriya upravlencheskoy mysli*. [A story of administrative thought]. Moscow: Infra-M. 731 p.
2. *Slovar' terminov MChS* [Dictionary of terms of the Ministry of Emergency Situations]. 2010. Federal State Unitary Air Enterprise of Emercom of Russia of EdwART.
3. Zatsarinny, A. A., and A. P. Suchkov. 2015. Sistemy situatsionnykh tsentrov spetsial'nogo naznacheniya. Osnovnye opredeleniya, ponyatiya i podkhody k sozdaniyu [Systems of the situational centers of a special purpose. Main definitions, concepts and approaches to creation]. *Mezhotraslevaya Informatsionnaya Sluzhba* [Interindustry Information Service] 4:31–41.

4. Zatsarinny, A. A. 2011. Organizatsionnye i sistemotekhnicheskie podkhody k postroeniyu sovremennykh situatsionnykh tsentrov [Organizational and sistemotekhnical approaches to creation of the modern situational centers]. *Sb. nauch.-tekhnich. statey "Metody postroeniya i tekhnologii funktsionirovaniya situatsionnykh tsentrov"* [Methods of construction and technology of functioning of the situational centers, Collection of Scientific and Technical Articles]. Ed. A. A. Zatsarinny. Moscow: IPI RAN. 10–25.
5. Korolev, V. I. 2009. Metodologiya postroeniya kompleksnoy zashchity informatsii na ob'ektakh informatizatsii [Methodology of creation of complex information security on objects of informatization]. *Sistemy Vysokoy Dostupnosti* [Systems of High Availability] 4(5):4–24.
6. GOST R ISO / IEC 27002-2012. Informatsionnaya tekhnologiya. Metody i sredstva obespecheniya bezopasnosti. Svod norm i pravil menedzhmenta informatsionnoy bezopasnosti [Information technology. Methods and means of ensuring of safety. Arch of norms and rules of management of information security].
7. GOST R 51275-99. Zashchita informatsii. Ob'ekt informatizatsii. Faktory, vozdeystvuyushchie na informatsiyu [Information security. Object of informatization. The factors influencing information].
8. Zatsarinny, A. A. 2014. Situatsionnye tsentry kak osnova informatsionno-analiticheskoy podderzhki prinyatiya resheniy v organakh gosudarstvennoy vlasti [Situational centers as a basis of information and analytical support of decision-making in Government]. *Sb. mat-lov 1-y Vseross. Konf. "Analitika razvitiya i bezopasnosti strany: Realii i perspektivy"* [Collection of materials of the 1st All-Russian Conference "Analytics of Development and Safety of the Country: Realities and Prospects"]. Moscow: Stolitsa. 277–295.
9. PraxO.S. 1.0, 2008. Podkhod sistemnoy inzhenerii k upravleniyu zhiznennym tsiklom. Ponyatiynnyy minimum [Approach of system engineering to management of life cycle. Conceptual minimum]. Available at: http://techinvestlab.ru/files/495344/se_js_minimum_praxos_1.doc (accessed February 9, 2016).
10. Korolev, V. I. 2015. Kontseptual'nye aspekty postroeniya informatsionnoy bezopasnosti korporatsiy [Conceptual aspects of creation of information security of corporations]. *Sistemy Vysokoy Dostupnosti* [Systems of High Availability] 2(11):50–64.
11. Gerasimenko, V. A. 1994. *Zashchita informatsii v avtomatizirovannykh sistemakh obrabotki dannykh* [Information security in the automated systems of data processing]. Moscow: Energoatomizdat. Vol. 1. 400 p.
12. Korolev, V. I., I. N. Sukhotin, and A. V. Korolev. 2005. Integrirovannyye sistemy bezopasnosti i vliyanie informatsionnykh riskov na deyatelnost' organizatsii [The integrated security systems and influence of information risks on activity of the organization]. *Ofitsial'nyy otchyot X Mezhdunar. Forumy "Tekhnologii Bezopasnosti."* *Sb. mat-lov.* [The Official Report of the 10th Forum (International) "Technologies of Safety." The collection of Materials]. Eds. I. K. Filonenko and N. V. Aleksandrova. Moscow: PROEKSPRO. 284–291.
13. Andreyev, V. 2012. Zashchita informatsii v situatsionnom tsentre — klyucheveye aspekty. Tipovye i spetsial'nye zadachi [Information security in the situational center — key aspects. Standard and special tasks]. *Connect* 5. Available at: <http://www.ситцентр.рф/pressa/andreev.pdf> (accessed February 9, 2016).

14. Sharay, V. A., and A. K. Malashikhin. 2014. Logiko-veroyatnostnye modeli ugroz informatsionnoy bezopasnosti situatsionnykh tsentrov [Logic and probabilistic model of threats of information security of the situational centers]. *Nauchnye trudy KubGTU* [Scientific works of KubGTU] 6. Available at: <http://ntk.kubstu.ru/file/306> (accessed February 9, 2016).
15. Nikiforov, D. 2015. Ob informatsionnoy bezopasnosti budushchey sistemy situatsionnykh tsentrov v Rossii [About information security of future system of the situational centers in Russia]. *Connect* 7–8. Available at: <http://www.aladdin-rd.ru/company/pressroom/articles/43615> (accessed February 9, 2016).
16. Strategiya natsional'noy bezopasnosti Rossijskoy Federatsii, utverzhdena Ukazom Prezidenta RF ot 31.12.2015 No. 683, st. 112 [Strategy of national security of the Russian Federation, approved by the Decree of the Russian President of 31.12.2015 No. 683, Art. 112].

Received February 9, 2016

Contributors

Zatsarinny Alexander A. (b. 1951) — Doctor of Science in technology, professor, Deputy Director, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences (FRC CSC); Institute of Informatics Problems of FRC CSC, 44-2 Vavilov Str., Moscow 119333, Russian Federation; AZatsarinny@ipiran.ru

Korolev Vadim I. (b. 1943) — Doctor of Science in technology, leading scientist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; professor, Department “Cybernetics and Information Security,” National Research Nuclear University “MEPhI”, 31 Kashirskoye Sh., Moscow 115409, Russian Federation; VKorolev@ipiran.ru

ПРИНЦИПЫ СОЗДАНИЯ СИСТЕМЫ МОНИТОРИНГА И АНАЛИЗА МИРОВОГО ПОТОКА НАУЧНО-ТЕХНИЧЕСКОЙ ИНФОРМАЦИИ

Ю. П. Калинин¹, Ал-др А. Хорошилов², Ал-ей А. Хорошилов³

Аннотация: Рассматриваются принципы создания системы мониторинга и анализа мирового потока научно-технической информации (НТИ), базирующиеся на современных представлениях о смысловой структуре текстов. Описываются методы и инструменты формализации смысловой структуры текстовой информации и методы создания декларативных средств для широкого спектра автоматизированных систем обработки и анализа текстовой информации.

Ключевые слова: автоматизированная обработка текстов; семантический анализ; формализованное описание текста; смысловая структура; лингвистическое программное обеспечение; декларативные средства

DOI: 10.14357/08696527160110

1 Введение

На современном этапе развития научно-технологического потенциала Российской Федерации особую актуальность приобретает необходимость принятия управленческих решений, направленных на повышение эффективности научно-исследовательских и опытно-конструкторских работ. Такие решения должны базироваться на использовании объективных аналитических данных о тенденциях развития мирового научно-технологического комплекса. Эти данные могут быть получены путем содержательного комплексного анализа мирового потока НТИ. Источниками такой информации могут служить зарубежные и отечественные отраслевые и ведомственные порталы, сайты научных учреждений, редакций журналов, научных конференций и сообществ, электронные средства массовой информации, а также множество других источников НТИ, размещенных в сети Интернет. В процессе смыслового анализа разноязычных гетерогенных распределенных научно-технических источников информации могут быть получены новые знания и определены приоритетные научные направления деятельности

¹Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, urkalinin@mail.ru

²Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, khoroshilov@mail.ru

³Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, alex_khoroshilov@mail.ru

отечественных и зарубежных научных коллективов и выявлены мировые перспективные опытно-конструкторские разработки.

Между тем отсутствие на IT-рынке апробированных технологий и сервисов высокоскоростной автоматической смысловой обработки разноязычной неструктурированной текстовой информации является существенным сдерживающим фактором для реализации этой задачи. В связи с этим необходимость создания промышленных технологий извлечения знаний из разнородных текстов является чрезвычайно актуальной. Эти технологии должны базироваться на современных представлениях о смысловой структуре разноязычных текстов и на новых высокопроизводительных методах и средствах анализа неструктурированной текстовой информации.

Комплекс перечисленных задач можно решить в процессе создания системы мониторинга и анализа мирового потока НТИ, базирующейся на применении перспективных информационных технологий и современных вычислительных ресурсов. Потребителями аналитической информации, полученной в результате функционирования такой системы, могут быть органы государственной власти, научные и образовательные организации, промышленные и внедренческие центры, технопарки, инвестиционные компании, экспертное сообщество, ученые и специалисты в различных областях науки и техники.

2 Тенденции развития систем анализа текстовой информации

Анализ материалов зарубежной и отечественной научно-технической литературы по рассматриваемой проблеме показывает, что в мире сейчас созданы и функционируют сотни поисково-аналитических систем различного назначения. Значительная их часть связана с анализом банковской деятельности и деятельностью высокотехнологичных компаний. Такие системы получили название системы Business Intelligence (BI). Этот термин, обозначающий особый класс информационных систем, был предложен еще в 1958 г. сотрудником IBM Хансон Питером Луном. Изначально BI рассматривались как инструменты, используемые для преобразования, хранения, анализа, моделирования, доставки и трассировки информации в ходе работы над задачами, связанными с принятием бизнес-решений на основе фактических данных. В дальнейшем спектр решаемых бизнес-задач значительно расширился, и сейчас под термином BI понимается информационно-аналитическая система, реализующая функции сбора, обработки, анализа и подготовки форм отчетности для различных структур бизнес-сообщества. Расширение функциональных возможностей таких систем отразилось в их названии, и такие системы все чаще стали называть системами Business Analytics (BA). Но при этом BA решают широкий круг аналитических задач, используя традиционные средства BI на качественно новом уровне.

Применяемые традиционные подходы для решения всего комплекса аналитических бизнес-задач сейчас исчерпали себя в силу ряда причин, главной из которых является их недостаточная динамичность и неприспособленность

к оперативной обработке больших объемов неструктурированных данных. Компенсировать возникшие угрозы можно только путем разумной интеграции проверенных временем традиционных решений с современными технологиями работы с неструктурированными данными. Сегодня такие технологии называют технологиями Embraced Search and Retrieval (ESR), реализующими две всеохватывающие (Embraced) функции — корпоративный поиск (Enterprise Search) и извлечение информации из данных (Retrieval).

Первыми, кто заметил необходимость расширения функциональных возможностей своих продуктов, были крупнейшие поставщики программного обеспечения, специализирующиеся на корпоративных решениях бизнес-аналитики — компании Oracle, IBM и HP. Поэтому начиная с 2011 г. эти фирмы начали интенсивно скупать компании, предлагающие решения в области ESR и работающие с неструктурированными данными. Одна из таких компаний — Autonomy (за контрольный пакет которой Hewlett-Packard выложила \$10 млрд) — является лидером в области смысловой обработки текстовой информации. Идеологическая основа технологий этой компании — концепция Meaning-Based Computing (MBC), позволяющая выделять смысл из разнородных текстов, независимо от формы их представления. В этой компании разработано около 500 различных технологий и процедур обработки текстов и определена иерархическая последовательность технологического процесса обработки текстовой информации, на нижнем уровне которой стоит простой поиск по ключевым словам, а на верхнем — технология MBC.

При создании системы мониторинга и анализа мирового потока научно-технических публикаций неизбежно возникнет проблема, связанная с необходимостью оперативной обработки огромных объемов текстовой информации в рамках имеющихся технологических возможностей. Особенно остро эта проблема стоит при решении задач семантического анализа текстов, и связано это прежде всего с тем, что при таком анализе тексты подвергаются сложной и многоступенчатой обработке с использованием словарей больших объемов. Обработка текстов выполняется последовательно комплексом процедур семантико-синтаксического и концептуального анализа. Используемый сейчас технологический процесс часто не позволяет обеспечить требуемые в настоящее время скорости обработки текстов. Между тем существенное повышение скоростей обработки текстов может лежать в плоскости уже апробированных решений проблемы big data (больших данных) [1, 2]. В качестве базового принципа обработки «больших данных» можно использовать массово-параллельную обработку, масштабируемую без деградации на множество узлов обработки. Такая обработка выполняется в рамках таких технологий, как NoSQL, MapReduce, Hadoop и др.

Разработка предлагаемой системы потребует решения ряда сложных научных проблем. Большая часть этих проблем (мультиагентный поиск и сбор разнородной информации из различных источников, консолидация и хранение этой информации) решена в процессе создания аналогичных систем, в которых авторы принимали непосредственное участие. В частности, в рамках проекта

«Мониторинг СМИ» (выполняемого в рамках задач Ситуационно-кризисного центра Федерального агентства по атомной энергии — «СКЦ Росатома») решалась задача сбора, консолидации, обработки и анализа текстов информации из более чем 6 тыс. интернет-источников, включая электронные средства массовой информации и профессиональные отраслевые отечественные и зарубежные сайты. В настоящее время разработанная система функционирует в промышленном режиме. Ежедневно на вход этой системы поступает около 100 тыс. документов, а общий объем накопленных документов составляет более 30 млн единиц. Система автоматически формирует реферативные сборники по выборкам документов в соответствии с информационными профилями пользователей.

Аналогичные задачи решались в рамках проекта системы ЕСУ НИОКР (Единой системы учета результатов научно-исследовательских, опытно-конструкторских работ гражданского назначения). Эта система сейчас также функционирует в промышленном режиме и обеспечивает автоматическую смысловую обработку полных текстов информационных фондов ФГАНУ ЦИТИС, представленных в виде коллекций 3,5 млн реферативных документов и 400 тыс. полнотекстовых документов. В процессе автоматической обработки документов обеспечивается возможность выявления в них семантических объектов (стран, городов, фамильно-именных групп, географических объектов и т. д.). В этой системе может выполняться в реальном масштабе времени семантический поиск документов.

В процессе создания этих систем решались задачи семантического поиска, фильтрации, рубрикации и кластеризации документов, автоматического аннотирования документов или групп документов, поиск похожих документов и дубликатов, выявление заимствований в текстах или сегментирование текстов на более мелкие смысловые фрагменты. При этом подавляющее большинство методов обработки неструктурированной текстовой информации было ориентировано на разнообразные лингвистические и онтологические модели предметных областей. Можно с уверенностью констатировать, что используемые в настоящее время модели предметной области несовершенны и не способны в полной степени отразить все многообразие наименований понятий и их связей в любой конкретной области, не говоря о таких масштабных областях, как, например, область биотехнологий или область авиационных и космических технологий.

При обработке мирового потока эта проблема осложняется еще как минимум двумя факторами — большими масштабами и наличием разнородных и разноязычных текстов. Первый фактор связан с огромными объемами текстовой информации, относящейся к широкому спектру тематических областей. Основной проблемой, которая не получила до настоящего времени должного решения, является проблема высокоэффективной семантической обработки разнородной и разноязычной текстовой информации. Современные технологии анализа текстов используют различные семантические инструменты, например инструмент онтологий. Следует отметить, что создание онтологических ресурсов требует больших временных затрат на их создание и, как показывает практика, их объ-

емы редко превышают несколько десятков тысяч. При этом если оперировать реальными численными значениями системы понятий мирового потока информации, то по данным международной организации INFOTERM (Международный терминологический центр, г. Вена, Австрия) число различных терминов в развитых естественных языках достигает 50 млн, а количество наименований товаров — 100 млн [3], поэтому необходимый перечень наименований понятий для решения этой задачи должен соотноситься с их числом в мировом потоке научно-технической информации. Кроме того, в онтологиях наименования понятий представлены в обобщенной форме, а в научно-технических текстах они встречаются в их всевозможных конкретных представлениях, которые не всегда присутствуют в составе конкретной онтологии. Выходом из создавшейся ситуации может быть использование таких лингвистических инструментов, как концептуальные словари большого объема, построенные по принципу «тематический словарь плюс политематический словарь». Такой инструмент можно создать в относительно сжатые сроки, и он может обеспечить достаточное покрытие текстов при их семантическом анализе.

Вторым фактором, усложняющим проблему смыслового анализа документов, является их представление на различных естественных языках, поэтому здесь необходимо также решить проблему правильной передачи смысла с одного естественного языка на другой. Эту задачу можно выполнить с помощью современной промышленной системы перевода, обеспечивающей адекватный перевод научно-технических текстов по широкому спектру тематик [4]. Словарная база таких систем также должна отражать понятийный состав широкого спектра тематических областей. Примером такой словарной базы может служить комплекс словарей системы МетаФраз, в которой только тематический словарь «Авиационные и космические технологии» включает 800 тыс. терминологических и фразеологических словосочетаний, а политематический словарь (служащий для перевода периферийной лексики) — 3,5 млн словосочетаний [5, 6].

3 Теоретические предпосылки для создания систем анализа текстовой информации

Предлагаемый подход к созданию систем анализа текстовой информации должен быть основан на правильных представлениях о смысловой структуре языка и речи. По современным представлениям наиболее устойчивыми единицами смысла являются понятия. Они занимают центральное место в языке и речи и являются теми базовыми строительными блоками, на основе которых формируются смысловые единицы более высоких уровней. Смысл текстовых документов выражается с помощью единиц смысла, входящих в их состав, поэтому центральной процедурой любых систем автоматической смысловой обработки текстов, в том числе и систем мониторинга текстов, должна быть процедура их семантико-синтаксического концептуального анализа [7–10]. В процессе анализа

документов должно выполняться распознавание их принадлежности к научным направлениям в соответствии с тематическими классами используемой системы классификации. Этот процесс основан на процессе распознавания смысловой близости классифицируемых документов и предметных рубрик классификатора. При этом для классифицируемых документов и рубрик классификатора необходимо составить формализованные концептуальные образы документа (КОД). Формализованные описания тематических рубрик классификаторов могут создаваться путем автоматического концептуального анализа ранее расклассифицированных массивов документов. В связи с этим основой декларативных средств автоматического концептуального анализа должны служить представительные политематические словари наименований понятий (концептуальные словари объемом несколько миллионов статей и покрывающие научно-технические тексты не менее чем на 99,5%). При анализе текстов необходимо также учитывать, что в них одни и те же объекты и процессы могут описываться с различной степенью общности и с помощью различных языковых средств. Поэтому при решении задач автоматической смысловой обработки текстовой информации необходимо учитывать такие явления, как синонимия, гипонимия (родовидовые отношения) и разнообразие средств выражения межфразовых связей.

Основной структурной единицей текста традиционно считается предложение. Предложения выступают в тексте не изолированно друг от друга, а в тесной смысловой связи. В основе этой связи лежат мыслительные образы тех конкретных или абстрактных объектов (ситуаций, явлений), которые человек имеет в виду, когда он порождает текст. Образы этих объектов имеют определенную структуру. Кроме того, они дополнительно структурируются человеком при их описании на естественном языке. Соответственно этому структурируется и текст [7, 10].

Исходя из вышесказанного, при решении задачи формализации смыслового содержания текстов необходимо методами семантико-синтаксического и концептуального анализа обработать текст, разделить его на предложения, выделить из него единицы смысла (наименования понятий) — слова и словосочетания, выражающие понятия, и установить между ними смысловые связи.

Эти задачи могут быть решены путем использования базового набора процедур семантико-синтаксического и концептуального анализа текстов, входящих в состав современных систем автоматической обработки текстовой информации. Такими средствами располагает базовое платформенное лингвистическое программное обеспечение (ПО) МетаФраз [11–17]. Кратко рассмотрим этот набор процедур семантико-синтаксического и концептуального анализа текстов.

4 Базовые процедуры анализа текста

Основным назначением базовых процедур анализа текста является структурирование и формализация его смыслового содержания, выявление понятийного состава предметной области, установление парадигматических, синтагматических и ассоциативных связей между наименованиями понятий и установление

их контекстного окружения. Ниже приводится описание основных процедур концептуального семантико-синтаксического анализа текстов.

4.1 Графематический анализ текста

Графематический анализ предназначен для предварительного анализа текста по представляющей его последовательности символов. Методы и алгоритмы, реализующие этот анализ, описаны в ряде работ [1, 18, 19]. В результате этого анализа определяется язык текста, устанавливаются местоположения слов, предложений, абзацев, фамильно-именной группы, дат и электронных адресов. Для автоматического определения указанной информации о формальной структуре текста в соответствующих методах графематического анализа используется следующий набор грамматических таблиц и словарей:

- словарь для установления языка текста;
- таблица признаков для выделения слов и разделителей в тексте;
- таблица признаков для выделения дат в цифровых форматах;
- таблица признаков для выделения фамильно-именной группы;
- таблица признаков для выделения электронных адресов;
- таблица признаков для разделения текста на предложения;
- таблица признаков для разделения текста на абзацы;
- таблица признаков для выделения примечаний.

Исходными данными служит последовательность символов исходного текста.

Выходными данными является информация о местоположении слов, разделителей, дат, установленных фамильно-именных групп, электронных адресов, предложений, заголовков и примечаний.

4.2 Морфологический анализ слов текста

Морфологический анализ слов естественных языков предназначен для определения структуры слов и назначения им грамматических признаков, необходимых для выполнения последующих процедур автоматической обработки текстовой информации (например, морфологического синтеза слов, синтаксического анализа и синтеза текстов и их концептуального анализа). Методы и алгоритмы, реализующие этот анализ, описаны в работах [1, 18, 19]. В результате этого анализа определяется структура слова и набор грамматической информации (ГИ), определяющей это слово вне контекста. Для автоматического определения указанной информации о слове в соответствующих методах морфологического анализа используется следующий набор грамматических таблиц и словарей:

- словарь слов-исключений (коротких и служебных слов) с назначенной ГИ;
- словарь конечных буквосочетаний слов русского языка;

- таблица флективных классов слов;
- таблица супплетивных форм слов;
- таблица установления наличия чередований в основах слов;
- таблица подстановок для реализации чередований в основах слов;
- таблица наборов ГИ слов;
- таблица окончаний русских слов.

Исходными данными служит буквенный код отдельного слова.

Выходными данными является информация о структуре слова и набор ГИ, определяющей это слово вне контекста.

4.3 Семантико-синтаксический анализ предложений текста

Семантико-синтаксический анализ текстов проводится с целью формализованного представления их структуры — выделения в них смысловых единиц и установления связей между ними. При этом структура текстов может интерпретироваться по-разному и описываться на различных формализованных языках. При описании синтаксической структуры текстов удобно опереться на какую-либо ее формализованную модель, например на модель дерева зависимостей. Согласно этой модели каждое предложение представляется в виде дерева, в узлах которого находятся слова. Слова соединяются друг с другом стрелками, выражающими отношения непосредственной доминации и направленными от подчиняющего (определяемого) слова к подчиненному (определяющему). Степень дифференциации этих отношений может быть разная, причем чем больше степень дифференциации, тем сложнее процесс описания текстов. Методы и алгоритмы, реализующие этот анализ, описаны в работах [1, 18, 19]. В результате этого анализа определяется синтаксическая структура предложения: производится членение на простые предложения, определяются главные и второстепенные члены предложения и устанавливаются смысловые связи между ними, строится дерево зависимости предложения и для каждого слова определяется однозначная ГИ, соответствующая контексту. Для автоматического определения указанной информации о структуре предложения в соответствующих методах семантико-синтаксического анализа используется следующий набор грамматических таблиц и словарей:

- таблица описаний правил синтаксического анализа;
- таблица описаний правил установления смысловых связей между словами предложения;
- таблица описаний правил определения однозначной ГИ слов с учетом контекста.

Исходными данными служат слова предложения и результаты их обработки процедурой морфологического анализа.

Выходными данными является информация о семантико-синтаксической структуре предложения.

4.4 Концептуальный анализ текстов

Концептуальный анализ текстов предназначен для определения смысловой структуры текстов, выявления понятийного (концептуального) состава текстов и установления связей между наименованиями понятий. Эту задачу невозможно решить только путем анализа синтаксической структуры текстов без привлечения семантических признаков. Сложность этой задачи связана, прежде всего, с вариативностью форм представления наименований понятий в текстах. Авторами было разработано несколько вариантов решения этой задачи [1, 20]. Наиболее эффективным решением оказался метод концептуального анализа текстов с контролем по тезаурусу (эталонному словарю), включающему более 1,8 млн понятий и свыше 400 тыс. связей между ними. В результате этого анализа в тексте выявляются наименования понятий, устанавливается концептуальная структура текста и строится таблица связей между наименованиями понятий. Для автоматического определения указанной информации о структуре текста в соответствующих методах концептуального анализа используется следующий набор грамматических таблиц и словарей:

- эталонный словарь наименований понятий;
- словарь смысловых связей между наименованиями понятий;
- словарь смысловых связей слов.

Исходными данными служит информация о семантико-синтаксической структуре предложения.

Выходными данными является таблица связей между наименованиями понятий.

Таблица связей наименований понятий представляет собой машинное представление смысловой структуры текста. Его визуальное представление можно сформировать в виде семантической карты текста, представляющей собой ориентированный граф, в узлах которого находятся объекты, события или темы документов, а дугами являются смысловые отношения между ними. Связи могут быть либо типизированными (определен семантический тип связи), либо логическими (установлен факт их наличия).

Семантическую карту можно построить с помощью пакета утилит Graphviz, разработанного специалистами лаборатории AT&T [21]. В качестве исходных данных для этой утилиты используется описание графа на специальном языке dot, а на выходе формируется граф в виде графического, векторного или текстового файла. При этом также возможен более сложный выход, например

с использованием координатной сетки, которую потом можно использовать для обозначения областей при показе на странице гипертекста.

Основными параметрами, по которым может оцениваться функционирование процедур семантического анализа текстов, являются качество анализа текстов и скорость его обработки. Качество анализа текстов определяется, прежде всего, использованием адекватной модели представления их смысловой структуры, эффективностью методов и алгоритмов анализа текстов, составом декларативных средств, обеспечивающих высокое покрытие анализируемых текстов.

Скорость обработки текстов зависит от быстродействия применяемых методов и алгоритмов семантической обработки, числа проходов по тексту при его обработке и от объемов грамматических таблиц и словарей, используемых при обработке текста.

Из предыдущих рассуждений следует, что ориентация на фразеологические словосочетания как на основную форму представления наименований понятий в естественных языках позволяет более точно учесть семантико-синтаксическую структуру текстов и построить более эффективную систему смысловой обработки текстовой информации. Построение такой системы неизбежно связано с выявлением разноязычного понятийного состава, который, по самым скромным оценкам, содержит несколько сот миллионов наименований понятий. Поэтому любая система анализа текстов в перспективе должна включать в свой состав систему мощных политематических словарей наименований понятий, содержащих нескольких миллионов (или десятков миллионов) словарных статей, состоящих преимущественно из фразеологических словосочетаний. В словарях должны содержаться также сведения об отношениях синонимии и о родовидовых отношениях между понятиями.

5 Предлагаемые технические решения по архитектуре системы мониторинга и анализа мирового потока научно-технической информации

Ниже опишем предлагаемые технические решения по архитектуре системы мониторинга и анализа мирового потока НТИ (далее — Системы). Система может быть реализована в виде программно-технологического комплекса, представляющего собой совокупность средств вычислительной техники, ПО и средств создания и заполнения машинной информационной базы Системы. Доступ пользователей к Системе осуществляется через единый портал обеспечения регламентированного защищенного доступа к информации. Информационный обмен должен осуществляться через единое информационное пространство и посредством использования стандартизированных протоколов и форматов обмена данными.

Все компоненты Системы должны функционировать в пределах единого логического пространства, обеспеченного интегрированными средствами серверов

данных и серверов приложений. Информационная совместимость со смежными системами должна обеспечиваться через портал, на уровне экспорта-импорта XML-документов. Для осуществления взаимного переноса и регистрации информации из/в смежные системы должен использоваться согласованный перечень полей обмена и формат файлов типа XML.

Логическую основу хранилища Системы составляют расширенная объектная модель документа и мультиагентное индексирование (на уровне терминов, словосочетаний, стохастических фрагментов текста), использующее (если необходимо) нормализацию лексики, словари исключений или предопределенной лексики, тезаурусы, словари естественного языка и словари словосочетаний предметных областей. Комплексное использование таких средств должно обеспечить не только эффективный семантический поиск, но и выявление дублей и плагиата.

Система должна обеспечивать реализацию основных функций в рамках следующих подсистем:

подсистема ведения информационного хранилища документов обеспечивает выполнение следующих функций:

- учет и обработка информации в информационном хранилище;
- управление информационным хранилищем;
- загрузка и выгрузка данных;
- хранение информации различной структуры, предназначенной для обработки в рамках различных приложений, включая задачу обеспечения возможности эффективного хранения, использования и представления хранимой текстовой, формульной, иллюстративной и мультимедийной информации;
- создание и ведение страховых копий данных и документов;
- формирование базы данных различных семантических представлений документов, полученных путем их мультиагентного индексирования (на уровне терминов, наименований понятий, стохастических фрагментов текста), базирующегося на механизмах формализации лексики с использованием тезаурусов и семантических словарей;

подсистема сбора документов должна обеспечивать выполнение следующих основных функций:

- мониторинг заданных интернет-источников и постоянное отслеживание изменений контента этих источников;
- загрузка документов из интернет-источников, их разбор и унификация форматов;
- загрузка документов электронных средств массовой информации, их разбор и унификация форматов;
- заполнение форм библиографического описания документов;

- проверка корректности представления информации перед вводом в информационное хранилище;

подсистема лингвистической обработки документов должна обеспечивать выполнение следующих основных функций:

- контроль поступления документов и отслеживание основных этапов лингвистической обработки и семантического анализа документов;
- выполнение цикла первоначальной лингвистической обработки содержания документов;
- заполнение форм формализованного описания первоначальной лингвистической обработки содержания документов;
- загрузка документов и их формализованных описаний в хранилище;
- разработка новых унифицированных учетных форм, устраняющих дублирование содержащейся в них информации;
- создание и модернизация программного и лингвистического обеспечения внутренней обработки документов;

подсистема лингвистического обеспечения системы должна обеспечивать выполнение следующих основных функций:

- выполнение цикла углубленного семантического анализа содержания документов;
- заполнение форм формализованного описания семантического анализа документов;
- реализация технологий распределенной первоначальной лингвистической обработки документов;
- реализация технологий распределенного углубленного семантического анализа документов;
- хранение и модернизация программных и декларативных средств автоматической лингвистической обработки и семантического анализа содержания документов;

подсистема информационного поиска и построения аналитических отчетов документов должна обеспечивать выполнение следующих основных функций:

- реализация классических механизмов поиска по четким и нечетким критериям, включая задачу поиска с реформулированием запроса по обратной;
- реализация контекстного поиска информации по семантическим запросам, в том числе с автоматическим переводом запроса / ответа на основе интеграции с системой машинного перевода (полностью новая задача);

- выполнение статистического анализа и систематизации информации о научно-исследовательских работах и перспективных технологических разработках, ведущихся в мире, а также автоматическое формирование накапливаемой онлайн-наукометрии;
- информационная поддержка процессов оценки состояния и тенденций развития различных научных направлений деятельности, а также автоматическое формирование сводных статистических портретов основных действующих лиц (на уровне отдельных ученых, коллективов, организаций и т. п.) в области создания наукоемкой продукции;
- формирование информационного обеспечения для оценки уровня и конкурентоспособности отечественной наукоемкой продукции;
- развитие аналитических сервисов в соответствии с потребностями пользователей по обработке хранящейся в Системе информации;
- формирование дайджестов, аналитических тематических и сводных отчетов по различным областям знаний и отраслям промышленной деятельности;
- экспорт выходных документов в другие приложения — в форматы xml, xls, pdf (полностью новая задача);

подсистема обеспечения информационной безопасности должна обеспечивать выполнение следующих основных функций:

- реализация единой политики информационной безопасности в Системе;
- реализация регламентированного доступа пользователей к Системе;
- реализация дифференцированной защиты информации;
- поддержание конфиденциальности (для выделенного сегмента), доступности, достоверности, целостности и подконтрольности обрабатываемой, хранимой и передаваемой в Системе информации, включая предотвращение несанкционированного уничтожения, искажения, копирования, блокирования защищаемой информации;
- принятие организационно-технических мер по отражению сетевых атак и противодействию попыткам получения несанкционированного доступа к серверам, рабочим станциям, сетевым и телекоммуникационным устройствам как извне, так и со стороны внутренних пользователей;

коммуникационная подсистема должна обеспечивать выполнение следующих основных функций:

- обеспечение функционирования в составе единой вычислительной сети, построенной по технологии Интернет/Интранет;
- обеспечение приема запросов внешних пользователей на выборку и поиск информации в Системе;

- обеспечение выдачи результатов обработки запросов внешним пользователям;
- обеспечение приема-передачи служебных операций, связанных с выполнением административных задач;

подсистема управления функционированием Системы должна обеспечивать выполнение следующих основных функций:

- обеспечение эффективного применения и оперативности реагирования на изменение состава, состояния и условий применения Системы (ее компонентов) на основе мониторинга состояния Системы и ее составных частей;
- обеспечение возможности управления событиями — получение информации о событиях и выполнение заранее определенных действий;
- обеспечение реализации задачи сбора и анализа информации о функционировании Системы;
- обеспечение реализации задачи сбора, резервного хранения и восстановления программного и информационного обеспечения и конфигурационной информации;

подсистема эксплуатации и технической поддержки должна обеспечивать выполнение следующих основных функций:

- администрирование системы хранения и обработки документов;
- контроль внутренних процессов в Системе;
- формирование и выдача оперативных отчетов о состоянии Системы;
- выполнение регламентных и профилактических работ на оборудовании Системы;
- монтаж и инсталляция вновь поступающего оборудования в технических и офисных помещениях;
- первичная инсталляция и настройка заданных параметров общего, встроенного, прикладного и специального ПО и ввод его в эксплуатацию;
- диагностика и устранение неисправностей и дефектов в структурированной кабельной системе;
- наращивание, модернизация и реконфигурация структурированной кабельной системы;
- документирование схемы сетевых соединений и поддержание документации в актуальном состоянии;

подсистема управления и визуализации (единый портал — графический интерфейс) должна обеспечивать выполнение следующих основных функций:

- визуализация аналитических материалов, результатов статистического анализа и систематизации информации о ведущихся в мире перспективных разработках, отчеты о мониторинге событий и публикаций в заданных предметных областях;
- обеспечение возможности получения авторизованными пользователями, в том числе органами государственной власти, доступа к различной информации о зарубежных и отечественных научных исследованиях и обобщенным результатам научно-технической деятельности в мировом масштабе;
- развитие информационной среды передачи и распространения результатов научно-технической деятельности в соответствии с международными классификаторами и рубриками;
- мониторинг функционирования единого портала, сбор и анализ статистических данных об использовании хранящихся информационных ресурсов;



Рис. 1 Структурная схема Системы

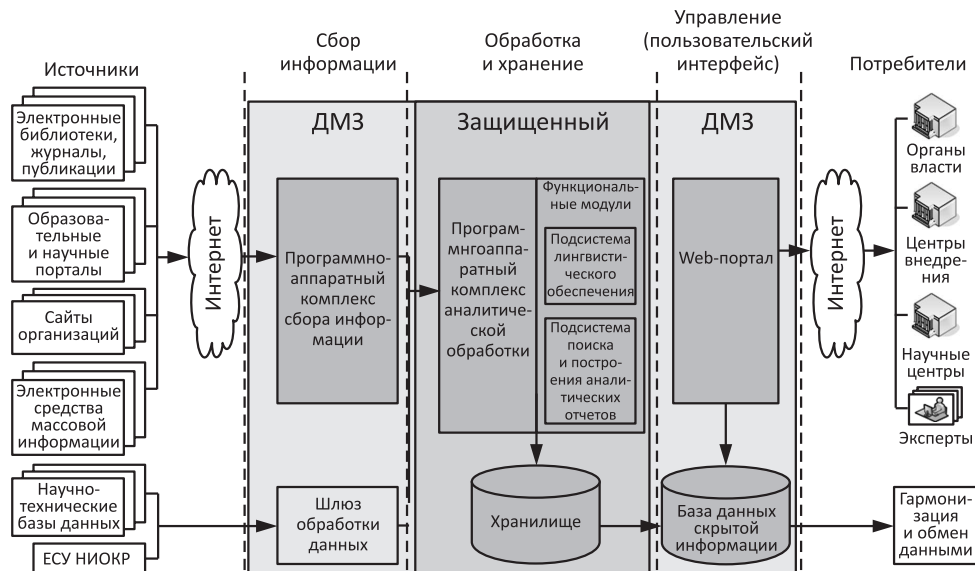


Рис. 2 Информационно-технологическая архитектура Системы

- возможность централизованного многопользовательского семантического поиска текстовой информации на русском языке по любым семантическим критериям.

Структурная схема Системы приведена на рис. 1, информационно-технологическая архитектура Системы показана на рис. 2.

Решения, закладываемые в Систему, должны быть реализованы на основе современных системно-технических принципов и подходов и обеспечивать:

- эволюционность развития;
- разработку на основе открытых стандартов, форматов, широко известных библиотек;
- высокую эксплуатационную и модернизационную пригодность;
- интеллектуальную поддержку работы пользователей (обеспечение правильности ввода информации, ее своевременности и полноты, контроль ввода по формату и соответствию новых данных уже введенным сведениям);
- облегченную масштабируемость Системы при ее развитии, росте нагрузки на ее сервисы, увеличение числа пользователей не должно требовать покупки дополнительных лицензий и (или) лицензионных отчислений, за исключением оплаты услуг по сопровождению (подписок);

- создание и использование единого корпоративного информационно-функционального пространства;
- работоспособность в гетерогенной (изменяемой) информационно-коммуникационной среде;
- доверенность ПО;
- эффективную реализацию и сопровождение полного жизненного цикла Системы.

Платформа, на которой будет разрабатываться Система, должна удовлетворять следующим базовым требованиям:

- поставка по лицензиям ПО с открытым кодом и/или свободного ПО;
- поддержка различных типов применяемого общего программного обеспечения (операционная система — ОС, система управления базами данных — СУБД, служба каталогов и т. п.) и аппаратного обеспечения;
- реализация в полном соответствии с современными стандартами разработки ПО;
- поддержка работы с промышленными СУБД;
- поставка применяемого ПО сторонних производителей с учетом вышеуказанных требований;
- полная поддержка исполнителем применяемого ПО сторонних производителей, возможно с привлечением услуг соответствующего производителя.

При опытной эксплуатации Система должна функционировать в режиме диагностики, позволяющем производить настройки и реконфигурацию, установку дополнительных элементов, переустановку ПО и т. д. После ее проведения Система может быть введена в промышленную эксплуатацию и сможет обеспечивать функционирование в штатном режиме 24 ч в сутки 365 дней в году. При этом в Системе должны быть предусмотрены программные и технические средства, обеспечивающие диагностику состояния Системы при возникновении нештатных ситуаций (сбои и отказы технических средств, ошибки в ПО), диагностические инструменты подсистемы должны позволять сохранять набор информации, необходимой для идентификации этих ситуаций, также должно обеспечиваться восстановление данных по регистрации и учету событий безопасности в базе данных до состояния на момент окончания последней нормально завершенной операции.

При создании такой Системы должна быть обеспечена возможность дальнейшего развития и модификации функционала и масштабирование при росте нагрузки на сервисы.

6 Методы решения семантических задач анализа текстов

Новизна предлагаемых решений при создании системы мониторинга и анализа НТИ состоит в комплексном использовании следующих методов семантического анализа текстов:

- методов распределения задач логико-семантического анализа естественно-языковых текстов и извлечения знаний;
- методов извлечения эксплицитно выраженных знаний и методики выявления скрытых смыслов в тексте;
- методов и средств формирования больших лексико-семантических ресурсов на основе прецедентов и машинного обучения.

В процессе создания такой системы необходимо также решить ряд сложных научно-технических задач:

- разработать концептуальную схему хранилища многоязычной системы мониторинга и анализа научно-технических публикаций;
- создать информационную технологию и автоматизированные средства оперативного мониторинга заданного информационного пространства;
- разработать многоязычную интегральную концептуальную модель широкого спектра предметных областей;
- разработать базовые процедуры генерации формализованного концептуального представления смыслового содержания разноязычных документов;
- разработать методы, алгоритмы и программы автоматического концептуального анализа разноязычных документов;
- разработать методы анализа контента различных типов источников (интернет-ресурсов, федеральных и ведомственных баз данных);
- исследовать методы оценки содержания контента различных типов источников и разработать критерии их включения в информационное пространство для оперативного мониторинга;
- исследовать и проанализировать понятийный состав разноязычных текстов в заданных предметных областях для разработки концептуальных словарей (словарей наименований понятий и смысловых ассоциативных связей между ними);
- исследовать и разработать методы автоматизированного создания, ведения и коррекции многоязычных словарей больших объемов;
- исследовать и разработать методы семантико-синтаксического и концептуального анализа разноязычных текстов и построение их формализованных описаний;

- исследовать и разработать методы отслеживания ключевой информации по профилям мониторинга на основе контекстно-тематических многоязычных словарей;
- исследовать и разработать методы обобщения многоязычных источников и выделения эксклюзивной информации;
- исследовать и разработать методы анализа, сопоставления, систематизации, обобщения и выявления тенденций развития отечественных и зарубежных научных направлений и технологических разработок.

Для решения перечисленных задач семантической обработки разноязычных документов в состав Системы предлагается включить подсистему лингвистического обеспечения, которую планируется разработать на основе базового платформенного лингвистического ПО МетаФраз путем его адаптации к функциональным задачам, программным интерфейсам и модели данных Системы. Лингвистическое ПО МетаФраз ранее уже было использовано в процессе разработки систем Мониторинг СМИ (СКЦ РОСАТОМ) и ЕСУ НИОКР (ЦИТИС).

Лингвистическое ПО МетаФраз разработано на основе теоретической концепции фразеологического концептуального анализа текстов, предложенной профессором Г. Г. Белоноговым в 1975 г. Эта концепция базируется на широкомасштабных исследованиях и адекватных представлениях о смысловой структуре текстов. Основной идеей этой концепции является обоснование использования в качестве основных единиц смысла устойчивых фразеологических и терминологических словосочетаний, обозначающих понятия, отношения между понятиями и типовые ситуации, представленные в предметной области. Эта концепция определяет принципы и методы выявления статистически обоснованного понятийного состава предметной области.

Разработанные на ее основе методы лингвистической аналогии и семантико-статистические методы автоматической обработки текстовой информации позволяют в значительной степени сократить временные издержки и снизить трудозатраты на создание комплекса программных и декларативных средств автоматического смыслового анализа документов.

Декларативные средства в соответствии с этой концепцией должны включать в свой состав интегральную словарную базу, содержащую политематические концептуальные словари и дополнительные тематические словари. Эти словари должны создаваться на основе автоматической обработки и анализа лексического состава и семантико-синтаксического анализа репрезентативных корпусов политематических и тематических текстов.

Разработанные технологии и процедуры создания концептуальных словарей обеспечивают включение в эти словари статистически обоснованного понятийного состава предметной области. Последние могут также создаваться в процессе адаптации словарной базы к лексике различных тематических областей по конкретным текстам документов. При этом также должны быть использованы средства

автоматизации, позволяющие выявить слова и словосочетания, наиболее часто используемые в текстах.

Таким образом, платформенное лингвистическое ПО МетаФраз, построенное на базе описанной теоретической концепции, обеспечивает весь технологический цикл преобразования текстового представления документа в его формализованное смысловое описание и обладает рядом характеристик, отличающих его от других аналогичных программных комплексов анализа содержания текстов, а именно:

- базируется на концепции фразеологического концептуального анализа текстов;
- оперирует мощными концептуальными политематическими и тематическими словарями;
- предоставляет пользователю единый инструментарий, позволяющий автоматизировать все этапы обработки и анализа документов;
- настраивается на обработку и анализ текстов по любой тематике с возможностью автоматизированного пополнения интегрального концептуального словаря.

7 Основное назначение системы мониторинга и анализа научно-технической информации и решаемые ею задачи

Основное назначение Системы заключается в обеспечении возможности выполнения мониторинга и анализа мирового потока НТИ по широкому спектру предметных областей и выявление тенденций развития отечественных и зарубежных научных направлений и технологических разработок. При этом Система решает следующие частные задачи:

- сбор и консолидация научных публикаций;
- смысловая обработка и систематизация;
- оперативный анализ;
- описание текущего состояния мирового научно-технологического комплекса;
- выявление тенденций развития научных направлений;
- информирование органов власти.

В результате создания такой системы могут быть реализованы следующие задачи:

- (1) получение объективной информации о тенденциях развития мирового научно-технологического комплекса, определение приоритетных научных направлений и перспективных технологических разработок, оценка научного и технологического потенциала научных коллективов и организаций, выявление дублирующихся и неэффективных направлений исследований и разработок;

- (2) создание единого хранилища данных отечественных и зарубежных научно-технических публикаций, баз данных ученых, научных коллективов и организаций;
- (3) информирование органов власти о состоянии и тенденциях научных направлений и технологических разработок:
 - регулярное оперативное оповещение;
 - аналитические отчеты по проблемно-ориентированным запросам;
 - обобщенная статистическая информация.

Выходными документами функционирования Системы могут быть следующие аналитические документы:

- аналитические отчеты о тенденциях развития мирового научно-технического комплекса;
- дайджесты:
 - по тематикам разработок и исследований;
 - по коллективам и организациям;
 - по приоритетным информационным источникам;
- информация о возникающих перспективных технологиях и научных разработках;
- динамика научной активности:
 - изменение текущих направлений исследований;
 - всплески активности (приоритетные разработки на текущий год);
- сопоставление данных о зарубежных и отечественных научных исследованиях;
- аналитические документы по запросам пользователей;
- отчеты о мониторинге событий и публикаций в заданных областях.

8 Заключение

Для успешного выполнения работ по созданию системы мониторинга и анализа мирового потока НТИ необходимы теоретические предпосылки и практический задел для разработки систем автоматического интеллектуально анализа разноязычных текстов. Такими предпосылками являются:

- адекватные теоретические представления о семантико-синтаксической структуре естественных языков;
- наличие у авторов проекта базовых процедур автоматического семантико-синтаксического фразеологического анализа и синтеза текстов;

- наличие процедур автоматического составления словарей наименований понятий и автоматического установления смысловых связей между ними;
- наличие процедур автоматического составления различного рода словарей по текстам документов.

В процессе исследований и разработок по фразеологическому машинному переводу авторами проекта был получен ряд фундаментальных результатов, которые могут быть успешно использованы при создании систем автоматической смысловой обработки текстовой информации: автоматической классификации и индексирования документов, их реферирования и анализа смысловой структуры, мультязычного поиска информации в сети Интернет, информационно-аналитических систем различного типа.

Разработанные авторами методы, алгоритмы и технологические решения внедрены в ряде промышленных систем, среди которых:

- автоматизированная система фразеологического машинного перевода «RETRANS 2005» (двуязычная система перевода) (2005–2006 гг.);
- многопользовательская сетевая автоматизированная система фразеологического машинного перевода «RETRANS 2008» (двуязычная система перевода) (2007–2008 гг.);
- многопользовательская сетевая автоматизированная система фразеологического машинного перевода «МетаФраз: Система перевода R8» (трехязычная система перевода) (2008–2010 гг.);
- многопользовательская сетевая автоматизированная система фразеологического машинного перевода «МетаФраз: Система перевода R10» (в различных версиях и комплектациях) (2010–2011 гг.);
- лингвистический процессор и средства интеграции для разработчиков приложений «МетаФраз: Лингвистический интеграционный комплект R10» (многофункциональный комплекс кроссплатформенной интеграции лингвистических средств смысловой обработки текстов и фразеологического машинного перевода) (2010–2014 гг.);
- подсистема лингвистического обеспечения системы ЕСУ НИОКР (2011–2012 гг.);
- смысловая поисково-аналитическая система авиационной промышленности «СПАС-Авиа» (2010–2014 гг.);
- система мониторинга СМИ СКЦ Росатома (2010–2014 гг.).

Литература

1. Богданов Ю. М., Пошатаев О. Н., Старовойтов А. В., Хорошилов А. А. Принципы создания высокопроизводительных систем обработки и анализа текстовой информации // Информатизация и связь, 2013. № 3. С. 74–81.

2. *Поштаев О. Н., Хорошилов А. А.* Методы анализа текстов в технологиях «Big Data» // Электронные библиотеки: перспективные методы и технологии, электронные коллекции: Тр. XV Всеросс. научной конф. RCDL'2013. — Ярославль: ЯрГУ им. П. Г. Демидова, 2013. С. 30–38.
3. *Белоногов Г. Г., Гиляревский Р. С., Хорошилов А. А.* Проблемы автоматической смысловой обработки текстовой информации // Научно-техническая информация. Сер. 2: Информационные процессы и системы, 2012. № 11. С. 24–28.
4. *Белоногов Г. Г., Гиляревский Р. С., Хорошилов А. А., Хорошилов-мл. А. А.* Автоматическое распознавание смысловой близости документов // Научно-техническая информация. Сер. 2: Информационные процессы и системы, 2011. № 7. С. 15–22.
5. *Белоногов Г. Г., Хорошилов Ал-др А., Хорошилов Ал-ей А.* Автоматизация составления англо-русских двуязычных фразеологических словарей по массивам двуязычных текстов (бинлинг) // Научно-техническая информация. Сер. 2: Информационные процессы и системы, 2010. № 5. С. 1–8.
6. *Khoroshilov Alexander A., Khoroshilov Alexei A.* Set-phrase machine translation based on multilingual dictionaries // WORLDCOMP'11: ICAI'11 Proceedings. — Las Vegas, Nevada, USA. — USA: CSREA Press, 2011. P. 888–891.
7. *Белоногов Г. Г., Калинин Ю. П., Хорошилов А. А.* Компьютерная лингвистика и перспективные информационные технологии. Теория и практика построения систем автоматической обработки текстовой информации. — М.: Русский мир, 2004. 247 с.
8. *Белоногов Г. Г., Хорошилов Ал-др А., Хорошилов Ал-ей А.* Единицы языка и речи в системах автоматической обработки текстовой информации // Научно-техническая информация. Сер. 2: Информационные процессы и системы, 2005. № 11. С. 21–29.
9. *Белоногов Г. Г.* Теоретические проблемы информатики. Т. 2: Семантические проблемы информатики. — М.: РЭА им. Г. В. Плеханова, 2008. 238 с.
10. *Белоногов Г. Г., Хорошилов Ал-др А., Хорошилов Ал-ей А. и др.* Система фразеологического машинного перевода МетаФраз. Современное состояние // Перевод: традиции и современные технологии. — М.: ВЦП, 2009. С. 18–26.
11. *Хорошилов Ал-др А., Хорошилов Ал-ей А., Никитин Ю. В. и др.* Автоматизированная словарная служба МетаФраз (MetaFraz ADS): Программный комплекс. Свидетельство об официальной регистрации программ для ЭВМ № 2014612497 от 26.02.2014.
12. *Хорошилов Ал-др А., Хорошилов Ал-ей А., Никитин Ю. В. и др.* Настройки системы перевода МетаФраз R10 (MF Lingware Add-in R10). Свидетельство об официальной регистрации программ для ЭВМ № 2014662528 от 02.12.2014.
13. *Хорошилов Ал-др А., Хорошилов Ал-ей А., Никитин Ю. В. и др.* Лингвистический интеграционный комплект МетаФраз R10 (MF Lingware Integration Kit R10). Свидетельство об официальной регистрации программ для ЭВМ № 2014662529 от 02.12.2014.
14. *Хорошилов Ал-др А., Хорошилов Ал-ей А., Никитин Ю. В. и др.* Сервер лингвистического ПО МетаФраз R10 (MF Lingware Server R10). Свидетельство об официальной регистрации программ для ЭВМ № 2014662743 от 08.12.2014.
15. *Хорошилов Ал-др А., Хорошилов Ал-ей А., Никитин Ю. В. и др.* Лингвистический комплекс МетаФраз R10 (MF Lingware Complexn R10). Свидетельство об официальной регистрации программ для ЭВМ № 2014663079 от 15.12.2014.

16. *Хорошилов Ал-др А., Хорошилов Ал-ей А., Никитин Ю. В. и др.* Система семантической обработки текстов МетаФраз R10 (MF Text Analyst R10). Свидетельство об официальной регистрации программ для ЭВМ № 2014663081 от 15.12.2014.
17. *Хорошилов Ал-др А., Хорошилов Ал-ей А., Никитин Ю. В. и др.* Система перевода МетаФраз R10 (MF Translation System R10). Свидетельство об официальной регистрации программ для ЭВМ № 2014663082 от 15.12.2014.
18. *Поштаев О. Н., Прохоров С. Н., Старовойтов А. В., Хорошилов А. А.* Методы автоматизированного составления и ведения словарей // Информатизация и связь, 2013. № 3. С. 91–97.
19. *Хорошилов Ал-др А., Никитин Ю. В., Хорошилов Ал-ей А., Будзко В. И.* Автоматическое создание формализованного представления смыслового содержания неструктурированных текстовых сообщений СМИ и социальных сетей // Системы высокой доступности, 2014. Т. 10. № 3. С. 36–51.
20. *Белогозов Г. Г., Гиляревский Р. С., Селедков С. Н., Хорошилов А. А.* О путях повышения качества поиска текстовой информации в системе Интернет // Научно-техническая информация. Сер. 2: Информационные процессы и системы, 2012. № 8. С. 15–22.
21. *Белогозов Г. Г., Гиляревский Р. С., Хорошилов Ал-др А., Хорошилов Ал-ей А.* Развитие систем автоматической обработки текстовой информации // Нейрокомпьютеры: разработка, применение, 2010. № 8. С. 4–13.

Поступила в редакцию 29.07.15

PRINCIPLES OF CREATION OF MONITORING SYSTEM AND ANALYSIS OF WORLD STREAM OF SCIENTIFIC AND TECHNICAL INFORMATION

Yu. P. Kalinin, Al-der A. Khoroshilov, and Al-ey A. Khoroshilov

Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilova Str., Moscow 119333, Russian Federation

Abstract: At the present stage of development of scientific and technological potential of the Russian Federation, it is necessary to improve the efficiency of research and development work. Such decisions should be based on using the objective analytical data on development trends of the world scientific and technological complex. These data can be obtained through the complex analysis of the global flow of scientific and technical information. The sources of such information may be foreign and domestic, branches and departmental portals, sites of scientific institutions, journals, scientific conferences and communities, electronic mass media, and other Internet sources of scientific and technical information. In the process of semantic analysis of multilingual heterogeneous distributed information sources, new knowledge can be obtained and the priority areas of activity of domestic and foreign research teams can be determined. The

new principles for development of the system for monitoring and analysis of the global flow of scientific and technical information are proposed. These principles are based on the modern perceptions of conceptual text structure. These perceptions are based on the modern conception of phraseological conceptual analysis of texts. The methods and tools for formalization of semantic structure of the text information and adaptation methods for declarative means for the procedures of linguistic processing and semantic analysis of the text information are considered in detail. The architecture of the system and the list of functions of all its subsystems are given.

Keywords: automated text processing; semantic analysis; formal description of text; semantic structure; data extraction; linguistic software

DOI: 10.14357/08696527160110

References

1. Bogdanov, Yu. M., O. N. Poshataev, A. V. Starovoytov, and A. A. Khoroshilov. 2013. Printsipy sozdaniya vysokoproizvoditel'nykh sistem obrabotki i analiza tekstovoy informatsii [Principles of creation of high-performance treatment systems and text information analysis]. *Informatizatsiya i Soyaz'* [Informatization and Communication] 3:74–81.
2. Poshataev, O. N., and A. A. Khoroshilov. 2013. Metody analiza tekstov v tekhnologiyakh “Big Data” [Methods of text analysis in “Big Data” technologies]. *Elektronnye Biblioteki: Perspektivnye Metody i Tekhnologii, Elektronnye Kollekt-sii: Tr. XV Vseross. nauchnoy konf. RCDL'2013* [E-libraries: Prospective Methods and Technology, Electronic Collections: 15th All-Russian Scientific Conference of REDL'2013 Proceedings]. Yaroslavl: Yaroslavl State University. 30–38.
3. Belonogov, G. G., R. S. Gilyarevskiy, and A. A. Khoroshilov. 2012. Problemy avtomaticheskoy smyslovoy obrabotki tekstovoy informatsii [Problems of automatic semantic text information processing]. *Nauchno-Tekhnicheskaya Informatsiya. Ser. 2: Informatsionnye protsessy i sistemy* [Scientific and Technical Information. Ser. 2: Information processes and systems] 2(11):24–28.
4. Belonogov, G. G., R. S. Gilyarevskiy, A. A. Khoroshilov, and A. A. Khoroshilov, Jr. 2011. Automatic recognition of semantic similarity of documents. *Nauchno-Tekhnicheskaya Informatsiya. Ser. 2: Informatsionnye protsessy i sistemy* [Scientific and Technical Information. Ser. 2: Information processes and systems] 2(7):15–22.
5. Belonogov, G. G., Al-dr A. Khoroshilov, and Al-ey A. Khoroshilov. 2010. Avtomatizatsiya sostavleniya anglo-russkikh dvuyazychnykh frazeologicheskikh slovarey po massivam dvuyazychnykh tekstov (bilingv) [Constructing automation of English-Russian bilingual phraseological vocabularies on arrays of bilingual texts (bilingv)]. *Nauchno-Tekhnicheskaya Informatsiya. Ser. 2: Informatsionnye protsessy i sistemy* [Scientific and Technical Information. Ser. 2: Information processes and systems] 2(5):1–8.
6. Khoroshilov, Alexander A., and Alexei A. Khoroshilov. 2011. Set-phrase machine translation based on multilingual dictionaries. *WORLDCOMP'11: ICAI'11 Proceedings*. Las Vegas, NV, USA: CSREA Press. 888–891.

7. Belonogov, G. G., Yu. P. Kalinin, and A. A. Khoroshilov. 2004. *Komp'uternaya lingvistika i perspektivnye informatsionnye tekhnologii. Teoriya i praktika postroeniya sistem avtomaticheskoy obrabotki tekstovoy informatsii* [Computational linguistics and perspective information technologies. Theory and practice of construction of automatic text processing systems]. Moscow: Russkiy Mir. 247 p.
8. Belonogov, G. G., Al-dr A. Khoroshilov., and Al-ey A. Khoroshilov. 2005. Edinitsy yazyka i rechi v sistemakh avtomaticheskoy obrabotki tekstovoy informatsii [Language and speech units in automatic text processing systems]. *Nauchno-Tekhnicheskaya Informatsiya. Ser. 2: Informatsionnye protsessy i sistemy* [Scientific and Technical Information. Ser. 2: Information processes and systems] 2(11):21–29.
9. Belonogov, G. G. 2008. *Teoreticheskie problemy informatiki. T. 2. Semanticheskie problemy informatiki* [Theoretical problems of informatics. Vol. 2. Semantic problems of informatics]. Moscow: G. V. Plekhanov REA. 2:238.
10. Belonogov, G. G., Al-dr A. Khoroshilov, Al-ey A. Khoroshilov, *et al.* 2009. Sistema frazeologicheskogo mashinnogo perevoda MetaFraz. Sovremennoe sostoyanie [System of phraseological machine translation of MetaFraz. Present state]. *Perevod: Traditsii i sovremennye tekhnologii* [Translation: Traditions and modern technologies]. Moscow: VCP. 18–26.
11. Khoroshilov, Al-dr A., Al-ey A. Khoroshilov, Yu. V. Nikitin, *et al.* 26.02.2014. Avtomatizirovannaya slovarnaya sluzhba MetaFraz (MetaFraz ADS): Programmnyy kompleks [Computer-aided dictionary service of MetaFraz (MetaFraz ADS): Program complex]. Patent RF No. 2014612497.
12. Khoroshilov, Al-dr A., Al-ey A. Khoroshilov, Yu. V. Nikitin, *et al.* 02.12.2014. Nadstroyki sistemy perevoda MetaFraz R10 (MF Lingware Add-in R10) [Suspensions of translating system of MetaFraz R10 (MF Lingware Add-in R10)]. Patent RF No. 2014662528.
13. Khoroshilov, Al-dr A., Al-ey A. Khoroshilov, Yu. V. Nikitin, *et al.* 02.12.2014. Lingvisticheskiy integratsionnyy komplekt MetaFraz R10 (MF Lingware Integration Kit R10) [Linguistic integrative kit of MetaFraz R10 (MF Lingware Integration Kit R10)]. Patent RF No. 2014662529.
14. Khoroshilov, Al-dr A., Al-ey A. Khoroshilov, Yu. V. Nikitin, *et al.* 08.12.2014. Server lingvisticheskogo PO MetaFraz R10 (MF Lingware Server R10) [Server of linguistic software of MetaFraz R10 (MF Lingware Server R10)]. Patent RF No. 2014662743.
15. Khoroshilov, Al-dr A., Al-ey A. Khoroshilov, Yu. V. Nikitin, *et al.* 15.12.2014. Lingvisticheskiy kompleks MetaFraz R10 (MF Lingware Complex R10) [Linguistic complex of MetaFraz R10 (MF Lingware Complex R10)]. Patent RF No. 2014663079.
16. Khoroshilov, Al-dr A., Al-ey A. Khoroshilov, Yu. V. Nikitin, *et al.* 15.12.2014. Sistema semanticheskoy obrabotki tekstov MetaFraz R10 (MF Text Analyst R10) [System of semantic text processing of MetaFraz R10 (MF Text Analyst R10)]. Patent RF No. 2014663081.
17. Khoroshilov, Al-dr A., Al-ey A. Khoroshilov, Yu. V. Nikitin, *et al.* 15.12.2014. Sistema perevoda MetaFraz R10 (MF Translation System R10) [Translating system of MetaFraz R10 (MF Translation System R10)]. Patent RF No. 2014663082.
18. Poshataev, O. N., S. N. Prokhorov, A. V. Starovoytov, and A. A. Khoroshilov. 2013. Metody avtomatizirovannogo sostavleniya i vedeniya slovarey [Methods of computer-aided compilation and maintenance of dictionaries]. *Informatizatsiya i Svyaz'* [Informatization and Communication] 3:91–97.

19. Khoroshilov, Al-dr A., Yu. V. Nikitin, Al-ey A. Khoroshilov, and V. I. Budsko. 2014. Avtomaticheskoe sozdanie formalizovannogo predstavleniya smyslovogo sodержaniya nestrukturirovannykh tekstovykh soobshcheniy SMI i sotsial'nykh setey [Automatic creation of formalized representation of semantic content of unstructured textual media advisory and social networks]. *Sistemy Vysokoy Dostupnosti* [Systems of High Availability] 10(3):36–51.
20. Belonogov, G. G., R. S. Gilyarevskiy, S.N. Seledkov, and A. A. Khoroshilov. 2012. O putyakh povysheniya kachestva poiska tekstovoy informatsii v sisteme Internet [About ways of increasing of quality of retrieval of text information in internet]. *Nauchno-Tekhnicheskaya Informatsiya. Ser. 2: Informatsionnye protsessy i sistemy* [Scientific and Technical Information. Ser. 2: Information processes and systems] 2(8):15–22.
21. Belonogov, G. G., R. S. Gilyarevskiy, Al-dr A. Khoroshilov, and Al-ey A. Khoroshilov. 2010. Razvitie sistem avtomaticheskoy obrabotki tekstovoy informatsii [Development of automatic text processing systems]. *Neyrokomp'yutery: Razrabotka, Primenenie* [Neurocomputers: Development, Application] 8:4–13.

Received July 29, 2015

Contributors

Kalinin Yury P. (b. 1948) — Doctor of Science in technology, professor, head of laboratory, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333; upkalinin@mail.ru

Khoroshilov Alexander A. (b. 1952) — Doctor of Science in technology, professor, leading scientist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333; khoroshilov@mail.ru

Khoroshilov Alexei A. (b. 1958) — Candidate of Science (PhD) in technology, scientist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333; alex_khoroshilov@mail.ru

МЕТОДИКА РАСЧЕТА НАДЕЖНОСТИ МЕЖДУНАРОДНОГО ТЕРМОЯДЕРНОГО ЭКСПЕРИМЕНТАЛЬНОГО РЕАКТОРА И ОПТИМИЗАЦИЯ ДЕЙСТВИЙ ПО СНИЖЕНИЮ РИСКОВ

*Г. М. Коновалов*¹

Аннотация: Статья посвящена одной из наиболее существенных проблем при создании международного термоядерного экспериментального реактора (ИТЭР) — обеспечению надежности. Описаны принципы, которыми руководствуются в настоящее время при анализе надежности реактора и его основных подсистем. В основу концепции повышения надежности положен так называемый RAMI-анализ (от англ. reliability, availability, maintainability, inspectability), охватывающий основные аспекты проблемы: увеличение коэффициентов надежности и готовности, возможность контроля, технического обслуживания, модернизации и замены. На примерах реальных подсистем реактора изложены основные этапы RAMI-анализа: функциональный анализ, определение коэффициентов надежности и готовности, анализ потенциальных отказов и степени их воздействия, составление списка смягчающих действий. Последний из перечисленных этапов связан с оптимизацией возможных действий по уменьшению вероятности возникновения риска и снижению времени на восстановление работоспособности. Предложена математическая модель, позволяющая автоматизировать процесс выбора наиболее эффективных и наименее затратных воздействий, сопряженный с обработкой больших массивов информации.

Ключевые слова: ИТЭР; токамак; анализ надежности больших систем; действия по смягчению рисков и повышению готовности; математические модели оптимизации надежности

DOI: 10.14357/08696527160111

1 Введение

В настоящее время во Франции силами семи участников (ЕС, Япония, США, Россия, Китай, Индия, Южная Корея) под общим руководством международной Организации ИТЭР (International Thermonuclear Experimental Reactor) ведется сооружение термоядерного реактора [1, 2]. Предполагается, что экспериментальный термоядерный реактор будет работать по три восьмичасовые смены непрерывно циклами по 16 мес., прерывающимися на восьмимесячные плановые остановки для обслуживания и профилактики. Разработка и строительство подобного рода установок являются крайне сложной задачей, сопряженной с огромными техническими рисками как на стадии проектирования и изготовления, так и на

¹Частное учреждение Государственной корпорации по атомной энергии «Росатом» «Проектный центр ИТЭР», g.konvalov@iterrf.ru

стадии эксплуатации. Кроме того, ИТЭР является ядерным объектом, и к нему применяются повышенные требования в плане надежности и безопасности.

Для увеличения надежности ИТЭР был разработан комплекс мер, называемый RAMI-анализом [3]. RAMI — аббревиатура от reliability (коэффициент надежности — вероятность того, что система будет выполнять свои функции без отказов на протяжении заданного интервала времени), availability (коэффициент готовности — вероятность того, что система будет в состоянии выполнять свои функции в данный момент времени), maintainability (ремонтпригодность — возможность ремонта, обслуживания и модернизации), inspectability (инспектируемость — возможность осуществления контроля).

Основная задача RAMI-анализа — определить возможные технические риски и предложить возможные методы и средства для снижения этих рисков и, как следствие, добиться максимального времени безотказной работы установки. RAMI-анализ проводится на стадии проектирования и начинается с эскизного анализа. Уже на самом начальном этапе можно предложить действия, направленные, например, на изменение конструкции или элементной базы, подготовку набора специфических, не предусмотренных промышленными стандартами тестов, составление списка запасных частей и т. д. с конечной целью повысить надежность системы и иногда даже минимизировать затраты на изготовление.

Реактор ИТЭР состоит из нескольких десятков крупных подсистем (вакуумная камера, магнитная система, система подачи энергии, диагностический комплекс, система управления и др.), для которых сформулированы требования по надежности. Процесс RAMI-анализа применяется ко всем подсистемам ИТЭР и состоит из четырех этапов.

1. Функциональный анализ — создание полной функциональной схемы подсистемы, описывающей все функции, начиная от главной и заканчивая функциями нижнего уровня и оборудованием, которое обеспечивает их выполнение.
2. Определение коэффициентов надежности и готовности подсистемы на основе данных о надежности используемого оборудования.
3. Анализ потенциальных отказов и степени их воздействия на подсистему — составление перечня возможных отказов оборудования, причин этих отказов, определение частоты возникновения и времени, требуемого на восстановление работоспособности системы, а также степени риска.
4. Составление списка смягчающих действий — определение возможных действий для уменьшения вероятности отказа и снижения времени на восстановление работоспособности.

Этапы 2–4 могут повторяться неоднократно до тех пор, пока коэффициенты готовности и надежности не достигнут требуемого заданного значения.

Ввиду уникальности сооружения многие аспекты обеспечения надежности требуют оригинальных решений. В частности, это относится к выбору воздействий, которые могут быть применены к различным элементам с целью

уменьшения рисков нежелательных отказов оборудования и повышению тем самым показателей надежности подсистем. Воздействия являются, вообще говоря, затратными, поэтому встает вопрос об их оптимизации.

Дальнейшее содержание статьи включает более подробное описание этапов РАМІ-анализа (разд. 2), а также постановку задачи оптимального выбора смягчающих воздействий, сопровождаемую примером ее решения для конкретной подсистемы (разд. 3).

2 Описание процесса проведения РАМІ-анализа

2.1 Функциональный анализ

Основной задачей функционального анализа является определение иерархии функций подсистемы, начиная от главной функции подсистемы и заканчивая функциями нижнего уровня и компонентами, обеспечивающими их выполнение, а также взаимосвязей между ними. Методика проведения функционального анализа базируется на так называемом IDEF0-подходе (IDEF0 — integration definition function language 0), который основан на методологии структурного анализа и проектирования (SADT — structured analysis and design technique) [4, 5].

В качестве примера на рис. 1 показана упрощенная функциональная IDEF0-схема одной из диагностических систем токамака¹ ИТЭР — диверторного монитора нейтронного потока (ДМНП).

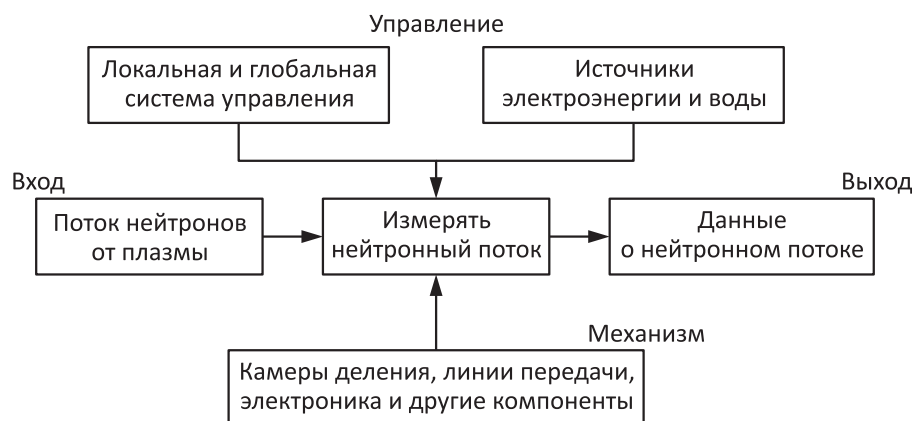


Рис. 1 Упрощенная IDEF0-модель ДМНП

¹Токамак (тороидальная камера с магнитными катушками) — тороидальная установка для магнитного удержания плазмы с целью достижения условий, необходимых для протекания управляемого термоядерного синтеза. Реактор ИТЭР по своей конструкции относится к токамакам.

Основная задача ДМНП — проводить измерения общего нейтронного потока от плазмы (на рис. 1 — *выход*). Сама система ДМНП является подсистемой общей системы диагностики токамака, которая помимо ДМНП включает в себя ряд других диагностик (анализатор атомов перезарядки, систему рефлектометрии со стороны сильного магнитного поля и др.). Основная функция ДМНП будет выполняться только в том случае, когда существует сам нейтронный поток от плазмы (доступен *вход*), работают системы управления и подачи электроэнергии и воды (осуществляется управление) и доступны *механизмы*, которыми выполняется собственно диагностика нейтронного потока.

Таким образом, полная функциональная схема дает представление о взаимодействии отдельных узлов и компонентов системы, что является базисом для последующих шагов РАМІ-анализа.

2.2 Определение коэффициентов надежности и готовности

В рамках подхода Организации ИТЭР для определения коэффициентов надежности и готовности используется диаграмма надежности (Reliability Block Diagrams). Для построения диаграммы надежности в качестве базиса используется функциональная схема подсистемы, на основании которой определяются зависимости между блоками диаграммы. Диаграммы надежности, описывающие многоуровневую иерархию функций, позволяют определить коэффициенты готовности и надежности для главной функции подсистемы, используя в качестве

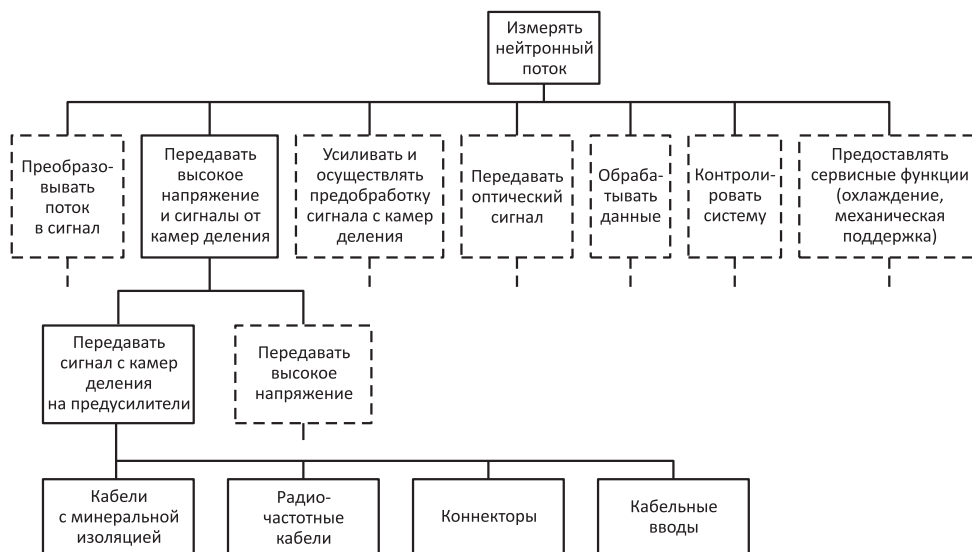


Рис. 2 Функциональная диаграмма для подфункции «Передать высокое напряжение и сигналы от камер деления» функции «Измерять нейтронный поток» системы ДМНП

входной информации только данные о компонентах, обеспечивающих функции нижнего уровня (базовые функции) (рис. 2).

Входной информацией для отдельных компонентов являются параметры надежности (среднее время наработки на отказ, вероятность отказа) и параметры обслуживания (среднее время на восстановление работоспособности). Эти параметры могут быть получены из различных источников, таких как спецификации производителя оборудования, базы данных, промышленные стандарты, предыдущий опыт изготовителя оборудования при создании аналогичных систем, предположения экспертов. В некоторых случаях доступные данные могут не отвечать специфическим условиям, в которых компонентам предстоит функционировать в ИТЭР, и необходима экспертная оценка для корректировки данных.

При построении диаграмм надежности необходимо учитывать такие аспекты, как параллельное подключение компонент, дублирование, возможность выхода из строя какого-либо компонента без нарушения функционирования основной функции и т. п. На рис. 3 приведен пример диаграммы надежности для базовой функции «Передавать сигнал с камер деления на предусилители», которая необходима для обеспечения функции «Передавать высокое напряжение и сигналы от камер деления» (см. также рис. 2).

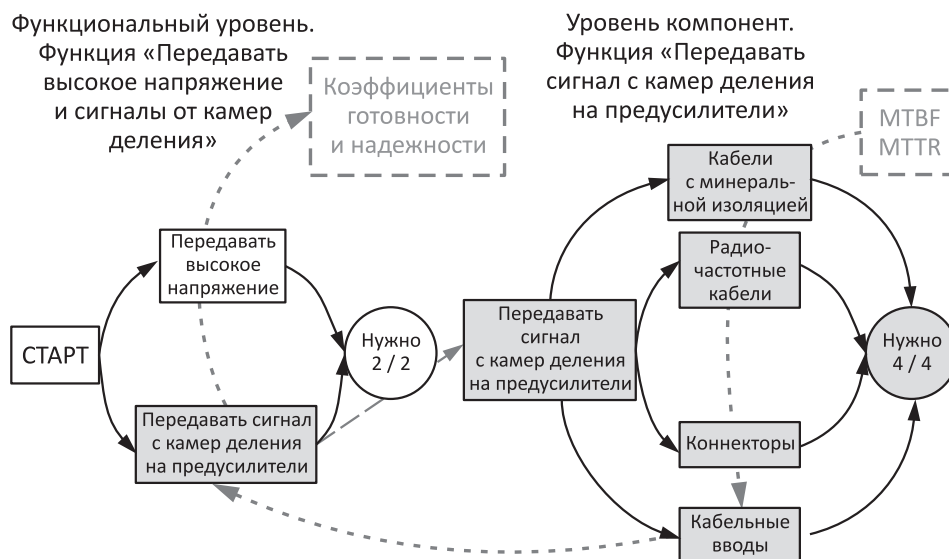


Рис. 3 Диаграмма надежности на примере базовой функции «Передавать сигнал с камер деления на предусилители» (MTBF — mean time between failures — среднее время наработки на отказ; MTTR — mean time to repair — среднее время восстановления работоспособности)

В реальности ДМНП представляет собой три независимых модуля, расположенных в разных местах токамака, причем каждый модуль состоит из трех камер деления с различной чувствительностью. Таким образом в диаграмме для функции «Передавать высокое напряжение и сигналы от камер деления», которая является функцией на уровень выше, чем функция «Передавать сигнал с камеры деления на предусилители» (см. рис. 2), описанная выше диаграмма должна быть учтена трижды. Аналогично трижды должна быть учтена функция «Передавать высокое напряжение и сигналы от камер деления» при построении диаграммы для функции верхнего уровня «Измерять нейтронный поток».

Таким образом, зная функциональные зависимости и потенциальные отказы компонентов, обеспечивающих выполнение базовых функций, можно оценить вероятность отказа основной функции подсистемы. Далее, зная вероятности отказов для основных функций подсистем, можно оценить надежность всей системы в целом.

Для расчета коэффициентов надежности и готовности, как правило, используется специализированное программное обеспечение (например, программа BlockSim компании ReliaSoft [6]), но методология расчета представляет отдельный интерес и ее обсуждение выходит за рамки данной статьи.

Аналогичные диаграммы, построенные для всех подсистем, могут быть объединены для определения коэффициентов надежности и готовности всей установки ИТЭР.

2.3 Анализ потенциальных отказов и степени их воздействия

Параллельно с построением диаграмм надежности в рамках подхода Организации ИТЭР проводится анализ потенциальных отказов и степени их воздействия для определения возможных функциональных отказов и оценки степени риска этих отказов. В совокупности эти действия носят название FMEC-анализ (Failure Modes, Effects, and Criticality analysis [3, 7]).

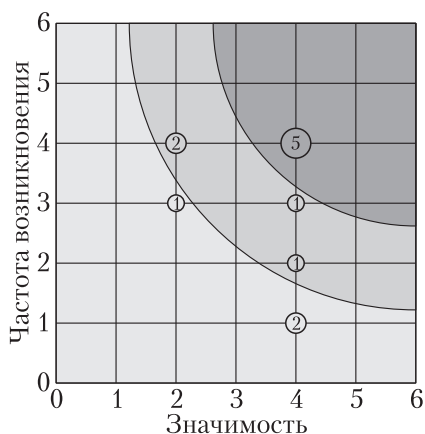
FMEC-анализ проводится на основе анализа функциональности и параметров надежности и обслуживания с соблюдением следующих основных принципов:

- необходимо описать все возможные отказы, причины, их вызывающие, и последствия этих отказов;
- отказы и последствия должны быть описаны количественно в терминах «частоты возникновения» (occurrence) и «значимости» (severity) в соответствии со шкалой, приведенной в табл. 1. Данная шкала должна использоваться для всех подсистем ИТЭР;
- степень риска (criticality) определяется как произведение частоты возникновения на значимость.

Согласно классификации, принятой в Организации ИТЭР, в зависимости от степени риска его можно отнести к существенному (major), среднему (medium) и незначительному (minor) (рис. 4).

Таблица 1 Шкала ИТЭР для количественной оценки частоты возникновения и значимости риска

Частота возникновения	Описание	Интенсивность отказов
1	Очень низкий	Менее $5 \cdot 10^{-4}$ в год (реже, чем 1 раз в 2000 лет)
2	Низкий	От $5 \cdot 10^{-4}$ до $5 \cdot 10^{-3}$ в год (реже, чем 1 раз в 200 лет)
3	Умеренный	От $5 \cdot 10^{-3}$ до $5 \cdot 10^{-2}$ в год (реже, чем 1 раз в 20 лет)
4	Высокий	От $5 \cdot 10^{-2}$ до $5 \cdot 10^{-1}$ в год (реже, чем 1 раз в 2 года)
5	Очень высокий	От $5 \cdot 10^{-1}$ до 5 в год (реже, чем 5 раз в год)
6	Критичный	Более 5 раз в год
Значимость	Описание	Время на восстановление работоспособности
1	Слабый	Менее 1 ч
2	Умеренный	От 1 ч до 1 дня
3	Серьезный	От 1 дня до 1 нед.
4	Опасный	От 1 нед. до 2 мес.
5	Критичный	От 2 мес. до 1 года
6	Катастрофический	Более 1 года

**Рис. 4** Диаграмма степеней рисков для ДМНП: темно-серая зона — существенные риски; серая — средние, светло-серая зона — незначительные риски

На примере ДМНП, представленном на рис. 4, можно видеть, что для данной подсистемы было определено 5 существенных рисков (степень риска больше 13). Для существенных рисков необходимо внедрение смягчающих действий (см. следующий раздел) для уменьшения вероятности возникновения рисков или уменьшения времени на восстановление работоспособности.

По правилам Организации ИТЭР для рисков средней степени смягчающие действия желательны (но не обязательны). Для незначительных рисков внедрение смягчающих действий полностью осуществляется по желанию разработчика подсистемы (например, для повышения коэффициента готовности всей подсистемы (см. разд. 3)).

2.4 Смягчающие действия

Для уменьшения степени рисков, определенных в процессе FMES-анализа, а также для увеличения коэффициентов надежности и готовности подсистемы должны применяться смягчающие действия. Они могут влиять как на частоту возникновения риска (предотвращающие), так и на значимость риска (защитные) или одновременно на оба параметра, тем самым уменьшая степень риска. Примеры некоторых возможных действий даны в табл. 2.

В случае ДМНП было предложено предусмотреть наличие запасных частей для снижения времени на замену вышедших из строя компонентов, что привело к снижению степени риска.

Таблица 2 Примеры смягчающих действий

Этап	Предотвращающие (уменьшающие частоту возникновения риска)	Защитные (уменьшающие значимость риска)
Проектирование	Предусмотреть дублирование потенциально ненадежного компонента	Предусмотреть невозможность каскада поломок оборудования из-за выхода из строя потенциально ненадежного компонента
Тестирование	Провести тестирование для определения надежности компонента	Проводить тесты для определения необходимости обслуживания
Функционирование	Отключать в процессе работы компоненты с целью предотвращения повреждения	Внедрить специальную рабочую процедуру для предотвращения возможного выхода из строя
Обслуживание	Увеличить частоту инспекций и периодических обслуживаний	Иметь в наличии запасные части для уменьшения времени ремонта

2.5 Повторный анализ после внедрения смягчающих действий

После внедрения смягчающих действий для каждой подсистемы должен проводиться повторный FMES-анализ и расчет коэффициентов надежности и готовности.

Критериями успешности RAMI-анализа служат:

- отсутствие существенных рисков;
- значения коэффициентов надежности и готовности не ниже требуемых.

На рис. 5 представлена диаграмма степеней рисков для ДМНП после внедрения смягчающих действий.

Как видно из рис. 5, первый критерий частично выполнен. Однако даже после внедрения смягчающих действий все равно осталось три существенных риска, что требует дальнейшей конструкторской доработки системы.

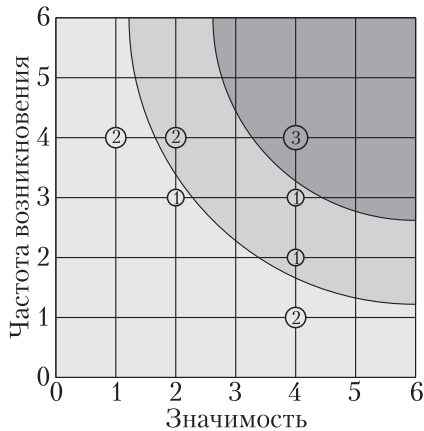


Рис. 5 Диаграмма степеней рисков после внедрения смягчающих действий для ДМНП: темно-серая зона — существенные риски; серая — средние, светло-серая зона — незначительные риски

проблемой при проведении РАМІ-анализа. Очевидно, что можно воздействовать абсолютно на все риски, но это влечет за собой очевидные технологические и экономические трудности. С другой стороны, можно внедрять действия по одному и каждый раз пересчитывать коэффициенты готовности и надежности до тех пор, пока они не станут соответствовать заданным ограничениям, но этот путь трудоемкий и требует больших временных затрат.

Таким образом, задача оптимизации выбора воздействий представляется в настоящее время весьма актуальной и значимой [8]. Ниже предлагается путь ее решения, опирающийся на математическое моделирование.

Рассмотрим систему, состоящую из n ненадежных независимых элементов. Нарботка на отказ i -го элемента задается плотностью распределения f_i . Работоспособность системы зависит от ее состояния, которое характеризуется вектором $s = (s_1, \dots, s_n)$, где $s_i = 0$, если i -й элемент исправен, и $s_i = 1$, если i -й элемент неисправен. Множество всех состояний обозначается S . Оно, очевидно, распадается на непересекающиеся подмножества S_0, \dots, S_n , где S_k состоит из тех и только тех векторов s , в которых ровно k единиц. Задана функция $H(s)$, $s \in S$, принимающая значения 0 (система исправна) или 1 (отказ системы). Функция H называется структурной функцией.

Пусть в начальный момент времени $t = 0$ все элементы системы исправны, так же как и сама система. Это означает, что начальное состояние системы есть $s(0) = \bar{0} = (0, \dots, 0) \in S_0$ и $H(\bar{0}) = 0$. Учитывая абсолютную непрерывность распределений наработки, с течением времени состояние системы $s(t)$ меняется

Как было сказано выше, первый критерий определяет минимум действий, которые необходимо применить к подсистеме. Однако этого минимума может оказаться недостаточно для обеспечения второго критерия.

3 Оптимизация смягчающих действий

Рассмотрим более внимательно аспект, связанный с выбором действий, улучшающих параметры надежности. На текущий момент вопрос о том, для каких рисков предлагать смягчающие действия, а для каких нет, является в значительной мере субъективным и почти целиком зависит от человека, который этот анализ проводит. Строгих критериев или рекомендаций по этому вопросу не существует. И это является существенной

следующим образом. В некоторые (случайные) моменты времени τ_1, τ_2, \dots происходят последовательные отказы элементов, при этом $s(\tau_1) \in S_1, s(\tau_2) \in S_2, \dots$. Момент τ первого отказа системы определяется как

$$\tau = \inf \{ \tau_k : s(\tau_k) \in S_k^* \},$$

где $S_k^* = \{s \in S_k : H(s) = 1\}$. Обозначим через \mathbb{T} математическое ожидание τ .

Другой характеристикой элементов является время восстановления. В случае отказа элемента i он может быть восстановлен за время, которое задается плотностью распределения g_i . Пусть $\sigma = (\sigma_1, \dots, \sigma_n) = s(\tau)$ — состояние системы в момент ее первого отказа и пусть I_σ — номера тех элементов, которые привели к отказу в состоянии σ , т. е. $I_\sigma = \{1 \leq i \leq n : \sigma_i = 1\}$. Определим время на восстановление системы после первого отказа как $\vartheta = \sup_{i \in I_\sigma} \vartheta_i$, где ϑ_i —

случайные величины с плотностью распределения g_i , и обозначим через Θ математическое ожидание ϑ . Для каждого элемента i определим еще два показателя: $\rho_i^{(1)} = R_1(f_i)$ и $\rho_i^{(2)} = R_2(g_i)$, где R_1 и R_2 — заданные функционалы. Эти показатели, которые носят название соответственно «риск по отказам» и «риск по восстановлению», характеризуют риски, связанные с частотой отказов элементов и временными затратами на их восстановление. Критичностью элемента i назовем число $\rho_i = \rho_i^{(1)} \cdot \rho_i^{(2)}$.

Предположим теперь, что надежность элементов поддается корректировке в следующем смысле. Пусть Y — заданное множество произвольной природы. Элементы Y содержательно означают воздействия, которые можно выбрать и применить независимо к каждому элементу системы до начала ее функционирования с целью повышения надежности. Пусть $u_i = (y_{i1}, \dots, y_{in_i}), y_{ij} \in Y$, — набор воздействий на элемент i . Тогда определяющие надежность плотности распределения наработки и времени восстановления элементов оказываются параметризованными векторами u_i : $f_i = f_i(u_i), g_i = g_i(u_i)$. Определенные выше показатели становятся зависимыми от набора управляющих параметров, примененных к элементам: $\mathbb{T} = \mathbb{T}(u), \Theta = \Theta(u), \rho(u) = \{\rho_i(u), i = 1, \dots, n\}$.

Применение воздействий сопряжено с затратами, которые зададим с помощью функции $c : Y \rightarrow (0, \infty)$, так что затраты на применение воздействий из набора u составляют

$$C(u) = \sum_{i=1}^n \sum_{j=1}^{n_i} c(y_{ij}).$$

Таким образом, проблема выбора воздействий, повышающих надежность системы, может быть сформулирована в терминах показателей надежности \mathbb{T} и ρ , показателя технического обслуживания Θ и стоимостного показателя C . В качестве варианта поставим задачу оптимального выбора управляющих воздействий с целью минимизировать стоимость воздействий при заданных уровнях

средней наработки на отказ системы, среднего времени восстановления системы и надежности отдельных элементов:

найти $\min_u C(u)$ при условиях $T(u) \leq T_1$, $\Theta(u) \leq T_2$, $\rho_i(u) \leq r_i$, $i = 1, \dots, n$,

где T_1 , T_2 , и r_i — заданные числа.

Для того чтобы поставленная задача, или любая подобная ей, стала практически значимой, необходимо, прежде всего, адекватно задать исходные параметры модели:

- множество управляющих воздействий Y ;
- плотности распределения вероятностей наработки и времени восстановления элементов и их зависимость от наборов управляющих воздействий: $f_i(u_i)$, $g_i(u_i)$, $i = 1, \dots, n$;
- функционалы рисков R_1 и R_2 ;
- ценовую функцию c ;
- ограничения на показатели надежности: T_1 , T_2 и r_i , $i = 1, \dots, n$.

Задание значений для параметров модели, а также окончательная формулировка оптимизационной задачи лежат в области экспертной деятельности. Однако после того как это сделано, определение структурной функции H , а также расчет показателей надежности T и Θ не представляют принципиальной трудности благодаря развитому аппарату теории надежности [9–11] и независимости случайных величин, входящих в модель. Возможность решения задачи математического программирования также не вызывает сомнений, хотя формат данной статьи не предполагает подробное обсуждение этого вопроса. Заметим только, что алгоритм оптимизации может существенно зависеть от параметров модели, прежде всего от вида плотностей распределения вероятностей. Для иллюстрации приведем пример, основанный на немного упрощенной реальной задаче.

Пусть наработки на отказ и времена восстановления элементов имеют экспоненциальное распределение соответственно с параметрами λ_i и μ_i , $i = 1, \dots, n$. Пусть также система имеет последовательную структуру надежности. Это означает, что отказ одного элемента влечет за собой отказ системы. Легко проверить, что при этих условиях выполняются следующие равенства:

$$\lambda = T^{-1} = \lambda_1 + \dots + \lambda_n; \quad \Theta = (\lambda_1 \Theta_1 + \dots + \lambda_n \Theta_n) T,$$

где $\Theta_i = (\mu_i)^{-1}$ — среднее время восстановления i -го элемента.

В предлагаемом примере система состоит из 28 элементов, среди которых есть однотипные (табл. 3). К каждому элементу может быть применено одно воздействие. (Физическое содержание воздействий разное для разных типов элементов, и оно здесь не раскрывается.) Воздействие может оказывать влияние на оба показателя надежности, т. е. на интенсивность отказа и на среднее время

Таблица 3 Характеристики элементов, подлежащих воздействиям

Тип элемента (количество)	Интенсивность отказов, 10^{-8} ч^{-1}		Среднее время восстановления, ч		Условная стоимость воздействия
	без воздействия	с воздействием	без воздействия	с воздействием	
Механическое крепление (1)	57,1	5,71	168	168	3
Привод револьверной камеры (1)	571	57,1	840	72	4
Приводы детекторного модуля нейтронного спектрометра и аттенюатора из гидрида лития (3)	571	57,1	840	72	4
Защита экран от радиации (1)	114	11,4	720	720	4
Защитный экран от магнитного поля (1)	114	11,4	720	720	4
Детектор из высокочистого германия (1)	1140	2,86	4320	4320	5
Детектор из бромида лантана (1)	1000	10	4320	4320	5
Датчики магнитного поля (2)	100	10	96	96	3
Датчики температуры (2)	40	4	96	96	3
Источники гамма-излучения (2)	2280	57,1	168	168	3
Система светодиодов (1)	57,1	57,1	129	23	3
Предусилители (2)	22,8	22,8	840	72	4
Защита кабелей от электромагнитного излучения (1)	114	11,4	168	168	3
Комплект кабелей (1)	57,1	57,1	168	23	3
Анализатор и процессор цифрового сигнала (3)	571	571	840	72	4
Хранилище данных локального контроллера (1)	571	571	840	23	4
Высоковольтный источник питания для детектора из бромида лантана (1)	5710	571	840	72	4
Спектрометрический усилитель (1)	5710	571	840	72	4
Анализатор и контроллер детектора из высокочистого германия (1)	5710	571	840	72	4
Интерфейс системы управления, сбора и обработки информации (1)	571	571	840	23	4

восстановления. Но у некоторых элементов под влиянием воздействия изменяется только один показатель, а значение другого остается прежним. Фактор, названный выше критичностью элемента, в этом примере не учитывается.

В отсутствие всяких воздействий система имеет интенсивность отказов $\lambda^* = 29,81 \cdot 10^{-5} \text{ год}^{-1}$, среднее время восстановления $\Theta^* = 972,62 \text{ ч}$, а затраты, очевидно, равны 0. В другом крайнем случае, когда воздействия применяются ко всем элементам, соответствующие значения показателей $\lambda_* = 5,15 \cdot 10^{-5} \text{ год}^{-1}$, $\Theta_* = 76,10 \text{ ч}$, а затраты составляют 104. Ясно, что, оптимизируя воздействия

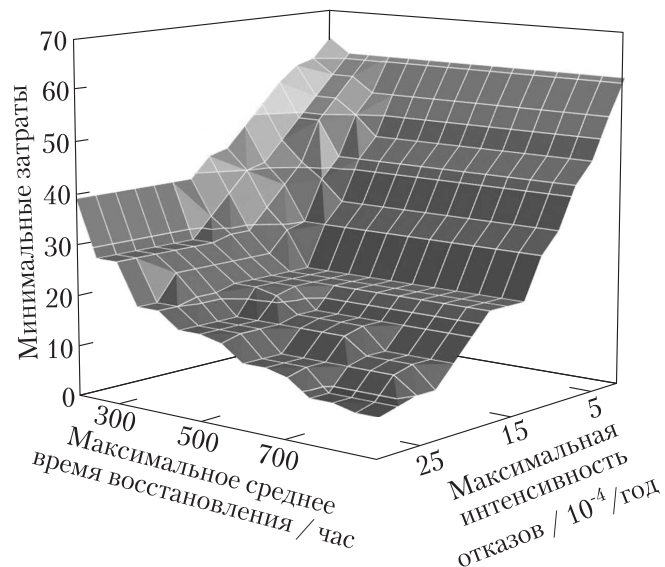


Рис. 6 Минимальная стоимость смягчающих воздействий при различных ограничениях на показатели надежности и технического обслуживания

по минимуму затрат, ограничения на интенсивность отказов и среднее время восстановления имеет смысл выбирать соответственно из интервалов $L = (\lambda_*, \lambda^*)$ и $T = (\Theta_*, \Theta^*)$. Задавая пару значений $(l \in L, t \in T)$, можно ставить вопрос о том, каковы воздействия, которые обеспечивают неравенства $\lambda \leq l$, $\Theta \leq t$ при минимальных затратах, обозначаемых $C(l, t)$. График функции $C(l, t)$ изображен на рис. 6. Он получен путем решения задачи на условный минимум в каждом из 225 узлов прямоугольной сетки внутри множества $L \times T$. На рис. 6 на горизонтальных осях отложены значения l и t , а вертикальная ось служит для значений $C(l, t)$.

Пусть, например, требуется, чтобы интенсивность отказов системы не превышала $l = 17,5 \cdot 10^{-5} \text{ год}^{-1}$, а среднее время восстановления было не больше чем $t = 525 \text{ ч}$. Тогда минимально необходимые затраты для этого составят 22 — столько требуется, чтобы применить воздействия к пяти элементам: HPGe detector, LaBr3 detector, HV power supply for LaBr3 detector, Spectrometric amplifier и HPGe analyzer with controller.

Заметим, что этот набор воздействий является в данном случае единственным решением дискретной оптимизационной задачи. Учитывая, что в рассматриваемом примере свыше 7 млн вариантов наборов, ясно, что «ручная» оптимизация затруднительна. Остается добавить, что в действительности задача имеет на порядок бóльшую размерность и содержит еще ряд дополнительных факторов,

которые не были упомянуты, но которые еще более усложняют выбор смягчающих воздействий.

4 Заключение

Подход Организации ИТЭР к анализу надежности включает 4 основных этапа и играет важную роль как инструмент контроля технических рисков. Согласно правилам Организации ИТЭР анализу подлежат все без исключения подсистемы ИТЭР, чтобы гарантировать надежную и безопасную эксплуатацию установки.

Многие подсистемы ИТЭР не имеют аналогов или же устанавливаемое оборудование никогда не использовалось в условиях ИТЭР. Это приводит к необходимости искать оригинальные инженерно-технические решения вопросов, связанных с обеспечением надежности, начиная с очевидного: «где брать параметры надежности и обслуживания для компонентов». Важная проблема возникает в связи с оптимизацией воздействий, уменьшающих вероятность отказов и времени восстановления. Применение предложенной в статье методики позволит сократить без ущерба для надежности и без того огромные затраты, связанные с проектом. Развитие изложенной в статье оптимизационной модели позволит учесть наличие неэкспоненциальных распределений, старение элементов и пр. Перспективой является создание информационной технологии нахождения эффективных решений по обеспечению надежности ИТЭР.

Литература

1. The ITER story. <http://www.iter.org/proj/iterhistory>.
2. *Holtkamp J.* An overview of the ITER project // *J. Fusion Eng. Design*, 2007. Vol. 82. No. 5-14. P. 427–434. doi: 10.1016/j.fusengdes.2007.03.029.
3. *Van Houtte D., Okayama K., Sagot F.* RAMI approach for ITER // *J. Fusion Eng. Design*, 2010. Vol. 85. No. 7-9. P. 1220–1224. doi:10.1016/j.fusengdes.2010.03.007.
4. Integration definition for function modeling (IDEF0). Federal Information Processing Standards Publication 183 (FIPS 183). — Gaithersburg, MD, USA: National Institute of Standards and Technology, 1993. 116 p.
5. *Марка Д., МакГоуэн К.* Методология структурного анализа и проектирования / Пер. с англ. — М.: МетаТехнология, 1993. 240 с. (*Marca D. A., McGowan C. L.* SATD: Structured analysis & design technique. — McGraw-Hill, 1988. 392 p.)
6. BlockSim: System Reliability and Maintainability Analysis Software Tool. <http://www.reliasoft.com/BlockSim/index.html>.
7. JPL-D-5703, Reliability analyses handbook. July 1990. http://everyspec.com/NASA/NASA-JPL/JPL-D-5703_JUL1990_15049.
8. *Gutierrez E.* Integral Asset Care. http://reliabilityweb.com/index.php/articles/integral_asset_care.
9. *Байхельт Ф., Франкен П.* Надежность и техническое обслуживание. Математический подход / Пер. с нем. — М.: Радио и связь, 1988. 392 с. (*Beichelt F., Franken P.*

Zuverlässigkeit und Instandhaltung. Mathematische Methoden. — Berlin: VEB Verlag Technik, 1983. 315 p.)

10. Кривулец В. Г., Полесский В. П. Об оценках надежности монотонной структуры // Проблемы передачи информации, 2001. Т. 37. Вып. 4. С. 112–129. doi:10.1023/A:1013831619318.
11. Острейковский В. А. Теория надежности. — М.: Высшая школа, 2003. 463 с.

Поступила в редакцию 02.12.15

METHODOLOGY OF RELIABILITY CALCULATION OF INTERNATIONAL THERMONUCLEAR EXPERIMENTAL REACTOR AND OPTIMIZATION OF RISK MITIGATION ACTIONS

G. M. Kononov

Institution “Project Center ITER” of the Russian Federation National Nuclear Corporation “ROSATOM,” 1-3 Kurchatova Sq., Moscow 123182, Russian Federation

Abstract: The article is devoted to one of the most important issues of creation of International Thermonuclear Experimental Reactor (ITER), namely, reliability. The principles of ITER reliability analysis approach which are applicable to the ITER and its subsystems are described. The basis of the reliability increasing is so-called RAMI-analysis (RAMI stands for Reliability, Availability, Maintainability, Inspectability). This analysis covers main aspects of the issue: increasing of availability and reliability, the ability to control, to maintain, to renovate, and so on. The main steps of the RAMI-analysis (functional analysis, calculation of availability and reliability, failure mode effects and criticality analysis, and determination of a list of mitigation actions) are shown on the base of real ITER subsystems. The determination of a list of mitigation actions is closely linked with optimization of possible actions which are focused on decreasing of failure possibility and decreasing of mean time to repair the failed component. A mathematical model allowing automating the selection of the most effective and the least expensive mitigation action process which is coupled with processing of large amounts of information is proposed.

Keywords: ITER; tokamak; large system reliability analysis; mitigation actions; mathematical models of reliability optimization

DOI: 10.14357/08696527160111

References

1. The ITER story. Available at: <http://www.iter.org/proj/iterhistory/> (accessed December 02, 2015).
2. Holtkamp, J. 2007. An overview of the ITER project. *J. Fusion Eng. Design* 82(5-14):427–434. doi: 10.1016/j.fusengdes.2007.03.029.

3. Van Houtte, D., K. Okayama, and F. Sagot. 2010. RAMI approach for ITER. *J. Fusion Eng. Design* 85(7-9):1220–1224. doi: 10.1016/j.fusengdes.2010.03.007.
4. Federal Information Processing Standards Publication 183 (FIPS 183). 1993. *Integration definition for function modeling (IDEF0)*. Gaithersburg, MD: National Institute of Standards and Technology. 116 p.
5. Marca, D. A., and C. L. McGowan. 1988. *SATD: Structured analysis & design technique*. McGraw-Hill. 392 p.
6. BlockSim: System Reliability and Maintainability Analysis Software Tool. Available at: <http://www.reliasoft.com/BlockSim/index.html/> (accessed December 02, 2015).
7. Reliability analysis handbook. Available at: http://everyspec.com/NASA/NASA-JPL/JPL_D-5703_JUL1990_15049/ (accessed December 02, 2015).
8. Integral Asset Care. Available at: http://reliabilityweb.com/index.php/articles/integral_asset_care/ (accessed December 02, 2015).
9. Beichelt, F., and P. Franken. 1983. *Zuverlässigkeit und Instandhaltung. Mathematische Methoden*. Berlin: VEB Verlag Technik. 315 p.
10. Krivulets, V. G., and V. P. Polesskiy. 2001. Ob otsenkakh nadezhnosti monotonoj struktury [On bounds for the monotone-structure reliability]. *Problemy Peredachi Informatsii* [Problems of Information Transmission] 37(4):380–396. doi:10.1023/A:1013831619318.
11. Ostreykovskiy, V. A. 2003. *Teoriya nadezhnosti* [Theory of reliability]. Moscow: Vysshaya Shkola. 463 p.

Received December 2, 2015

Contributor

Konovalov Grigory M. (b. 1980) — Candidate of Science (PhD) in physics and mathematics, leading scientist, Institution “Project Center ITER” of the Russian Federation National Nuclear Corporation “ROSATOM,” 1-3 Kurchatova Sq., Moscow 123182, Russian Federation; g.konovalov@iterrf.ru

ИНФОРМАЦИОННАЯ СИСТЕМА ПОДДЕРЖКИ ПРИНЯТИЯ ПРОЦЕССУАЛЬНЫХ РЕШЕНИЙ

Р. Р. Рзаев¹, Ф. Б. Агаев², А. И. Гююшев³, З. Р. Джамалов⁴

Аннотация: Предлагается подход к формированию системы информационной поддержки принятия процессуальных решений (СИППР), основанный на применении механизма нечеткого вывода, реализованного в нейросетевом логическом базисе. В рамках данного подхода разработан метод, позволяющий преодолеть семантическую неопределенность в оценочных понятиях процессуального права. В качестве примера выбрана статья «Нарушение авторских или смежных прав» Уголовного кодекса Азербайджанской Республики (УК АР), на базе которой был предложен формализм для оценочного понятия «значительный ущерб» в увязке с применяемой санкцией. Для вынесения адекватного оценочному понятию наказания предлагается шкала градации возможных санкций, полученная на основе описания соответствующей правовой нормы в терминах нечетких импликативных правил.

Ключевые слова: оценочное понятие; правовая норма; нечеткое множество; многослойная нейронная сеть; нечеткий вывод

DOI: 10.14357/08696527160112

1 Введение

Нейронные сети и нечеткие логические системы (НЛС) служат универсальным средством моделирования причинно-следственных связей. Как правило, их объединение дает возможность создавать принципиально новые аппаратные и программные средства, которые позволяют существенно расширить классы решаемых задач управления и принятия решений в условиях неопределенности, неточности и шума, коими характеризуется естественная среда. Хотя нейронные сети и НЛС имеют формальное сходство, между ними есть существенные различия. По своей природе НЛС является структурированным численно оценивающим механизмом, построенным в виде нечетких импликативных правил вида

¹Институт систем управления Национальной академии наук Азербайджана, ramingza@yahoo.com

²Аппарат Уполномоченного по правам человека (омбудсмена) Азербайджанской Республики, faig_agayev@yahoo.com

³Институт систем управления Национальной академии наук Азербайджана, geyushev@hotmail.com

⁴Бакинский государственный университет, zjamalov@mail.ru

«если... тогда...». Для представления композиционного правила вывода взвешивается выход каждого правила в соответствии со степенью принадлежности его входов и по всем выходам правил вычисляется центрост, обеспечивающий генерацию подходящего выхода НЛС.

Чаще всего проектирование НЛС проводится методами подбора (*trail-and-error design*) [1, 2]. При этом большинство подходов подразумевает субъективный выбор функций принадлежности и лингвистических правил на базе эвристических знаний в области изучения человеко-операционных систем или существующих контроллеров с последующим тестированием проектируемой НЛС на предмет генерации подходящего выхода. В противном случае функции принадлежности и/или логические правила подлежат настройке.

Исследования в этом направлении предусматривают процедуру самообучения НЛС, которая включает модификацию правил, основанную на концепции «*linguistic phase plane*» [3], и метод «логического испытания» (*logic examination*) для процесса конвертации входных-выходных данных в нечеткие правила управления, основанный на концепции «нечеткой идентификации» [4].

Существенное развитие технологии оптимизации НЛС нашло свое отражение еще в работах [5, 6], тем не менее до сих пор ведутся вполне успешные активные исследования в области проектирования НЛС. В настоящее время ввиду того, что внутренние (или скрытые) слои нейронных сетей остаются в определенном смысле «непрозрачными» для пользователей, большинство исследований концентрируется вокруг формирования оптимальных структур и размеров сетей. Тем не менее свое широкое применение нейронные сети нашли начиная с середины 1980-х гг., когда американский математик Д. Румелхарт предложил алгоритм обучения *error back-propagation* [7]. Именно с этих пор нейронные сети, принося свои способности к обучению в теорию автоматического управления, стали объектом активных исследований проектировщиков НЛС [8, 9].

В настоящей статье предпринимается попытка создания СИППР на базе НЛС, реализованной посредством обобщенной нейросетевой модели. Данная модель, представленная в виде *feedforward* многослойной нейронной сети, поддерживает идеологию нечеткого логического управления в нейросетевом логическом базисе, при этом СИППР формируется в автоматическом режиме посредством обучения и тестирования на основе наборов обучающих входных-выходных данных.

В нейросетевой модели входные и выходные нейроны представляют входные состояния и выходные управляющие сигналы/оценки соответственно, а нейроны из «скрытых» слоев олицетворяют функции принадлежности и нечеткие имплицативные правила. Все это позволяет симулировать способы человеческих рассуждений в рамках нейросетевой структуры, а также хранить согласованные правила механизма вывода, как и в случае традиционной системы логического вывода. Более того, предлагаемая архитектура СИППР в рамках оптимизации одной единственной целевой функции позволяет достаточно просто формулировать как параметрическое, так и структурное обучение. Такой подход обес-

печивает оригинальное решение задачи многокритериальной оптимизации при проектировании СИППР.

В [10] описана традиционная (обобщенная) модель системы нечеткого логического управления и принятия решений, которая была реализована на базе нейронной сети с *feedforward* многослойной топологической структурой. В контексте вышесказанного необходимо адаптировать структуру и функции этой системы под решение задачи преодоления семантической неопределенности в оценочных понятиях процессуального права и тем самым сформулировать новый подход к созданию гибкой СИППР.

2 Концепция системы информационной поддержки принятия процессуальных решений в нейросетевом логическом базисе

Руководствуясь рассмотренной в [10] типовой моделью системы нечеткого логического управления и принятия решений, представим обобщенную схему СИППР так, как это показано на рис. 1. Данная схема включает в себя три основных компоненты: фаззификатор, базовые нечеткие правила и механизм нечеткого вывода и дефаззификатор.

Фаззификатор осуществляет процедуру фаззификации посредством заранее установленных функций принадлежности, описывающих входные данные в виде нечетких множеств как значений входных лингвистических переменных. Базовые правила представляют собой нечеткие импликации вида «если... , тогда...», которые на начальном этапе описывают эвристические знания в области процессуального права. Механизм нечеткого вывода, реализуя композицию данных правил, индуцирует нечеткий вывод об оценочном понятии в правовой норме, чтобы выработать адекватное ему процессуальное решение. Дефаззификатор

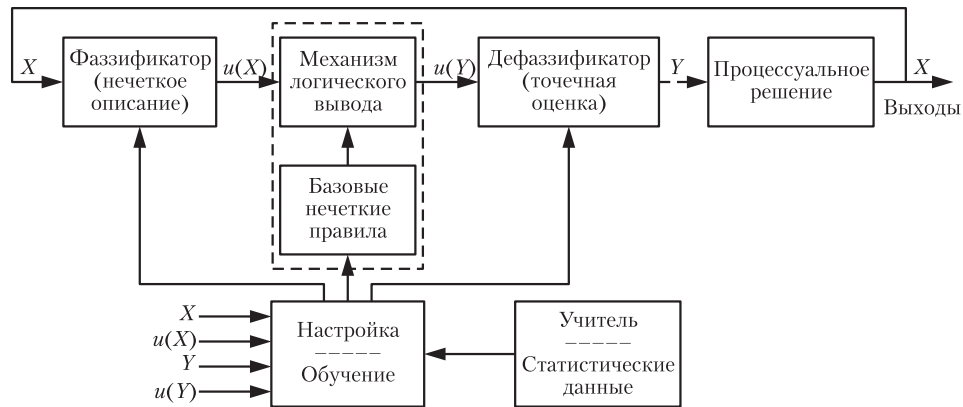


Рис. 1 Концептуальная структура СИППР

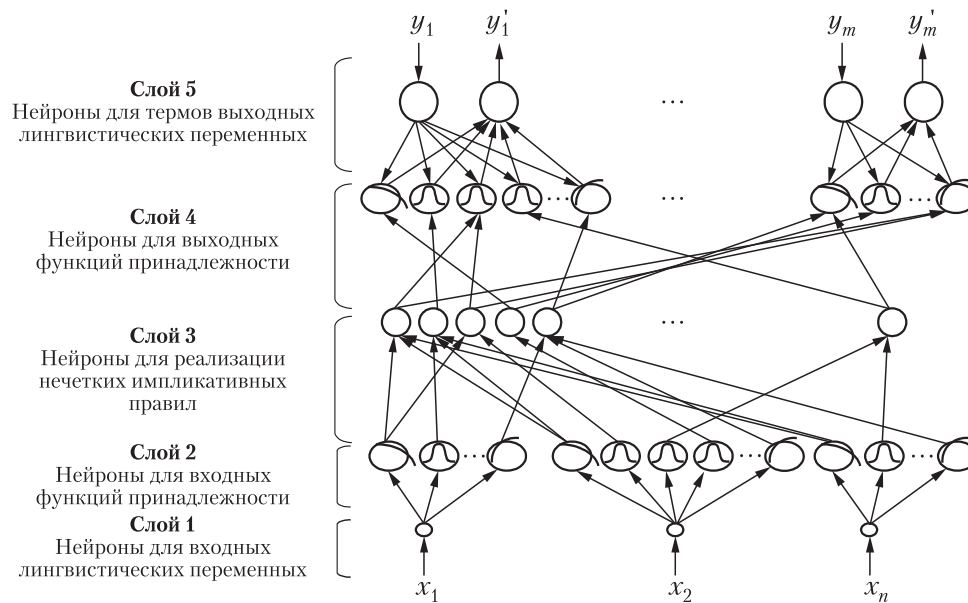


Рис. 2 Системы информационной поддержки принятия процессуальных решений в нотации *feedforward* пятислойной нейронной сети

осуществляет процедуру дефаззификации нечетких выходов, т. е. представление нечетких выводов об оценочном понятии в виде обычных чисел посредством, например, центроидного метода. При этом основной проблемой проектирования СИППР является выбор подходящих входных-выходных функций принадлежности и набора нечетких логических правил.

Основанная на базовой структуре и концепции НЛС СИППР в нейросетевом логическом базисе с коннекционной топологической структурой и способностью к обучению призвана устранить данную проблему. На рис. 2 представлена структура такой системы, которая состоит из пяти слоев.

Нейроны 1-го (входного) слоя представляют собой входные лингвистические переменные, и поэтому их можно интерпретировать как рецепторы. Пятый слой является выходным, и поэтому все его нейроны, по сути, исполняют роль эффекторов. Нейроны 2-го и 4-го слоев активируются как функции принадлежности, чтобы представлять термы (значения) соответствующих лингвистических переменных. Каждый нейрон 3-го слоя имитирует одно нечеткое логическое правило, а совокупность таких нейронов в составе данного слоя формирует базовый набор нечетких логических правил. Связи между 3-м и 4-м слоями в совокупности функционируют как коннекционный (ассоциативный) механизм вывода. Входные связи 3-го слоя определяют причины (предусловия) для нечет-

ких логических правил, а входные связи 4-го слоя определяют следствия, т. е. в совокупности они формируют причинно-следственные связи в рамках механизма нечеткого вывода. При этом каждый нейрон-правило имеет не более одной входной связи, исходящей из некоторого нейрона — термина входной лингвистической переменной.

В общем виде вход для каждого нейрона данной коннекционной нейронной сети формулируем как

$$\text{Input} = f(u_1^k, u_2^k, \dots, u_p^k; w_1^k, w_2^k, \dots, w_p^k),$$

где k — номер слоя; p — номер входной связи; u_i^k ($i = 1, \dots, p$) — i -й сигнал из k -го слоя; w_i^k ($i = 1, \dots, p$) — вес i -й связи из k -го слоя. В рамках принятых обозначений выходы нейронов в результате их активаций обозначим как

$$\text{Output} = o_i^k = a(f)(i = 1, \dots, p),$$

где $a(\cdot)$ является функцией активации. В частности, в качестве функции активации можно выбрать сигмоидную функцию вида $a = 1/(1+e^{-f})$, где $f = \sum_{i=1}^p w_i^k u_i^k$.

3 Симуляция системы информационной поддержки принятия решений в нотации пакета MATLAB (на примере статьи 165 «Нарушение авторских или смежных прав» Уголовного кодекса Азербайджанской республики)

Цель любого симулятора состоит в демонстрации построения (включая процесс обучения) и функционирования исследуемой системы. В данном случае симулятор имитирует создание СИППР в нотации *feedforward* пятислойной нейронной сети, с помощью которой оценочное понятие в правовой норме численно интерпретируется для поддержки выбора процессуального решения. Система информационной поддержки принятия решений, обладая способностью к обучению через обучающие примеры, призвана функционировать в автоматическом режиме. Основной целью является демонстрация способности СИППР численно интерпретировать оценочные понятия в правовых нормах на базе существующих процессуальных решений в области правоприменения.

Итак, в качестве примера рассмотрим статью 165 «Нарушение авторских или смежных прав» УК АР [11], которая формулируется следующим образом:

«Незаконное использование объектов авторского или смежных прав, то есть издание под своим именем или иное присвоение авторства на чужое научное, литературное, художественное или иное произведение, его незаконное переиздание или распространение, а равно принуждение к соавторству, если эти

деяния причинили значительный ущерб, наказываются штрафом в размере от 100 до 500 AZN либо общественными работами на срок от 160 до 240 часов. Те же деяния, совершенные неоднократно, или группой лиц по предварительному сговору, или организованной группой, наказываются штрафом в размере от 500 до 1000 AZN либо лишением свободы на срок до трех лет с конфискацией имущества».

В данном случае составом правонарушения A является «незаконное использование объектов авторского или смежных прав», характеризующееся признаками B_1 — «издание под своим именем или иное присвоение авторства на чужое научное, литературное, художественное или иное произведение, его незаконное переиздание или распространение, а равно принуждение к соавторству». Санкцией же S_1 является «наказание штрафом в размере от ста до пятисот манатов» либо «назначение общественных работ на срок от ста шестидесяти до двухсот сорока часов». В случае наличия дополнительных признаков состава правонарушения B_2 — «деяния, совершенные неоднократно, или группой лиц по предварительному сговору, или организованной группой» — применяется санкция S_2 — «наказание штрафом в размере от пятисот до тысячи манатов либо лишением свободы на срок до трех лет с конфискацией имущества». С учетом принятых обозначений эту норму можно записать как

$$\left. \begin{array}{l} (A \approx B_1) \Rightarrow S_1 ; \\ (A \approx (B_1 \wedge B_2)) \Rightarrow S_2 . \end{array} \right\} \quad (1)$$

Казалось бы, с формулировкой данной нормы права и с ее логическим формализмом (1) все выглядит достаточно определенно. Но это только на первый взгляд. Камнем преткновения здесь является оценочное понятие «значительный ущерб», вернее его семантическая неопределенность. Его толкование наряду с другими терминами лингвистической переменной «ущерб», такими как «незначительный ущерб», «большой ущерб», «ущерб в особо крупном размере», «существенный ущерб» и т. д., занимает значительное место в правоприменении и в правотворчестве. Все дело в том, что необдуманная оценка величины ущерба может, с одной стороны, привести к недостаточному обеспечению охраны авторских или смежных прав, а с другой — к необоснованному расширению принимаемой санкции. Более того, в разных ситуациях термы лингвистической переменной «ущерб» могут трактоваться по-разному. Поэтому для каждого конкретного случая целесообразно формировать соответствующую шкалу градации оценочных понятий, используемых в процессе правоприменения.

3.1 Классификация правонарушений

Очевидно, что процесс правоприменения является многокритериальной процедурой, подразумевающей применение композиционного правила агрегирования

оценки по каждому конкретному делу. Для оценки правонарушения и формулирования соответствующей санкции выберем пять оценочных понятий:

- (1) u_1 — «НЕЗНАЧИТЕЛЬНЫЙ УЩЕРБ»;
- (2) u_2 — «СУЩЕСТВЕННЫЙ УЩЕРБ»;
- (3) u_3 — «БОЛЬШОЙ УЩЕРБ»;
- (4) u_4 — «ЗНАЧИТЕЛЬНЫЙ УЩЕРБ»;
- (5) u_5 — «УЩЕРБ В ОСОБО КРУПНОМ РАЗМЕРЕ».

Проще говоря, под множеством $C = (u_1, u_2, u_3, u_4, u_5)$ будем понимать совокупность признаков, по которым классифицируются санкции. Тогда, полагая используемые в правовой норме критерии нечеткими множествами, оценку правонарушения произведем с использованием достаточного набора нечетких импликативных правил вида «если... , тогда...» и на их основе установим соответствующую шкалу градации возможных санкций. Итак, перефразируем положения ст. № 165 УК АР следующим образом:

- e_1 : «Если имеет место издание под своим именем или иное присвоение авторства на чужое научное, литературное, художественное или иное произведение, а равно и принуждение к соавторству, то назначается минимальное наказание (штраф в размере от 100 до 500 AZN)»;
- e_2 : «Если вдобавок к вышеприведенным требованиям имеет место незаконное переиздание или распространение чужого научного, литературного, художественного или иного произведения, то назначается наказание выше минимального (назначение общественных работ на срок от 160 до 240 часов)»;
- e_3 : «Если дополнительно к условиям, оговоренным в e_2 , те же деяния были совершены неоднократно группой лиц по предварительному сговору или организованной группой, то назначается максимальное наказание (лишение свободы сроком на 3 года с конфискацией имущества)»;
- e_4 : «Если в рассматриваемом деле имеют место признаки, оговоренные в e_3 , однако вместо группы лиц ущерб нанесен со стороны одного лица, нарушавшего авторские права неоднократно, то назначается наказание средней тяжести (штраф в размере от 500 до 1000 AZN)»;
- e_5 : «Если имеет место принуждение к соавторству, но не подтверждено незаконное переиздание или распространение чужого научного, литературного, художественного или иного произведения группой лиц по предварительному сговору или организованной группой, то назначается наказание средней тяжести (штраф в размере от 500 до 1000 AZN)»;
- e_6 : «Если не подтверждено издание под своим именем или иное присвоение авторства на чужое научное, литературное, художественное или иное произведение, а равно и принуждение к соавторству, то состав правонарушения отсутствует».

Таблица 1 Переменные НЛС

Входные переменные	X ₁	Имя переменной	Издание под своим именем или иное присвоение авторства на чужое произведение
		Значение переменной	A = ИМЕЕТ МЕСТО
		Пределы значений	[0, 1]
	X ₂	Имя переменной	Принуждение к соавторству
		Значение переменной	B = ИМЕЕТ МЕСТО
		Пределы значений	[0, 1]
	X ₃	Имя переменной	Переиздание или распространение чужого произведения
		Значение переменной	C = ИМЕЕТ МЕСТО
		Пределы значений	[0, 1]
	X ₄	Имя переменной	Нарушение авторских прав группой лиц неоднократно по предварительному сговору или организованной группой
		Значение переменной	D = ИМЕЕТ МЕСТО
		Пределы значений	[0, 1]
	X ₅	Имя переменной	Нарушение авторских прав одним лицом неоднократно
		Значение переменной	E = ИМЕЕТ МЕСТО
		Пределы значений	[0, 1]
Выходная переменная Y	Имя переменной	Наказание (уровень применяемой санкции)	
	Терм-множество	{Y ₀ = ОТСУТСТВУЕТ, Y ₁ = МИНИМАЛЬНОЕ, Y ₂ = ВЫШЕ МИНИМАЛЬНОГО, Y ₃ = СРЕДНЕЙ ТЯЖЕСТИ, Y ₄ = МАКСИМАЛЬНОЕ}	
	Пределы значений	[0, 1]	

Эти рассуждения позволяют сформировать полный набор лингвистических переменных и правил для построения НЛС. Для удобства все переменные сведены в табл. 1.

Сама система правил в символьной форме будет иметь следующий вид:

$$\begin{aligned}
 e_1 &: (X_1 = A) \&(X_2 = B) \Rightarrow (Y = Y_1); \\
 e_2 &: (X_1 = A) \&(X_2 = B) \&(X_3 = C) \Rightarrow (Y = Y_2); \\
 e_3 &: (X_1 = A) \&(X_2 = B) \&(X_3 = C) \&(X_4 = D) \Rightarrow (Y = Y_4); \\
 e_4 &: (X_1 = A) \&(X_2 = B) \&(X_3 = C) \&(X_5 = E) \Rightarrow (Y = Y_3); \\
 e_5 &: (X_2 = B) \&(X_3 = \neg C) \&(X_4 = \neg D) \Rightarrow (Y = Y_3); \\
 e_6 &: (X_1 = \neg A) \&(X_2 = \neg B) \Rightarrow (Y = Y_0).
 \end{aligned}$$

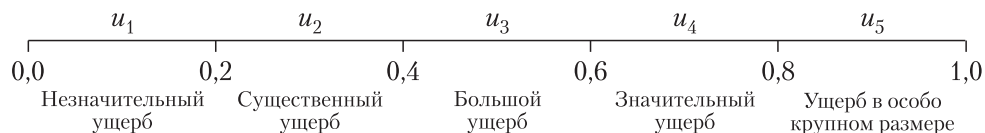


Рис. 3 Уровни ущерба для оценки правонарушения в масштабе [0, 1]

Переменная Y задана на дискретном множестве $J = \{0; 0,1; 0,2; \dots; 1\}$, а ее термы описаны следующими нечеткими множествами:

$$\begin{aligned}
 Y_1 &= \text{МИНИМАЛЬНОЕ}, \quad \mu_{Y_1}(x) = x, \quad x \in J; \\
 Y_2 &= \text{ВЫШЕ МИНИМАЛЬНОГО}, \quad \mu_{Y_2}(x) = \sqrt{x}; \\
 Y_3 &= \text{СРЕДНЕЙ ТЯЖЕСТИ}, \quad \mu_{Y_3}(x) = x^2; \\
 Y_4 &= \text{МАКСИМАЛЬНОЕ}, \quad \mu_{Y_4}(x) = \begin{cases} 1, & x = 1; \\ 0, & x < 1; \end{cases} \\
 Y_0 &= \text{ОТСУТСТВУЕТ}, \quad \mu_{Y_0}(x) = 1 - x; \quad x \in J.
 \end{aligned}$$

Фаззификация термов переменных X_k осуществлена с применением функций вида

$$\mu(u) = \exp \left\{ -\frac{(u - 1)^2}{\sigma_k^2} \right\} \quad (k = 1, \dots, 5),$$

восстанавливающих нечеткие множества по опорному вектору $(u_1, u_2, u_3, u_4, u_5)$, где $u_i = \{(d_{i-1} + d_i)/2\} \in [d_0, d_i]$. При этом значения для σ_k подбираются исходя из степени важности признаков правонарушения в рамках правовой нормы. Очевидно, что простым преобразованием $x = d_0 + t(d_5 - d_0)$ ($t \in [0, 1]$) отрезок $[d_0, d_5]$ можно легко свести к отрезку $[0, 1]$. В итоге оценка нарушения авторских прав с точки зрения нанесенного ущерба, градуированного на отрезке $[0, 1]$ (рис. 3), где $d_i = 0,2i$ ($i = 0, \dots, 5$), позволила описать значения входных переменных в виде соответствующих нечетких множеств:

$$\begin{aligned}
 A &= \frac{0,0392}{u_1} + \frac{0,1409}{u_2} + \frac{0,3679}{u_3} + \frac{0,6977}{u_4} + \frac{0,9608}{u_5}; \\
 B &= \frac{0,0183}{u_1} + \frac{0,0889}{u_2} + \frac{0,2910}{u_3} + \frac{0,6412}{u_4} + \frac{0,9518}{u_5}; \\
 C &= \frac{0,0063}{u_1} + \frac{0,0468}{u_2} + \frac{0,2096}{u_3} + \frac{0,5698}{u_4} + \frac{0,9394}{u_5}; \\
 D &= \frac{0,0013}{u_1} + \frac{0,0183}{u_2} + \frac{0,1299}{u_3} + \frac{0,4797}{u_4} + \frac{0,9216}{u_5};
 \end{aligned}$$

$$E = \frac{0,0001}{u_1} + \frac{0,0043}{u_2} + \frac{0,0622}{u_3} + \frac{0,3679}{u_4} + \frac{0,8948}{u_5}.$$

Для реализации правил e_1 – e_6 был использован оператор импликации Лукасевича

$$I(x, y) = \min(1, 1 - x + y),$$

а для фаззификации нечетких выходов как уровней наказания E_k ($k = 1, \dots, 5$) применено равенство [12]

$$F(E_k) = \frac{1}{\alpha_{\max}} \int_0^{\alpha_{\max}} M(E_{k\alpha}) d\alpha,$$

в котором $E_{k\alpha} = \{x | \mu_{E_k}(x) \geq \alpha, x \in J\}$; $M(E_{k\alpha}) = \sum_{j=1}^n x_j/n$; α_{\max} — максимальное значение на $E_{k\alpha}$. В результате получены следующие дефаззифицированные значения выходов нечеткой модели правовой нормы «Нарушение авторских или смежных прав»:

- при *незначительном* ущербе (u_1) число $F(E_1) = 0,3268$ как верхняя граница интервала, соответствующего вынесению санкции E_1 ;
- при *существенном* ущербе (u_2) число $F(E_2) = 0,3714$ как верхняя граница интервала, соответствующего вынесению санкции E_2 ;
- при *большом* ущербе (u_3) число $F(E_3) = 0,4578$ как верхняя граница интервала, соответствующего вынесению санкции E_3 ;
- при *значительном* ущербе (u_4) число $F(E_4) = 0,6431$ как верхняя граница интервала, соответствующего вынесению санкции E_4 ;
- при ущербе *в особо крупном размере* (u_5) число $F(E_5) = 0,9608$ как верхняя граница интервала, соответствующего вынесению санкции E_5 .

В принятых допущениях итоговая шкала для оценки и выбора санкций по рассматриваемому делу о нарушении авторских прав может выглядеть так, как это показано на рис. 4.

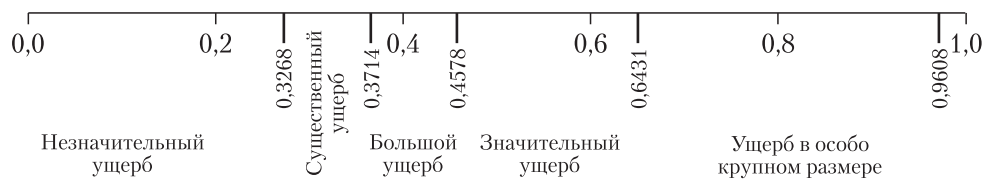


Рис. 4 Шкала для выбора санкций на базе уровней возможных ущербов

3.2 Оценка правонарушения

После того как была установлена обоснованная шкала для классификации санкций применительно к незаконному использованию авторского или смежных прав, можно приступить собственно к реализации СИППР. При этом, опираясь на логическую структуру (1) правовой нормы, за основу следует выбрать следующие непротиворечивые и вполне обоснованные рассуждения:

- d_1 : если впервые в результате издания путем присвоения авторства на чужое произведение, или переиздания, или распространения, а равно принуждения к соавторству причинен незначительный ущерб, то состав преступления отсутствует;
- d_2 : если впервые в результате издания путем присвоения авторства на чужое произведение, или переиздания, или распространения, а равно принуждения к соавторству причинен существенный ущерб, то применяется минимальное наказание в виде выплаты штрафа от 100 до 500 AZN;
- d_3 : если впервые в результате издания путем присвоения авторства на чужое произведение, или переиздания, или распространения, а равно принуждения к соавторству причинен большой ущерб, то применяется санкция выше минимальной, т. е. наказание в виде выполнения общественных работ сроком от 160 до 240 ч;
- d_4 : если неоднократно в результате издания путем присвоения авторства на чужое произведение, или переиздания, или распространения, а равно принуждения к соавторству причинен значительный ущерб, то применяется наказание средней тяжести в виде выплаты штрафа от 500 до 1000 AZN;
- d_5 : если неоднократно в результате издания путем присвоения авторства на чужое произведение, или переиздания, или распространения, а равно принуждения к соавторству причинен ущерб в особо крупном размере, то применяется максимальное наказание в виде лишения свободы сроком до трех лет с конфискацией имущества;
- d_6 : если указанные деяния совершены группой лиц по предварительному сговору или организованной группой и при этом причинен значительный ущерб, то также применяется максимальное наказание в виде лишения свободы сроком до трех лет с конфискацией имущества.

Данные рассуждения устанавливают причинно-следственные связи между характеристиками состава правонарушения, с одной стороны, и собственно применяемой санкцией, с другой. Входными характеристиками здесь являются:

- переменная x_1 — размер нанесенного ущерба, принимающая 5 значений, формализуемых с помощью гауссовских функций принадлежности с вершинами соответственно в точках с абсциссами $u_i = [0,2(i - 1) + 0,2i]/2$ ($i = 1, \dots, 5$);

– переменная x_2 — наличие подобного прецедента в прошлом и/или наличие других участников преступления, принимающая 3 значения, формализуемых с помощью гауссовских функций принадлежности с вершинами соответственных точек с абсциссами 0, 0,5 и 1.

Выходной характеристикой является переменная y — уровень наказания, с 5 термами, формализуемыми с помощью гауссовских функций принадлежности с вершинами соответственно в точках: 0,3268, 0,3714, 0,4578, 0,6431 и 0,9608.

Реализация правил d_1-d_6 в нотации MATLAB с помощью FIS-редактора типа Мамдани позволила сформировать выборку обучающих пар (табл. 2) для построения СИППР.

Загрузка данного обучающего множества в редактор ANFIS типа Сугено сгенерировала структуру СИППР в нотации *feedforward* пятислойной нейронной сети вида, описанного на рис. 2. В данном случае структура СИППР включает 15 правил, 2 входные и одну выходную переменные. Для ее настройки использован гибридный метод, представляющий собой комбинацию метода наименьших квадратов и метода убывания обратного градиента. Так, на рис. 5 представлены оптимизированные функции принадлежности, описывающие термы входных лингвистических переменных.

На рис. 6 представлен графический интерфейс просмотра правил сгенерированной СИППР, а также фрагмент оптимального набора базовых правил.

В табл. 3 перечислены «обученные» нечеткие импликативные правила в несколько упрощенном виде. Например, правило 4, выделенное на рис. 6, интерпретируется так:

Если x_1 является $X_{12} = \text{СУЩЕСТВЕННЫЙ}$ (т. е. нечетким множеством, описываемым гауссовской функцией $\mu_{X_{12}}(u) = \exp\{-(u - 0,3075)^2/0,0703^2\}$) и x_2 является $X_{21} = \text{ВПЕРВЫЕ ИМЕЕТ МЕСТО}$

Таблица 2 Набор входных-выходных данных для построения СИППР

№	Входы		Выход
	x_1	x_2	y
1	0,100	0,000	0,383
2	0,300	0,000	0,407
3	0,500	0,000	0,479
4	0,700	0,000	0,644
5	0,900	0,000	0,711
6	0,100	0,500	0,440
7	0,300	0,500	0,483
8	0,500	0,500	0,644
9	0,700	0,500	0,643
10	0,900	0,500	0,855
11	0,100	1,000	0,556
12	0,300	1,000	0,644
13	0,500	1,000	0,711
14	0,700	1,000	0,855
15	0,900	1,000	0,711
16	0,265	0,868	0,538
17	0,507	0,241	0,492
18	0,779	0,150	0,723
19	0,823	0,232	0,738
20	0,852	0,068	0,718
21	0,874	0,732	0,802
22	0,206	0,859	0,535
23	0,214	0,377	0,426
24	0,177	0,786	0,514
25	0,845	0,195	0,745
26	0,625	0,105	0,532
27	0,808	0,914	0,779
28	0,148	0,059	0,393
29	0,133	0,768	0,483
30	0,852	0,323	0,788

Примечания: x_1 — размер нанесенного ущерба; x_2 — наличие прецедента в прошлом или наличие других участников преступления; y — уровень наказания

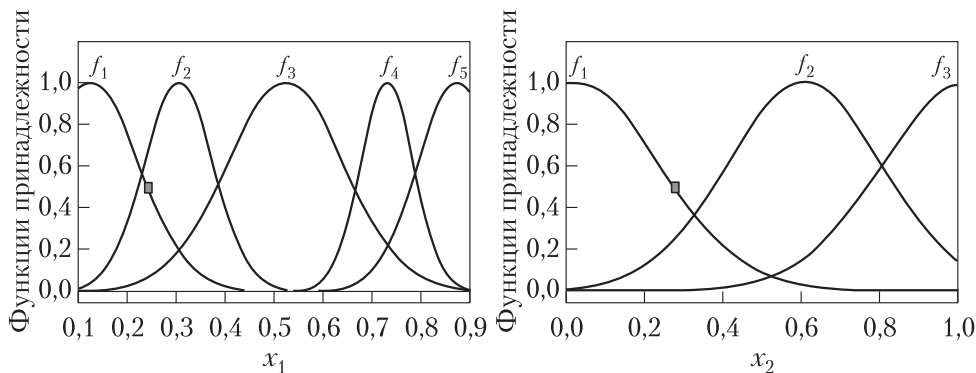


Рис. 5 Оптимизированные функции принадлежности входов НЛС

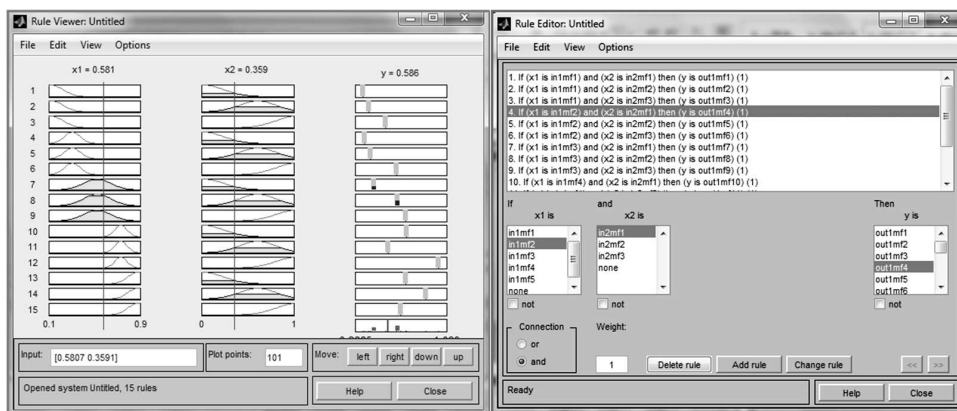


Рис. 6 Графический интерфейс просмотра правил сгенерированной СИППР

(т. е. нечетким множеством, описываемым гауссовской функцией $\mu_{X_{21}}(u) = \exp\{- (u - 0,6090)^2 / 0,1983^2\}$), тогда y является $Y_2 = \text{ВЫПЛАТА ШТРАФА}$ (от 100 до 500 AZN).

Как видно из табл. 3, выбраны 5 оценок правонарушения на основе установленных 5 интервалов, охватывающих дефаззифицированные значения нечетких выходов y : $[0; 0,4]$, $[0,4; 0,55]$, $[0,55; 0,7]$, $[0,7; 0,85]$ и $[0,85; 1]$.

4 Заключение

В процессе симуляции СИППР были использованы три критерия оценки состава правонарушения, на базе которых посредством применения адаптированной FIS типа Сугено удалось построить и изучить адекватную причинно-следст-

Таблица 3 Настроенные по гибриднему алгоритму логические правила

№	Входы		Дефаззифицированное значение	Выход Оценка правонарушения (санкция)
	x_1	x_2		
1	X_{11}	X_{21}	0,3859	Y_1 — состав преступления отсутствует
2	X_{11}	X_{22}	0,4334	Y_2 — выплата штрафа от 100 до 500 манат
3	X_{11}	X_{23}	0,5678	Y_3 — общественная работа сроком от 160 до 240 ч
4	X_{12}	X_{21}	0,4005	Y_2 — выплата штрафа от 100 до 500 манат
5	X_{12}	X_{22}	0,4453	Y_2 — выплата штрафа от 100 до 500 манат
6	X_{12}	X_{23}	0,6503	Y_3 — общественная работа сроком от 160 до 240 ч
7	X_{13}	X_{21}	0,4704	Y_2 — выплата штрафа от 100 до 500 манат
8	X_{13}	X_{22}	0,6572	Y_3 — общественная работа сроком от 160 до 240 ч
9	X_{13}	X_{23}	0,7226	Y_4 — выплата штрафа от 500 до 1000 манат
10	X_{14}	X_{21}	0,7302	Y_4 — выплата штрафа от 500 до 1000 манат
11	X_{14}	X_{22}	0,5859	Y_3 — общественная работа сроком от 160 до 240 ч
12	X_{14}	X_{23}	0,9796	Y_5 — лишение свободы сроком до трех лет с конфискацией имущества
13	X_{15}	X_{21}	0,7242	Y_4 — выплата штрафа от 500 до 1000 манат
14	X_{15}	X_{22}	0,8815	Y_5 — лишение свободы сроком до трех лет с конфискацией имущества
15	X_{15}	X_{23}	0,6810	Y_3 — общественная работа сроком от 160 до 240 ч

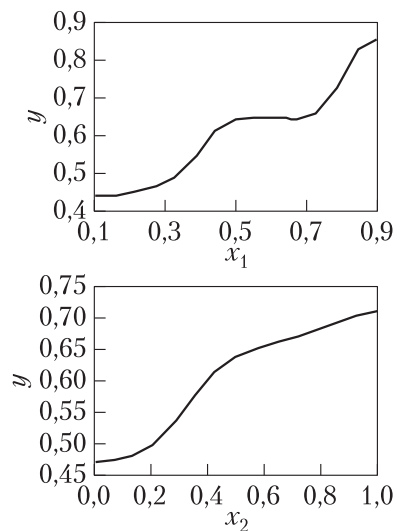
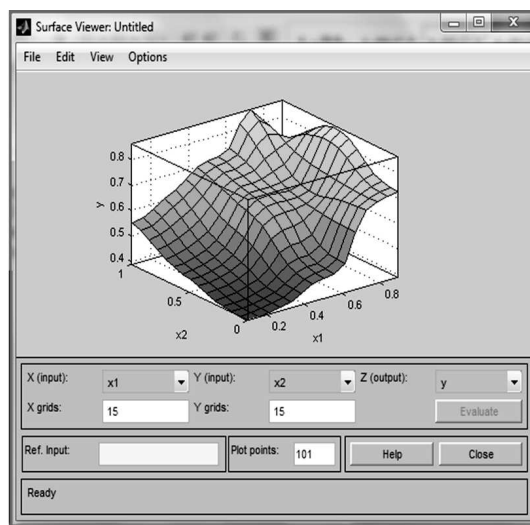


Рис. 7 Зависимость санкции от состава правонарушения

венную связь между характеристиками состава правонарушения и результатом правоприменения (рис. 7).

Численная интерпретация степени правонарушения и соответственно уровня выносимой санкции не превышает величину 0,85 при максимальной 1, относительно умеренно увеличивается с ростом x_2 и резко увеличивается при росте x_1 . Предлагаемая СИППР позволяет достаточно быстро и относительно легко диверсифицировать свои функции и на другие виды правонарушений. Для этого необходимо формализовать соответствующие правовые нормы и сформировать достаточную статистику оценок по различным сценариям состава правонарушений. В перспективе подобная система способна функционировать и в автономном режиме, так как при ее разработке и адаптации достаточно привлекать эвристические знания в области правоприменения.

Литература

1. Industrial applications of fuzzy control / Ed. M. Sugeno. — Amsterdam: Elsevier Science Publs. B. V., 1985. P. 231–239.
2. Bernard J. A. Use of rule-based system for process control // Control Syst. Magazine, 1988. Vol. 8. No. 5. P. 3–13.
3. Braae M., Rutherford D. A. Selection of parameters for a fuzzy logic controller // Fuzzy Sets Syst., 1979. Vol. 2. No. 3. P. 185–199.
4. Tong R. M. Synthesis of fuzzy models for industrial processes // Int. J. General Syst., 1978. Vol. 4. P. 143–162.
5. Procyk T. J., Mamdani E. H. A linguistic self-organizing process controller // Automatica, 1979. Vol. 15. No. 1. P. 15–30.
6. Takagi T., Sugeno M. Derivation of fuzzy control rules from human operator's control actions // Fuzzy Information Knowledge Representation Decision Analysis: IFAC Symposium Proceedings. — Marseilles, France, 1983. P. 55–60.
7. Rumelhart D. E., Hinton G. E., Williams R. J. Learning internal representations by error propagation // Parallel Distributed Proc., 1986. Vol. 1. P. 318–362.
8. Lin C. T., Lee C. S. G. Neural-network-based fuzzy logic control and decision system // IEEE Trans. Comp., 1991. Vol. 40. No. 12. P. 1320–1336.
9. Kosko B. Neural networks and fuzzy systems. — Prentice Hall, NJ, USA: Englewood Cliffs, 1992. 456 p.
10. Lin C. T., Lee C. S. G. Supervised and unsupervised learning with fuzzy similarity for neural network-based fuzzy logic control systems // Fuzzy sets, neural networks, and soft computing / Eds. R. R. Yager, L. A. Zadeh. — New York, NY, USA: Van Nostrand Reinhold, 1994. P. 85–125.
11. Уголовный кодекс Азербайджанской Республики, утвержденный Законом Азербайджанской Республики от 30 декабря 1999 г. № 787-IQ. Вступил в силу с 1 сентября 2000 г. согласно Закону Азербайджанской Республики от 26 мая 2000 г. № 886-IQ. <http://www.legislationonline.com>.
12. Рзаев Р. Р. Интеллектуальный анализ данных в системах поддержки принятия решений. — Saarbrücken, Germany: LAP Lambert Academic Publishing GmbH & Co. KG, 2013. 130 с.

Поступила в редакцию 22.09.15

INFORMATION SYSTEM OF PROCEDURAL DECISION-MAKING SUPPORT

R. R. Rzayev¹, F. B. Agayev², A. I. Goyushov¹, and Z. R. Jamalov³

¹Institute of Control Systems of the Azerbaijan National Academy of Sciences, 9 B. Vahabzadeh Str., Baku AZ1141, Azerbaijan Republic

²Information and International Relations Department of the Office of the Commissioner for Human Rights (Ombudsman) of the Republic of Azerbaijan, 40 Uzeyir Hajibeyov Str., Baku AZ1148, Azerbaijan Republic

³Baku State University, 23 Zakhid Khalilov Str., Baku AZ1148, Azerbaijan Republic

Abstract: An approach to the formation of the system of procedural decision-making information support is suggested, which is based on the application of the fuzzy inference mechanism implemented within the neural network logical basis. Using this approach, the method allowing overcoming the semantic uncertainty in criterion concepts of a procedural law is proposed. As an example, the authors chose the Article “Violation of the author’s or adjacent rights” of the Criminal Code of the Azerbaijan Republic, on the basis of which the formalism for the criterion concept “extensive damage” coupled with the applied sanction was suggested. For imposition of adequate to criterion concept sentence, the paper proposed the scale of possible sanctions obtained on the basis of the description of the correspondent legal norm in terms of fuzzy implicative rules.

Keywords: appraisal concept; legal norm; fuzzy set; multilayer neural network; fuzzy conclusion

DOI: 10.14357/08696527160112

References

1. Sugeno, M., ed. 1985. *Industrial applications of fuzzy control*. Amsterdam: Elsevier Science Pubs. B. V. 231–239.
2. Bernard, J. A. 1988. Use of rule-based system for process control. *Control Syst. Magazine* 8(5):3–13.
3. Braae, M., and D. A. Rutherford. 1979. Selection of parameters for a fuzzy logic controller. *Fuzzy Sets Syst.* 2(3):185–199.
4. Tong, R. M. 1978. Synthesis of fuzzy models for industrial processes. *Int. J. General Syst.* 4:143–162.
5. Procyk, T. J., and E. H. Mamdani. 1979. A linguistic self-organizing process controller. *Automatica* 15(1):15–30.
6. Takagi, T., and M. Sugeno. 1983. Derivation of fuzzy control rules from human operator’s control actions. *Fuzzy Information Knowledge Representation Decision Analysis: IFAC Symposium Proceedings*. Marseilles, France. 55–60.

7. Rumelhart, D. E., G. E. Hinton, and R. J. Williams. 1986. Learning internal representations by error propagation. *Parallel Distributed Proc.* 1:318–362.
8. Lin, C. T., and C. S. G. Lee. 1991. Neural-network-based fuzzy logic control and decision system. *IEEE Trans. Comp.* 40(12):1320–1336.
9. Kosko, B. 1992. *Neural networks and fuzzy systems*. New York, NY: Prentice Hall. 456 p.
10. Lin, C. T., and C. S. G. Lee. 1994. Supervised and unsupervised learning with fuzzy similarity for neural network-based fuzzy logic control systems. *Fuzzy sets, neural networks, and soft computing*. Eds. R. R. Yager and L. A. Zadeh. New York, NY: Van Nostrand Reinhold. 85–125.
11. Criminal Code of the Azerbaijan Republic approved by the Act of Azerbaijan Republic from December 30, 1999, No. 787-IQ. Enured on September 1, 2000, under the Law of Azerbaijan Republic on May 26, 2000, No. 886-IR. Available at: <http://www.legislationonline.com> (accessed May 15, 2015).
12. Rzaev, R. R. 2013. *Intellectual'niy analiz dannykh v sistemakh podderzhki prinyatiya resheniy* [Intellectual analyses in decision support systems]. Saarbrucken, Germany: LAP Lambert Academic Publishing GmbH & Co. KG. 130 p.

Received September 22, 2015

Contributors

Rzayev Ramin R. (b. 1961) — Doctor of Science in technology, professor, principal scientist, Institute of Control Systems of the Azerbaijan National Academy of Sciences, 9 B. Vahabzadeh Str., Baku AZ1141, Azerbaijan Republic; raminrza@yahoo.com

Agayev Faig B. (b. 1986) — Chief of the Legal education, scientific-analytical, Information and International Relations Department of the Office of the Commissioner for Human Rights (Ombudsman) of the Republic of Azerbaijan, 40 Uzeyir Hajibeyov str., Baku AZ1148, Azerbaijan Republic; ombudsman@ombudsman.gov.az

Goyushov Aydin I. (b. 1978) — doctorate student, Institute of Control Systems of the Azerbaijan National Academy of Sciences, 9 B. Vahabzadeh Str., Baku AZ1141, Azerbaijan Republic; geyushev@hotmail.com

Jamalov Zeynal R. (b. 1959) — Candidate of Science (PhD) in physics and mathematics, Associate Professor, Baku State University, 23 Zakhid Khalilov Str., Baku AZ1148, Azerbaijan Republic; zjamalov@mail.ru

НОРМАЛЬНЫЕ И ОРТОГОНАЛЬНЫЕ СУБОПТИМАЛЬНЫЕ ФИЛЬТРЫ ДЛЯ НЕЛИНЕЙНЫХ СТОХАСТИЧЕСКИХ СИСТЕМ НА МНОГООБРАЗИЯХ*

*И. Н. Сеницын*¹

Аннотация: Для стохастических систем (СтС) на гладких многообразиях (МСтС) с винеровскими и пуассоновскими шумами в уравнениях состояния и винеровскими шумами в наблюдениях разработана теория синтеза нормальных и ортогональных субоптимальных фильтров (СОФ) по среднеквадратическому критерию. Получены точные фильтрационные уравнения для МСтС. Особое внимание уделено модифицированным фильтрам на основе ненормированных распределений. Обсуждаются вопросы упрощения точных фильтрационных уравнений. Приводятся уравнения субоптимальных фильтров на основе методов нормальной аппроксимации (МНА) и статистической линеаризации (МСЛ). Для решения задач в реальном времени использование нормальных СОФ не обеспечивает необходимой точности, поэтому в основу синтеза положены методы ортогональных разложений (МОР) и квазимоментов (МКМ) для апостериорной одномерной плотности. Получены уравнения точности и чувствительности алгоритмов. В качестве тестовых примеров рассмотрены фильтры для одномерных нелинейных СтС с аддитивным и мультипликативным белым шумом. Даны некоторые обобщения разработанных алгоритмов.

Ключевые слова: апостериорное одномерное распределение; винеровский шум; квазимомент (КМ); коэффициент ортогонального разложения (КОР); метод квазимоментов (МКМ); метод нормальной аппроксимации (МНА); метод ортогональных разложений (МОР); модифицированный МНА (ММНА); модифицированный МОР (ММОР); нормальный фильтр; ортогональный СОФ (ОСОФ); первая функция чувствительности; пуассоновский шум; стохастическая система на многообразиях (МСтС); субоптимальный фильтр (СОФ)

DOI: 10.14357/08696527160113

1 Введение

Методы ортогональных разложений и квазимоментов в [1, 2] были развиты для аналитического моделирования одно- и многомерных распределений в МСтС

* Работа выполнена при поддержке РФФИ (проект 15-07-02244).

¹ Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, sinitsin@dol.ru

и дано их применение для задач надежности и безопасности технических систем. В [3] представлена теория СОФ на базе МНА и МСЛ, а также МОР и МКМ для МСтС с винеровскими шумами в уравнениях наблюдений и винеровскими и пуассоновскими шумами в уравнениях состояния. В основу СОФ были положены точные нелинейные уравнения для апостериорного одномерного распределения.

Рассмотрим развитие [3], положив в основу точные линейные уравнения для ненормированного апостериорного одномерного распределения. В разд. 2 получены точные уравнения для нормированного и ненормированного апостериорных одномерных распределений. Раздел 3 содержит сведения о приближенных методах нелинейной фильтрации. Особое внимание в разд. 3 уделено СОФ на базе МНА, МСЛ, а также модифицированного МНА (ММНА) на основе ненормированного распределения. В разд. 4 получены уравнения ортогональных СОФ (ОСОФ) на базе МОР и МКМ, а в разд. 5 — уравнения модифицированных ОСОФ (МОСОФ) на базе модифицированного МОР (ММОР) для ненормированных распределений. Вопросы точности и чувствительности ОСОФ и МОСОФ рассматриваются в разд. 6. Тестовый пример дан в разд. 7. Заключение содержит выводы и некоторые обобщения.

2 Точные уравнения для апостериорного одномерного распределения

2.1 Уравнения процессов. Вспомогательные формулы

Как известно [4–8], на практике часто возникают задачи непрерывного определения состояния системы по результатам непрерывных наблюдений. Так как наблюдения всегда сопровождаются случайными ошибками, то следует говорить не об определении состояния системы, а о его оценивании (фильтрации, экстраполяции, интерполяции и т. д.) путем статистической обработки результатов наблюдений. Будем рассматривать задачи фильтрации состояния систем, моделями которых могут служить стохастические дифференциальные уравнения с винеровскими и пуассоновскими шумами.

Замечание 2.1. В некоторых случаях стохастические дифференциальные уравнения модели изучаемой системы могут иметь неизвестные параметры и, как правило, всегда содержат параметры, известные с ограниченной точностью. Поэтому возникает задача непрерывного оценивания неизвестных параметров системы (точнее, ее модели) по результатам непрерывных наблюдений. Предположим, что правые части уравнений зависят от конечного множества неизвестных параметров, которые будем рассматривать как компоненты вектора параметров θ . Одним из возможных подходов в таких случаях является следующий прием: неизвестный векторный параметр θ считают стохастическим процессом $\Theta = \Theta_t$, который определяется дифференциальным уравнением $\dot{\Theta}_t = 0$, и включают компоненты этого векторного процесса в вектор состояния системы («расширяют» вектор состояния путем включения в него неизвестных параметров в качестве дополнительных компонент). Таким образом, задача непрерывного оценивания

неизвестных параметров модели системы сводится к задаче непрерывного оценивания состояния системы с расширенным вектором состояния. От неизвестных параметров могут зависеть и уравнения наблюдения. Эти параметры следует включить в вектор θ и, следовательно, в расширенный вектор состояния.

Пусть векторный стохастический процесс (СтП) $[X_t^T Y_t^T]^T$ определяется системой векторных стохастических дифференциальных уравнений Ито:

$$dX_t = \varphi(X_t, Y_t, \Theta, t) dt + \psi'(X_t, Y_t, \Theta, t) dW_0 + \int_{R_0^q} \psi''(X_t, Y_t, \Theta, t, v) P^0(dt, dv), \quad X(t_0) = X_0; \quad (1)$$

$$dY_t = \varphi_1(X_t, Y_t, \Theta, t) dt + \psi'_1(X_t, Y_t, \Theta, t) dW_0 + \int_{R_0^q} \psi''_1(X_t, Y_t, \Theta, t, v) P^0(dt, dv), \quad Y(t_0) = Y_0. \quad (2)$$

Здесь $Y_t = Y(t)$ — n_y -мерный наблюдаемый СтП, $Y_t \in \Delta^y$ (Δ^y — гладкое многообразие наблюдений); $X_t = X(t)$ — n_x -мерный ненаблюдаемый СтП (вектор состояния), $X_t \in \Delta^x$ (Δ^x — гладкое многообразие состояний); $W_0 = W_0(t)$ — n_w -мерный винеровский СтП ($n_w \geq n_y$) интенсивности $\nu_0 = \nu_0(\Theta, t)$; $P^0(\Delta, A) = P(\Delta, A) - \mu_P(\Delta, A)$, $P(\Delta, A)$ представляет собой для любого множества A простой пуассоновский СтП, а $\mu_P(\Delta, A)$ — его математическое ожидание, причем

$$\mu_P(\Delta, A) = MP(\Delta, A) = \int_{\Delta} \nu_P(\tau, A) d\tau;$$

$\nu_P(\Delta, A)$ — интенсивность соответствующего пуассоновского потока событий, $\Delta = (t_1, t_2]$; интегрирование по v распространяется на все пространство R^q с выколотым началом координат; Θ — вектор случайных параметров размерности n_Θ ; $\varphi = \varphi(X_t, Y_t, \Theta, t)$, $\varphi_1 = \varphi_1(X_t, Y_t, \Theta, t)$; $\psi' = \psi'(X_t, Y_t, \Theta, t)$ и $\psi'_1 = \psi'_1(X_t, Y_t, \Theta, t)$ — известные функции, отображающие $R^{n_x} \times R^{n_y} \times R$ соответственно в R^{n_x} , R^{n_y} , $R^{n_x n_w}$ и $R^{n_y n_w}$; $\psi'' = \psi''(X_t, Y_t, \Theta, t, v)$ и $\psi''_1 = \psi''_1(X_t, Y_t, \Theta, t, v)$ — известные функции, отображающие $R^{n_x} \times R^{n_y} \times R^q$ в R^{n_x} и R^{n_y} . Требуется найти оценку \hat{X}_t СтП X_t в каждый момент времени t по результатам наблюдения СтП $Y(\tau)$ до момента t , $Y_{t_0}^t = \{Y(\tau) : t_0 \leq \tau < t\}$.

Предположим, что

- уравнение состояния имеет вид (1);
- уравнение наблюдения (2), во-первых, не содержит пуассоновского шума ($\psi''_1 \equiv 0$), а во-вторых, коэффициент при винеровском шуме ψ'_1 в уравнениях наблюдения не зависит от состояния ($\psi'_1(X_t, Y_t, \Theta, t) = \psi'_1(Y_t, \Theta, t)$).

В этом случае уравнения задачи нелинейной фильтрации имеют следующий вид:

$$dX_t = \varphi(X_t, Y_t, \Theta, t) dt + \psi'(X_t, Y_t, \Theta, t) dW_0 + \\ + \int_{R_0^q} \psi''(X_t, Y_t, \Theta, t, v) P^0(dt, dv), \quad X(t_0) = X_0; \quad (3)$$

$$dY_t = \varphi_1(X_t, Y_t, \Theta, t) dt + \psi_1(Y_t, \Theta, t) dW_0, \quad Y(t_0) = Y_0. \quad (4)$$

Будем считать, что выполнены условия существования и единственности СтП $[X_t^T \ Y_t^T]^T$, определяемого (3) и (4) при соответствующих начальных условиях [4, 5, 7].

В дальнейшем для стохастического уравнения

$$dZ = a dt + b dW_0 + \int_{R_0^q} c P^0(dt, dv) \quad (5)$$

потребуется обобщенная формула Ито [7, 8] для дифференциала нелинейной функции $U = U(Z, t)$:

$$dU = \left\{ U_t + U_z^T a + \frac{1}{2} \text{tr} [U_{zz} b \nu b^T] \right\} dt + \\ + \int_{R_0^q} [U(Z + c, t)^T - U(Z, t)^T - U_z^T c] \mu_P(dt, dv) + \\ + U_Z^T b dW_0 + \int_{R_0^q} [U(Z + c, t) - U(Z, t)] P^0(dt, dv). \quad (6)$$

Здесь a , b и c — известные функции Z и t .

2.2 Фильтрационные уравнения для нормированного распределения

Как известно [4–8], для любых СтП X_t и Y_t оптимальная оценка \hat{X}^t , минимизирующая средний квадрат ошибки в каждый момент времени t , представляет собой апостериорное математическое ожидание СтП X_t : $\hat{X}^t = \mathbf{M} [X_t | Y_{t_0}^t]$. Чтобы найти это условное математическое ожидание, необходимо знать $p_t = p_t(x)$ — апостериорное одномерное распределение СтП X_t .

В основе уравнений оптимальной (в смысле минимума средней квадратической ошибки) фильтрации для уравнений (3) и (4) в силу (6) лежит следующая формула для стохастического дифференциала апостериорного математического ожидания скалярной функции $f = f(X, t)$ вектора состояния [3]:

$$\begin{aligned}
 d\hat{f} = & dM_{\Delta x}^{p_t} \left[f_t(X, t) + f_x(X, t)^T \varphi(X, Y, t) + \frac{1}{2} \operatorname{tr} \left\{ f_{xx}(X, t) (\psi' \nu_0 \psi'^T)(X, Y, t) \right\} + \right. \\
 & \left. + \int_{R_0^g} \left\{ f(X + \psi'', t) - f(X, t) - f_x(X, t)^T \psi''(X, Y, t) \right\} \nu_P(t, dv) \mid Y_{t_0}^t \right] dt + \\
 & + M_{\Delta x}^{p_t} \left\{ f(X, t) \left[\varphi_1(X, Y, t)^T - \hat{\varphi}_1^T \right] + f_x(X, t)^T (\psi \nu_0 \psi_1^T)(X, Y, t) \mid Y_{t_0}^t \right\} \times \\
 & \times \psi_1 \nu_0 \psi_1^T)^{-1}(Y, t) (dY - \hat{\varphi}_1 dt). \quad (7)
 \end{aligned}$$

Здесь для краткости аргумент Θ опущен; $X = X_t$, $Y = Y_t$, $\nu = \nu_0$ и ν_P — интенсивности W_0 и P^0 ; $\hat{\varphi}_1$ — апостериорное математическое ожидание φ_1 при заданной условной плотности $p_t = p_t(x, \Theta)$:

$$\hat{\varphi}_1 = M_{\Delta x}^{p_t} [\varphi_1(X, Y, t)].$$

Полагая в (5) $f(X, t) \equiv g_t(\lambda, \Theta) = M_{\Delta x}^{p_t} [\exp(i\lambda^T X)]$, получим точное нелинейное фильтрационное уравнение для характеристической функции $g_t(\lambda, \Theta)$:

$$\begin{aligned}
 dg_t(\lambda, \Theta) = & M_{\Delta x}^{p_t} \left[\left\{ i\lambda^T \varphi(X_t, Y_t, \Theta, t) - \right. \right. \\
 & \left. \left. - \frac{1}{2} \lambda^T (\psi \nu_0 \psi^T)(X_t, Y_t, \Theta, t) \lambda + \gamma(\lambda, X_t, Y_t, \Theta, t) \right\} e^{i\lambda^T X_t} \mid Y_{t_0}^t \right] dt + \\
 & + M_{\Delta x}^{p_t} \left[\left\{ \varphi_1(X_t, Y_t, \Theta, t)^T - \hat{\varphi}_1^T + i\lambda^T (\psi \nu_0 \psi_1^T)(X_t, Y_t, \Theta, t) e^{i\lambda^T X_t} \mid Y_{t_0}^t \right\} \times \right. \\
 & \left. \times (\psi_1 \nu_0 \psi_1^T)^{-1}(Y_t, \Theta, t) (dY_t - \hat{\varphi}_1 dt) \right], \quad (8)
 \end{aligned}$$

где

$$\begin{aligned}
 \gamma = & \gamma(\lambda, X_t, Y_t, \Theta, t) = \\
 & = \int_{R_0^g} \left[e^{i\lambda^T \psi''(X_t, Y_t, \Theta, t, v)} - 1 - i\lambda^T \psi''(X_t, Y_t, \Theta, t, v) \right] \nu_P(\Theta, t, v) dv. \quad (9)
 \end{aligned}$$

Функции $g_t(\lambda, \Theta)$ и $p_t(x, \Theta)$ связаны между собой преобразованием Фурье [4].

Отсюда для гауссовской МСтС (3), (4) ($\psi'' \equiv 0$) уравнение (7) при $\gamma = 0$ упрощается и приобретает вид:

$$\begin{aligned}
 dg_t(\lambda, \Theta) = & \\
 = & M_{\Delta x}^{p_t} \left[\left\{ i\lambda^T \varphi(X_t, Y_t, \Theta, t) - \frac{1}{2} \lambda^T (\psi \nu_0 \psi^T)(X_t, Y_t, \Theta, t) \lambda \right\} e^{i\lambda^T X_t} \mid Y_{t_0}^t \right] dt +
 \end{aligned}$$

$$+ \mathbf{M}_{\Delta^x}^{p_t} \left[\left\{ \varphi_1 (X_t, Y_t, \Theta, t)^T - \hat{\varphi}_1^T + i\lambda^T (\psi\nu_0\psi_1^T) (X_t, Y_t, \Theta, t) e^{i\lambda^T X_t} \mid Y_{t_0}^t \right\} \times \right. \\ \left. \times (\psi_1\nu_0\psi_1^T)^{-1} (Y_t, \Theta, t) (dY_t - \hat{\varphi}_1 dt) \right]. \quad (10)$$

2.3 Фильтрационные уравнения для ненормированного распределения

Предположим дополнительно к условиям существования и единственности решения (3) и (4) условия существования непрерывных первых производных функций φ и ψ_1 , а также вторых производных функций ψ' и ψ'' по компонентам X_t . Тогда, применяя формулу (7) к ненормированной апостериорной характеристической функции $\tilde{d}_t(\lambda)$, получим:

$$d\tilde{g}_t(\lambda) = \tilde{\mathbf{M}}_{\Delta^x}^{\tilde{p}_t} \left\{ i\lambda^T \varphi (Y_t, X_t, \Theta, t) - \frac{1}{2} \lambda^T (\psi' \nu_0 \psi'^T) (Y_t, X_t, \Theta, t) \lambda + \right. \\ \left. + \gamma (\lambda, X_t, Y_t, \Theta, t) e^{i\lambda^T X_t} \mid Y_{t_0}^t \right\} + \\ + \tilde{\mathbf{M}}_{\Delta^x}^{\tilde{p}_t} \left[\varphi_1 (Y_t, X_t, \Theta, t)^T + i\lambda^T (\psi' \nu_0 \psi_1^T) (Y_t, X_t, \Theta, t) e^{i\lambda^T X_t} \mid Y_{t_0}^t \right] \times \\ \times (\psi'_1 \nu_0 \psi_1'^T)^{-1} (Y_t, \Theta, t) dY_t, \quad (11)$$

где $\gamma(\lambda, X_t, Y_t, \Theta, t)$ определена (9);

$$\hat{\varphi}_1 = \tilde{\mathbf{M}}_{\Delta^x}^{\tilde{p}_t} \varphi (X_t, Y_t, \Theta, t).$$

Кроме того, здесь введены следующие обозначения: $p_t(x)$ и $\tilde{p}_t(x)$ — апостериорные плотности, связанные с апостериорными характеристическими функциями $g_t(\lambda)$ и $\tilde{g}_t(\lambda)$ соотношениями:

$$\tilde{g}_t(\lambda) = \mu_t \mathbf{M}_{\Delta^x}^{p_t} e^{i\lambda^T X_t} = \mu_t g_t(\lambda); \quad \tilde{p}_t(x) = \mu_t p_t(x).$$

Функция μ_t не зависит от $X_t = x$ (но, конечно, может зависеть от Y_t и $Y_{t_0}^t$) и определяется дифференциальным уравнением

$$d\mu_t = \tilde{\varphi} dt + \tilde{\psi} dY_t, \quad (12)$$

где $\tilde{\varphi}$ и $\tilde{\psi}$ — некоторые функции Y_t и $Y_{t_0}^t$; $\mathbf{M}_{\Delta^x}^{p_t}$ и $\tilde{\mathbf{M}}_{\Delta^x}^{\tilde{p}_t}$ — операторы математических ожиданий.

Полагая в (11) $\psi'' = 0$ и $\gamma = 0$, приходим к соответствующему уравнению для гауссовской МСтС (3), (4):

$$\begin{aligned}
 d\tilde{g}_t = & M_{\Delta^x}^{\tilde{p}_t} \left[i\lambda^T \varphi(Y_t, X_t, \Theta, t) - \frac{1}{2} \lambda^T \left(\psi' \nu_0 \psi'^T \right) (Y_t, X_t, \Theta, t) \lambda + \right. \\
 & \left. + \gamma(\lambda, X_t, Y_t, \Theta, t) e^{i\lambda^T X_t} \mid Y_{t_0}^t \right] dt + \\
 & + M_{\Delta^x}^{\tilde{p}_t} \left[\varphi_1(Y_t, X_t, \Theta, t)^T + i\lambda^T \left(\psi'_0 \nu_0 \psi'_1{}^T \right) (Y_t, X_t, \Theta, t) e^{i\lambda^T X_t} \mid Y_{t_0}^t \right] \times \\
 & \times \left(\psi'_1 \nu_0 \psi'_1 \right)^{-1} (Y_t, \Theta, t) dY_t. \quad (13)
 \end{aligned}$$

2.4 Частный случай уравнений (3), (4)

Если функция ψ'' в (3) допускает представление

$$\psi'' = \psi' \omega(\Theta, v), \quad (14)$$

где $P^0(\Delta, A) = P^0((0, t], dv)$, то уравнения (3), (4) примут следующий вид:

$$\left. \begin{aligned}
 \dot{X}_t = & \varphi(X_t, Y_t, \Theta, t) + \psi'(X_t, Y_t, \Theta, t) V(\Theta, t), \quad X(t_0) = X_0; \\
 \dot{Y}_t = & \varphi(X_t, Y_t, \Theta, t) + \psi_1(Y_t, \Theta, t) V_0(\Theta, t), \quad Y(t_0) = Y_0.
 \end{aligned} \right\} \quad (15)$$

Здесь $V_0(\Theta, t) = \dot{W}_0(\Theta, t)$; $V(\Theta, t) = \dot{W}(\Theta, t)$,

$$\bar{W}(\Theta, t) = W_0(\Theta, t) + \int_{R_0^q} \omega(\Theta, v) P^0((0, t], dv);$$

$\nu_P(\Theta, t, v) dv = [\partial \mu(\Theta, t, v) / \partial t] dv$ — интенсивность пуассоновского потока скачков, равных $\omega(\Theta, t)$. При этом логарифмические производные от одномерных характеристических функций определяются известными формулами:

$$\begin{aligned}
 \chi^{W_0}(\rho; t) &= -\frac{1}{2} \rho^T \nu_0(\Theta, t) \rho; \\
 \chi^{\bar{W}}(\rho; t) &= -\frac{1}{2} \rho^T(\Theta, t) \rho^T + \int_{R_0^q} \left[e^{i\rho^T \omega(\Theta, v)} - 1 - i\rho^T \omega(\Theta, v) \right] \nu_P(\Theta, t, v) dv.
 \end{aligned}$$

В таком случае уравнение для апостериорной одномерной характеристической функции имеет вид (8), где функция (9) допускает следующую запись:

$$\begin{aligned}
 \gamma = & \int_{R_0^q} \left[e^{i\lambda^T \psi'(X_t, Y_t, \Theta, t) \omega(\Theta, v)} - 1 - \right. \\
 & \left. - i\lambda^T \psi'(X_t, Y_t, \Theta, t) \omega(\Theta, v) \right] \nu_P(\Theta, t, v) dv. \quad (16)
 \end{aligned}$$

Замечание 2.2. Этот случай для нормированного распределения рассмотрен в [9], а для ненормированного распределения в [10].

Таким образом, можно сформулировать следующие утверждения разд. 2.

Теорема 2.1. Пусть для МСтС (3), (4) выполнены условия существования и единственности решения, а матрица $\sigma_1 = \psi_1 \nu_0 \psi_1^T$ не вырождена. Тогда при условии ограниченности соответствующих математических ожиданий точное фильтрационное уравнение для условной одномерной нормированной характеристической функций имеет вид (8).

Теорема 2.2. В условиях теоремы 2.1 при отсутствии пуассоновских шумов точное фильтрационное уравнение для условной одномерной нормированной характеристической функции имеет вид (10).

Теорема 2.3. Если в условиях теоремы 2.1 дополнительно предположить существование непрерывных первых производных функций φ и ψ_1 и вторых производных функций ψ' и ψ'' по компонентам X_t , то при условии ограниченности соответствующих математических ожиданий точное фильтрационное уравнение для ненормированной условной одномерной характеристической функции будет иметь вид (11).

Теорема 2.4. В условиях теоремы 2.3 при отсутствии пуассоновских шумов точное фильтрационное уравнение для ненормированной условно одномерной характеристической функции будет иметь вид (13).

Теорема 2.5. Пусть для МСтС (15) выполнены условия существования и единственности решения, имеет место представление (14), а матрица $\sigma_1 = \psi_1 \nu_0 \psi_1^T$ не вырождена. Тогда при условии ограниченности соответствующих математических ожиданий точное фильтрационное уравнение имеет вид (8) при условии (16).

Теорема 2.6. Если в условиях теоремы 2.5 дополнительно предположить существование непрерывных первых производных функций φ и ψ_1 , а также вторых производных функций ψ' и ψ'' по компонентам X_t , то при условии ограниченности соответствующих математических ожиданий точное фильтрационное уравнение для ненормированного одномерного распределения будет иметь вид (11) при условии (16).

3 Приближенные методы нелинейной фильтрации. Субоптимальные фильтры на основе метода нормальной аппроксимации (статистической линеаризации) и модифицированного метода нормальной аппроксимации

3.1 Вводные замечания

Точное решение фильтрационных уравнений возможно только в случаях, когда уравнения гауссовской дифференциальной МСтС линейны или линейны

лишь относительно вектора состояния X_t при независимой от состояния функции ψ . Эти уравнения дают точное решение задачи оптимальной нелинейной фильтрации. Однако это решение не может быть реализовано практически. Для нахождения оптимальной оценки вектора состояния необходимо решить фильтрационное уравнение для апостериорной характеристической функции (или фильтрационное уравнение для апостериорной плотности вектора состояния X_t) после получения результатов наблюдений, затем вычислить оптимальную оценку вектора X_t . Но методов точного решения этих уравнений в общем случае пока еще не существует.

Численное решение фильтрационных уравнений в задачах реального времени (или онлайн-оценивания) тоже невозможно, так как для этого требуется много времени, а решать их необходимо каждый раз после получения результатов наблюдений. Кроме того, практическое применение точной теории оптимальной нелинейной фильтрации имеет смысл только в тех случаях, когда оценки можно вычислять в реальном масштабе времени по мере получения результатов наблюдений. Точная теория дает оптимальные оценки в каждый момент t по результатам наблюдений, полученным к этому моменту, без использования последующих результатов наблюдений. Если эти оценки не могут быть вычислены в тот же момент t или хотя бы с фиксированным приемлемым запаздыванием и их вычисление приходится откладывать на будущее, то нет никакого смысла отказываться от использования наблюдений, получаемых после момента t , для оценивания состояния системы в момент t . Поэтому для статистической обработки результатов после окончания наблюдений, т. е. для офлайн-оценивания, целесообразно применять известные из математической статистики методы постобработки информации [5].

Необходимость обработки результатов наблюдений в реальном масштабе времени непосредственно в процессе эксперимента привела к появлению ряда приближенных методов оптимальной нелинейной фильтрации, называемых обычно методами *субоптимальной фильтрации* [6–8]. Одни приближенные методы основаны на приближенном решении фильтрационных уравнений, а другие — на превращении формул для стохастических дифференциалов оптимальной оценки \hat{X}_t и апостериорной ковариационной матрицы ошибки R_t в стохастические дифференциальные уравнения для \hat{X}_t и R_t путем разложения функций φ , φ_1 и ψ_1 или φ , φ_1 , $\psi'\psi''$, ψ и ψ_1 в степенные ряды и отбрасывания остаточных членов.

Для приближенного решения уравнения для апостериорной одномерной характеристической функции $g_1(\lambda, \Theta)$ вектора X_t можно использовать методы, основанные на параметризации одномерных распределений СтП, определяемого стохастическим дифференциальным уравнением [6–8]. Эти методы позволяют изучить стохастические дифференциальные уравнения для параметров апостериорного распределения. Простейшим таким методом является МНА апостериорного распределения. Исключительно важное практическое значение имеют

квазилинейные фильтры, получаемые с помощью методов эквивалентной линеаризации [6–8].

3.2 Нормальный субоптимальный фильтр для гауссовских систем

Так как нормальное (гауссовское) распределение, аппроксимирующее апостериорное одномерное распределение X_t , полностью определяется математическим ожиданием \hat{X}_t и ковариационной матрицей R_t вектора X_t , то при аппроксимации апостериорного одномерного распределения вектора X_t нормальным распределением все математические ожидания в правых частях формул для $d\hat{X}_t$ и dR_t будут определенными функциями \hat{X}_t , R_t и t . Для гауссовских МСТС ($\psi'' = 0$ и $\psi_1'' = 0$), пользуясь формулой (7), можно показать, что фильтрационные уравнения для нормального СОФ (НСОФ) будут представлять собой стохастические дифференциальные уравнения, определяющие \hat{X}_t и R_t :

$$\left. \begin{aligned} d\hat{X}_t &= f\left(\hat{X}_t, Y_t, R_t, \Theta, t\right) dt + \\ &\quad + h\left(\hat{X}_t, Y_t, R_t, \Theta, t\right) \left[dY_t - f^{(1)}\left(\hat{X}_t, Y_t, R_t, \Theta, t\right) dt\right]; \\ dR_t &= \left\{ f^{(2)}\left(\hat{X}_t, Y_t, R_t, \Theta, t\right) - \right. \\ &\quad \left. - h\left(\hat{X}_t, Y_t, R_t, \Theta, t\right) (\psi_1 \nu_0 \psi_1^T) (Y_t, \Theta, t) h\left(\hat{X}_t, Y_t, R_t, \Theta, t\right)^T \right\} dt + \\ &\quad + \sum_{r=1}^{n_y} \rho_r \left(\hat{X}_t, Y_t, R_t, \Theta, t\right) \left[dY_r - f_r^{(1)}\left(\hat{X}_t, Y_t, R_t, \Theta, t\right) dt\right]. \end{aligned} \right\} (17)$$

Здесь введены следующие обозначения:

$$\left. \begin{aligned} f &= f\left(\hat{X}_t, Y_t, R_t, t\right) = M_{\Delta^x}^N [\varphi(Y_t, X_t, \Theta, t)] = \hat{\varphi}; \\ f^{(1)} &= f^{(1)}\left(\hat{X}_t, Y_t, R_t, \Theta, t\right) = \\ &= \left\{ f_r^{(1)}\left(\hat{X}_t, Y_t, R_t, \Theta, t\right) \right\} = M_{\Delta^x}^N [\varphi_1(Y_t, X, \Theta, t)] = \hat{\varphi}_1^T; \\ h &= h\left(\hat{X}_t, Y_t, R_t, t\right) = \\ &= \left\{ M_{\Delta^x}^N \left[\hat{X}_t \varphi_1(Y_t, X_t, \Theta, t)^T + \psi \nu_0 \psi_1^T(Y_t, X_t, \Theta, t) \right] - \right. \\ &\quad \left. - \hat{X}_t f^{(1)}\left(\hat{X}_t, Y_t, R_t, \Theta, t\right)^T \right\} (\psi_1 \nu_0 \psi_1^T)^{-1} (Y_t, \Theta, t); \end{aligned} \right\} (18)$$

$$\rho_r = \rho_r \left(\hat{X}_t, Y_t, R_t, \Theta, t \right) = \mathbf{M}_{\Delta x}^N \left\{ \left(X_t - \hat{X}_t \right) \left(X_t^T - \hat{X}_t^T \right) a_r \left(Y_t, X_t, \Theta, t \right) + \left(X_t - \hat{X}_t \right) b_r \left(Y_t, X_t, \Theta, t \right)^T \left(X_t^T - \hat{X}_t^T \right) + b_r \left(Y_t, X_t, \Theta, t \right) \left(X_t^T - \hat{X}_t^T \right) \right\} \quad (r = 1, \dots, n_y), \quad (19)$$

$$f^{(2)} = f^{(2)} \left(\hat{X}_t, Y_t, R_t, \Theta, t \right) = \mathbf{M}_{\Delta x}^N \left\{ \left(X_t - \hat{X}_t \right) \varphi \left(Y_t, X_t, \Theta, t \right)^T + \varphi \left(Y_t, X_t, \Theta, t \right) \left(X_t^T - \hat{X}_t^T \right) + \psi \nu_0 \psi^T \left(Y_t, X_t, \Theta, t \right) \right\}; \quad (20)$$

где a_r — r -й элемент матрицы-строки $(\varphi_1^T - \hat{\varphi}_1^T)(\psi_1 \nu_0 \psi_1^T)$, а b_{kr} — элемент k -й строки и r -го столбца $(\psi \nu_0 \psi^T)(\psi_1 \nu_0 \psi_1^T)^{-1}$; b_r — r -й столбец матрицы $(\psi \nu_0 \psi^T)(\psi_1 \nu_0 \psi_1^T)$, $b_r = [b_{1r} \dots b_{n_x r}]^T$.

Количество уравнений для апостериорного одномерного распределения определяется по формуле:

$$Q_{\text{МНА}} = n_x + \frac{n_x(n_x + 1)}{2} = \frac{n_x(n_x + 3)}{2}.$$

За начальные значения \hat{X}_t и R_t при интегрировании уравнений (17), естественно, следует принять условные математическое ожидание и ковариационную матрицу величины X_0 относительно Y_0 :

$$\hat{X}_0 = \mathbf{M}_{\Delta x}^N [X_0 | Y_0]; \quad R_0 = \mathbf{M}_{\Delta x}^N \left[\left(X_0 - \hat{X}_0 \right) \left(X_0^T - \hat{X}_0^T \right) | Y_0 \right]. \quad (21)$$

Если нет информации об условном распределении X_0 относительно Y_0 , то начальные условия можно взять в виде:

$$\hat{X}_0 = \mathbf{M} X_0; \quad R_0 = \mathbf{M} (X_0 - \mathbf{M} X_0) (X_0^T - \mathbf{M} X_0^T).$$

Если же и об этих величинах нет никакой информации, то начальные значения \hat{X}_t и R_t приходится задавать произвольно.

3.3 Нормальный субоптимальный фильтр для негауссовских систем

В основе соответствующей теоремы для МСтС (3), (4) с пуассоновскими шумами в (3) и невырожденной матрицей $\sigma_1 = \psi_1 \nu_0 \psi_1^T$ лежат уравнения теоремы 2.3. При этом если учесть формулу (7), то потребуется ограниченность функций f , $f^{(1)}$, h и ρ_r , определяемых (19) и (18), и функции

$$\bar{f}^{(2)} = f^{(2)} + \mathbf{M}_{\Delta x}^N \left[\int_{R_0^q} \psi'' \psi''^T \nu_P(\Theta, t, dv) \right].$$

Замечание 3.1. Для гладких функций φ , φ_1 , ψ' и ψ'_1 и гауссовских МСтС (3), (4) СОФ на основе МНА называется просто гауссовским фильтром [4–6].

3.4 Модифицированные нормальные субоптимальные фильтры

Обобщая [8] и повторяя соответствующие выкладки для ненормированного апостериорного одномерного распределения:

$$\left. \begin{aligned} \tilde{p}_t(x) &\approx \tilde{g}_t^*(x) = \frac{\mu_t}{\sqrt{((2\pi)^n |R_t|)}} \exp \left\{ -\frac{1}{2} (x^\top - \hat{X}_t^\top) R_t^{-1} (x - \hat{X}_t) \right\}; \\ \tilde{g}_t(\lambda) &\approx \tilde{g}_t^*(\lambda) = \mu_t \exp \left\{ i\lambda^\top \hat{X}_t^\top - \frac{1}{2} \lambda^\top R_t \lambda \right\}, \end{aligned} \right\} (22)$$

получаем уравнения НСОФ

$$\hat{X}_t = \frac{m_t}{\mu_t}; \quad R_t = \frac{\Gamma_t}{\mu_t} - \frac{m_t m_t^\top}{\mu_t^2}; \quad (23)$$

$$\left. \begin{aligned} dm_t &= \mu_t f_1(\hat{X}_t, Y_t, R_t, \Theta, t) dt + \mu h_1(\hat{X}_t, Y_t, R_t, \Theta, t) dY_t; \\ d\Gamma_t &= \mu_t f_2(\hat{X}_t, Y_t, R_t, \Theta, t) dt + \mu_t \sum_{r=1}^{n_y} h_{2r}(\hat{X}_t, Y_t, R_t, \Theta, t) dY_r; \\ d\mu_t &= \mu_t h_3(\hat{X}_t, Y_t, R_t, \Theta, t) dY_t. \end{aligned} \right\} (24)$$

Здесь введены следующие обозначения:

$$\left. \begin{aligned} \sigma_0 &= \psi \nu_0 \psi^\top; \quad \bar{\sigma}_0 = \sigma_0 + \int_{R_0^q} \psi'' \psi''^\top \nu_P(\Theta, t, dv); \\ \sigma_1 &= \psi \nu_0 \psi_1^\top; \quad \sigma_2 = \psi_1 \nu_0 \psi_1^\top, \end{aligned} \right\} (25)$$

где σ_{1s} — s -я строка матрицы σ_1 ; m_t — вектор с компонентами m_{e_1}, \dots, m_{e_n} ; e_n — n -мерный вектор, все компоненты которого равны нулю, кроме s -й, равной 1; Γ_t — матрица с элементами $m_{e_s + e_h}$ ($s, h = 1, \dots, n$). При этом

$$\left. \begin{aligned} f_1 &= f_1(\hat{X}_t, Y_t, R_t, \Theta, t) = M_{\Delta_x}^N [\varphi(Y_t, X, \Theta, t)]; \\ h_1 &= h_1(\hat{X}_t, Y_t, R_t, \Theta, t) = M_{\Delta_x}^N [\sigma_1(Y_t, X, \Theta, t) + X \varphi_1(Y_t, X, \Theta, t)]; \\ f_2 &= f_2(\hat{X}_t, Y_t, R_t, \Theta, t) = \\ &= M_{\Delta_x}^N \left[X \varphi(Y_t, X, \Theta, t)^\top + \varphi(Y_t, X, \Theta, t) X^\top + \right. \\ &\quad \left. + \bar{\sigma}_0(Y_t, X, R_t, \Theta, t) \right] \sigma_2(Y_t, \Theta, t)^{-1}; \end{aligned} \right\} (26)$$

$$\left. \begin{aligned} h_{2r} &= h_{2r}(\hat{X}_t, Y_t, R_t, \Theta, t) = M_{\Delta_x}^N \left[X X^T a_r(Y_t, X, \Theta, t) + \right. \\ &\quad \left. + X b_r(Y_t, X, \Theta, t)^T + b_r(Y_t, X, \Theta, t) X^T \right] \quad (r = 1, \dots, n_x); \\ h_3 &= h_3(\hat{X}_t, Y_t, R_t, \Theta, t) = M_{\Delta_x}^N \left[\varphi_1(Y_t, X, \Theta, t)^T \right] \sigma_2(Y_t, \Theta, t)^{-1}. \end{aligned} \right\} \quad (27)$$

Начальными условиями для (24) служат

$$m(t_0) = m_0 = \hat{X}_0; \quad \Gamma(t_0) = \Gamma_0 = R_0 + \hat{X}_0 \hat{X}_0^T; \quad \mu(t_0) = 1, \quad (28)$$

где \hat{X}_0 и R_0 — условное математическое ожидание и ковариационная матрица вектора X_0 относительно Y_0 .

Уравнения НСОФ на основе ММНА несколько проще, чем соответствующие уравнения на основе МНА при нормированном распределении, хотя и содержат на одно уравнение больше:

$$Q_{\text{ММНА}} = 1 + \frac{n_x(n_x + 3)}{2}.$$

3.5 Квазилинейный субоптимальный фильтр на основе метода статистической линеаризации

Для МСтС (1), (2) при $\psi' = \psi'(\Theta, t)$, $\psi'' = \psi''(\Theta, t, v)$, $\psi'_1 = \psi'_1(\Theta, t)$ и $\psi''_1 = \psi''_1(\Theta, t, v)$ (т. е. с аддитивными винеровскими и пуассоновскими шумами) уравнения НСОФ проще получаются, если нелинейные функции φ и φ_1 на основе гауссовского (нормального) распределения заменить на статистически линеаризованные [6–8]:

$$\left. \begin{aligned} \varphi &= \varphi(X_t, Y_t, \Theta, t) \approx \varphi_0 + k_x^\varphi (X_t - m_t^x) + k_y^\varphi (Y_t - m_t^y); \\ \varphi_1 &= \varphi_1(X_t, Y_t, \Theta, t) \approx \varphi_{10} + k_x^{\varphi_1} (X_t - m_t^x) + k_y^{\varphi_1} (Y_t - m_t^y), \end{aligned} \right\} \quad (29)$$

а затем использовать уравнения линейной фильтрации [6–8]. Входящие в (29) коэффициенты статистической линеаризации зависят от математических ожиданий, дисперсий и ковариаций

$$Z_t = \begin{bmatrix} X_t \\ Y_t \end{bmatrix}; \quad m_t^z = \begin{bmatrix} m_t^x \\ m_t^y \end{bmatrix}; \quad K_t^z = \begin{bmatrix} K_t^x & K_t^{xy} \\ K_t^{xy} & K_t^y \end{bmatrix}.$$

Они определяются из уравнений:

$$\dot{Z}_t = A^z Z_t + A_0^z + B_0^z V, \quad V = \dot{W};$$

$$\left. \begin{aligned} \dot{m}_t^z &= A^z m_t^z + A_0^z, \quad m_{t_0}^z = m_0^z; \\ \dot{K}_t^z &= B^z K_t^z + K_t^z (B^z)^\top + B_0^z \nu^m (B_0^z)^\top, \quad K_{t_0}^z = K_0^z. \end{aligned} \right\} \quad (30)$$

Здесь введены следующие обозначения:

$$\begin{aligned} A_0^z &= \begin{bmatrix} a_0 \\ b_0 \end{bmatrix}; \quad A^z = \begin{bmatrix} a_1 & a \\ b_1 & b \end{bmatrix}; \quad B_0^z = \begin{bmatrix} \bar{\psi} \\ \bar{\psi}_1 \end{bmatrix}; \\ a &= k_y^\varphi; \quad a_1 = k_x^\varphi; \quad a_0 = \varphi_0 - k_x^\varphi m_t^x - k_y^\varphi m_t^y; \\ b &= k_y^{\varphi_1}; \quad b_1 = k_x^{\varphi_1}; \quad b_0 = \varphi_0 - k_x^{\varphi_1} m_t^x - k_y^{\varphi_1} m_t^y; \\ \psi dW_0 + \int_{R_0^g} \psi'' P^0(dt, dv) &= \bar{\psi} dW; \quad \psi'_1 dW_0 + \int_{R_0^g} \psi''_1 P^0(dt, dv) = \bar{\psi}_1 dW, \end{aligned} \quad (31)$$

где ν^W — интенсивность СтП с независимыми приращениями, состоящего из винеровской и пуассоновской частей (31). Тогда уравнения квазилинейного НСОФ будут иметь вид:

$$\dot{\hat{X}}_t = aY_t + a_1 \hat{X}_t + a_0 + \beta_t \left[Z_t - (bY_t + b_1 \hat{X}_t + b_0) \right]; \quad (32)$$

$$\beta_t = \left(R_t b_1^\top + \bar{\psi} \nu^W \bar{\psi}_1^\top \right) \left(\bar{\psi}_1 \nu^W \bar{\psi}_1^\top \right)^{-1}; \quad (33)$$

$$\begin{aligned} \dot{R}_t &= a_1 R_t + R_t a_1^\top + \bar{\psi} \nu^W \bar{\psi}^\top - \\ &- \left(R_t b_1^\top + \bar{\psi} \nu^W \bar{\psi}_1^\top \right) \left(\bar{\psi}_1 \nu^W \bar{\psi}_1^\top \right)^{-1} \left(b_1 R_t + \bar{\psi}_1 \nu^W \bar{\psi}^\top \right). \end{aligned} \quad (34)$$

Таким образом, приходим к следующим утверждениям разд. 3.

Теорема 3.1. Пусть выполнены условия теоремы 2.1. Тогда НСОФ для (3), (4) описывается уравнениями (17) и (21) при условиях ограниченности функций (18)–(20).

Теорема 3.2. Пусть выполнены условия теоремы 2.2. Тогда НСОФ для гауссовской системы (3) описывается уравнениями (17) и (21) при условиях ограниченности функций (18)–(20) при $\psi'' = 0$.

Теорема 3.3. Пусть выполнены условия теоремы 2.3. Тогда НСОФ для (3), (4) описывается уравнениями (22)–(24) и (28) при условиях ограниченности функций (25)–(27).

Теорема 3.4. Пусть выполнены условия теоремы 2.4. Тогда НСОФ для гауссовской системы (3) описывается уравнениями (22)–(24) и (28) при $\bar{\sigma}_0 = \sigma_0$ и ограниченности функций (25)–(27) и $\psi'' = 0$.

Теорема 3.5. Пусть МСтС (1), (2) содержит только аддитивные винеровские и пуассоновские шумы и допускает замену статистически линеаризованной,

а матрица $\sigma_1 = \bar{\psi}_1 \nu^W \bar{\psi}_1^T$ не вырождена. Тогда в основе алгоритма квазилинейного НСОФ лежат уравнения (32)–(34) при начальных условиях (21).

Замечание 3.2. Уравнения теоремы 2.3 сохраняют вид, если коэффициенты статистической линеаризации в (29) вычислять для известного эквивалентного (негауссовского) распределения. При этом уравнения (30), (32) и (34), как известно [6–8], имеют место для любого негауссовского СтП.

Замечание 3.3. Из теоремы 3.1 немедленно следуют уравнения НСОФ для фильтрации стационарных процессов в установившемся режиме для стационарных МСтС, если приравнять нулю правые части уравнений (30), (32) и (34).

4 Ортогональные субоптимальные фильтры на основе методов ортогональных разложений и квазиомоментов и модифицированного метода ортогональных разложений

4.1 Гауссовские системы. Ортогональный субоптимальный фильтр на основе метода ортогональных разложений

При аппроксимации апостериорной одномерной плотности отрезком ее ортогонального разложения [1, 2]

$$p_t(x, \Theta) \approx p^*(x; \Theta, \vartheta) = w(x; \Theta) \left[1 + \sum_{l=3}^N \sum_{|\nu|=l} c_\nu p_\nu(x) \right] \quad (35)$$

естественно принять за параметры, образующие вектор ϑ , апостериорные математическое ожидание \hat{X}_t , ковариационную матрицу R_t вектора X_t , а также коэффициенты ортогонального разложения (КОР) c_ν ($|\nu| = 3, \dots, N$). Здесь КОР определяется формулой

$$c_\nu = \left[q_\nu \left(\frac{\partial}{i \partial \lambda} \right) g_t(\lambda, \Theta) \right]_{\lambda=0}. \quad (36)$$

Заметим, что полином q_ν зависит от \hat{X}_t и R_t .

На основе (7) и (8) для гауссовской МСтС (3), (4) при $\psi'' = 0$ получим, что ОСОФ определяется следующими уравнениями:

$$\left. \begin{aligned} d\hat{X}_t &= f dt + h (dY_t - f^{(1)} dt) ; \\ dR_t &= \left(f^{(2)} - h \psi_1 \nu_0 \psi_1^T h^T \right) dt + \sum_{r=1}^{n_y} \rho_r \left(dY_r - f_r^{(1)} dt \right) . \end{aligned} \right\} \quad (37)$$

Здесь введены обозначения:

$$\left. \begin{aligned} f &= f(Y_t, \vartheta, \Theta, t) = \mathbf{M}_{\Delta x}^{p*} [\varphi(Y_t, X, \Theta, t)] ; \\ f^{(1)} &= \left\{ f_r^{(1)} \right\} = f^{(1)}(Y_t, \vartheta, \Theta, t) = \mathbf{M}_{\Delta x}^{p*} [\varphi_1(Y_t, X, \Theta, t)] ; \\ h &= h(Y_t, \vartheta, \Theta, t) = \left\{ \mathbf{M}_{\Delta x}^{p*} [X\varphi_1(Y_t, X, \Theta, t)^T + \right. \\ &\quad \left. + (\psi\nu_0\psi_1^T)(Y_t, X, \Theta, t)] - \hat{X}_t f^{(1)T} \right\} (\psi_1\nu_0\psi_1^T)^{-1}(Y_t, \Theta, t) ; \\ \rho_r &= \rho_r(Y_t, \vartheta, \Theta, t) = \mathbf{M}_{\Delta x}^{p*} \left[(X - \hat{X}_t) (X^T - \hat{X}_t^T) a_r(Y_t, X, \Theta, t) + \right. \\ &\quad \left. + (X - \hat{X}_t) b_r(Y_t, X, \Theta, t)^T + b_r(Y_t, X, \Theta, t) (X^T - \hat{X}_t^T) \right] \\ &\quad (r = 1, \dots, n_y) . \end{aligned} \right\} \quad (38)$$

$$\begin{aligned} f^{(2)} &= f^{(2)}(Y_t, \vartheta, \Theta, t) = \mathbf{M}_{\Delta x}^{p*} \left[(X - \hat{X}_t) \varphi(Y_t, X, \Theta, t)^T + \right. \\ &\quad \left. + \varphi(Y_t, X, \Theta, t) (X^T - \hat{X}_t^T) + (\psi\nu_0\psi^T)(Y_t, X, \Theta, t) \right] ; \end{aligned} \quad (40)$$

Далее перепишем (37) покоординатно:

$$\left. \begin{aligned} d\hat{X}_s &= f_s dt + h_s (dY_t - f^{(1)} dt) = A^{\hat{X}_s} dt + B^{\hat{X}_s} dY_t \quad (s = 1, \dots, n_x) ; \\ dR_{sq} &= \left(f_{sq}^{(2)} - h_s \psi_1 \nu_0 \psi_1^T h_q^T \right) dt + \eta_{sq} (dY_t - f^{(1)} dt) = \\ &= A^{R_{sq}} dt + B^{R_{sq}} dY_t , \end{aligned} \right\} \quad (41)$$

где $\hat{X}_s(t_0) = X_{s0}$, $R_{sq}(t_0) = R_{sq0}$, $s, q = 1, \dots, n_x$; η_{sq} — матрица-строка, элементами которой служат соответствующие элементы матрицы $\rho_1, \dots, \rho_{n_1}$:

$$\eta_{sq} = \eta_{e_s + e_q} = [\rho_{1sq} \cdots \rho_{msq}] \quad (s, q = 1, \dots, n_x) .$$

Здесь и далее для краткости индекс t сохраним только у Y_t . По формуле дифференцирования Ито для винеровского СтП, учитывая (41), находим в силу (36) стохастический дифференциал:

$$\begin{aligned} dc_\kappa &= \left[d \left\{ q_\kappa \left(\frac{\partial}{\partial \lambda} \right) g_t(\lambda, \Theta) \right\} \right]_{\lambda=0} = \\ &= \sum_{s=1}^{n_x} \left[\partial q_\kappa \left(\frac{\partial}{\partial \lambda} \right) \partial \hat{X}_s \cdot g_t(\lambda, \Theta) \right]_{\lambda=0} d\hat{X}_s + \\ &+ \sum_{s,u=1}^{n_x} \left[\partial q_\kappa \left(\frac{\partial}{\partial \lambda} \right) \partial R_{su} \cdot g_t(\lambda, \Theta) \right]_{\lambda=0} dR_{su} + \left[q_\kappa \left(\frac{\partial}{\partial \lambda} \right) dg_t(\lambda, \Theta) \right]_{\lambda=0} + \end{aligned}$$

$$\begin{aligned}
 & + \left\{ \frac{1}{2} \sum_{s,u=1}^{n_x} \left[\frac{\partial^2 q_\kappa(\partial/i\partial\lambda) \cdot g_t(\lambda, \Theta)}{\partial \hat{X}_s \partial \hat{X}_u} \right]_{\lambda=0} h_s \psi_1 \nu_0 \psi_1^T h_u^T + \right. \\
 & \quad + \frac{1}{2} \sum_{s,u,k,l=1}^{n_x} \left[\frac{\partial^2 q_\kappa(\partial/i\partial\lambda) \cdot g_t(\lambda, \Theta)}{\partial R_{su} \partial R_{kl}} \right]_{\lambda=0} \eta_{su} \psi_1 \nu_0 \psi_1^T \eta_{kl}^T + \\
 & \quad \left. + \sum_{s,k,l=1}^{n_x} \left[\frac{\partial^2 q_\kappa(\partial/i\partial\lambda) \cdot g_t(\lambda, \Theta)}{\partial \hat{X}_s \partial R_{kl}} \right]_{\lambda=0} h_s \psi_1 \nu_0 \psi_1^T \eta_{kl}^T \right\} dt.
 \end{aligned}$$

Подставив сюда выражения (41) и (7) дифференциалов $d\hat{X}_s$, dR_{sq} и $dg_t(\lambda, \Theta)$ и вспомнив, что для любого полинома $P(x)$ $P[(\partial/i\partial\lambda)g_t(\lambda)]_{\lambda=0} = P(\alpha)$, получаем стохастические дифференциальные уравнения:

$$\begin{aligned}
 dc_\kappa = & \left\{ F_\kappa + \sum_{s=1}^{n_x} \frac{\partial q_\kappa(\alpha)}{\partial \hat{X}_s} f_s + \sum_{s,u=1}^{n_x} \frac{\partial q_\kappa(\alpha)}{\partial R_{su}} \left(f_{su}^{(2)} - h_s \psi_1 \nu_0 \psi_1^T h_u^T \right) + \right. \\
 & + \frac{1}{2} \sum_{s,u=1}^{n_x} \frac{\partial^2 q_\kappa(\alpha)}{\partial \hat{X}_s \partial \hat{X}_u} h_s \psi_1 \nu_0 \psi_1^T h_u^T + \frac{1}{2} \sum_{s,u,k,l=1}^{n_x} \frac{\partial^2 q_\kappa(\alpha)}{\partial R_{su} \partial R_{kl}} \eta_{su} \psi_1 \nu_0 \psi_1^T \eta_{kl}^T + \\
 & \left. + \sum_{s,k,l=1}^{n_x} \frac{\partial^2 q_\kappa(\alpha)}{\partial \hat{X}_s \partial R_{kl}} h_s \psi_1 \nu_0 \psi_1^T \eta_{kl}^T \right\} dt + \left\{ H_\kappa + \sum_{s=1}^{n_x} \frac{\partial q_\kappa(\alpha)}{\partial \hat{X}_s} h_s + \right. \\
 & \left. + \sum_{s,u=1}^{n_x} \frac{\partial q_\kappa(\alpha)}{\partial R_{su}} \eta_{su} \right\} (dY_t - f^{(1)} dt) = A^{c_\kappa} dt + B^{c_\kappa} dY_t, \\
 c_\kappa(t_0) = & c_{\kappa 0} \quad (|\kappa| = 3, \dots, N). \quad (42)
 \end{aligned}$$

Здесь в дополнение к прежним обозначениям принято:

$$\left. \begin{aligned}
 F_\kappa = F_\kappa(Y_t, \Theta, \vartheta, t) = & \sum_{s=1}^{n_x} M_{\Delta^x}^{p^*} \left[\varphi_s(Y_t, X, \Theta, t) \frac{\partial q_\kappa(X)}{\partial X_s} \right] + \\
 & + \frac{1}{2} \sum_{s,u=1}^{n_x} M_{\Delta^x}^{p^*} \left[\sigma_{su}(Y_t, X, \Theta, t) \frac{\partial^2 q_\kappa(X)}{\partial X_s \partial X_u} \right], \\
 H_\kappa = H_\kappa(Y_t, \vartheta, \Theta, t) = & \left\{ M_{\Delta^x}^{p^*} \left[\varphi_1(Y_t, X, \Theta, t)^T q_\kappa(X) \right] + \right. \\
 & + \sum_{s=1}^{n_x} M_{\Delta^x}^{p^*} \left[\left(\psi \nu_0 \psi_1^T \right)_s(Y_t, X, \Theta, t) \frac{\partial q_\kappa(X)}{\partial X_s} \right] - \\
 & \left. - c_\kappa f^{(1)T} \right\} (\psi_1 \nu_0 \psi_1^T)^{-1}(Y_t, \Theta, t),
 \end{aligned} \right\} \quad (43)$$

где через $(\psi \nu_0 \psi_1^T)_s$ обозначена s -я строка матрицы $\psi \nu_0 \psi_1^T$; $\sigma = \psi \nu_0 \psi_1^T = \{\sigma_{su}\}$.

Функции f_s , $f^{(1)}$, $f_{su}^{(2)}$, h_s , η_{su} , F_κ и H_κ в уравнениях (41) и (42) представляют собой линейные комбинации величин c_ν ($|\nu| = 3, \dots, N$) с коэффициентами, зависящими от \hat{X}_t и R_t . Величины $\partial q_\kappa(\alpha)/\partial \hat{X}_s$, $\partial q_\kappa(\alpha)/\partial R_{su}$, $\partial^2 q_\kappa(\alpha)/\partial \hat{X}_s \partial \hat{X}_u$, $\partial^2 q_\kappa(\alpha)/\partial R_{su} \partial R_{kl}$ и $\partial^2 q_\kappa(\alpha)/\partial \hat{X}_s \partial R_{kl}$ после замены моментов их выражениями через c_ν тоже будут линейными комбинациями величин c_ν с коэффициентами, зависящими от \hat{X}_t и R_t .

4.2 Гауссовские системы. Ортогональный субоптимальный фильтр на основе метода квазимоментов

В частном случае разложений (35) по полиномам Эрмита КОР c_κ представляют собой квазимоменты. В этом случае, как показано в [4, 5], для производных полиномов Эрмита G_κ формулы (43) приводятся к виду:

$$\left. \begin{aligned} F_\kappa &= \sum_{s=1}^{n_x} \kappa_s M_{\Delta^x}^{p*} [\varphi_s(Y_t, X, \Theta, t) G_{\kappa-e_s}(X-m)] + \\ &+ \frac{1}{2} \sum_{s=1}^{n_x} \kappa_s (\kappa_s - 1) M_{\Delta^x}^{p*} [\sigma_{ss}(Y_t, X, \Theta, t) G_{\kappa-2e_s}(X-m)] + \\ &+ \sum_{u=2}^{n_x} \sum_{s=1}^{u-1} \kappa_s \kappa_u M_{\Delta^x}^{p*} [\sigma_{su}(Y_t, X, \Theta, t) G_{\kappa-e_s-e_u}(X-m)] ; \\ H_\kappa &= \left\{ M_{\Delta^x}^{p*} [\varphi_1(Y_t, X, \Theta, t)^T G_\kappa(X-m)] + \right. \\ &+ \sum_{s=1}^{n_x} \kappa_s M_{\Delta^x}^{p*} \left[(\psi \nu_0 \psi_1^T) (Y_t, X, t) G_{\kappa-e_s}(X-m) \right] - \\ &\left. - f^{(1)T} c_\kappa \right\} (\psi_1 \nu_0 \psi_1^T)^{-1}(Y_t, t), \end{aligned} \right\} (44)$$

где

$$\begin{aligned} \frac{\partial q_\kappa(\alpha)}{\partial \hat{X}_s} &= -\kappa_s c_{\kappa-e_s}; \\ \frac{\partial q_\kappa(\alpha)}{\partial R_{ss}} &= -\frac{1}{2} \frac{\partial q_\kappa(\alpha)}{\partial \hat{X}_s^2} = -\frac{1}{2} \kappa_s (\kappa_s - 1) c_{\kappa-2e_s}; \\ \frac{\partial q_\kappa(\alpha)}{\partial R_{su}} &= -\frac{\partial^2 q_\kappa(\alpha)}{\partial \hat{X}_s \partial \hat{X}_u} = -\kappa_s \kappa_u c_{\kappa-e_s-e_u}; \\ \frac{\partial^2 q_\kappa(\alpha)}{\partial R_{ss}^2} &= \frac{1}{4} \kappa_s (\kappa_s - 1) (\kappa_s - 2) (\kappa_s - 3) c_{\kappa-4e_s}; \end{aligned}$$

$$\begin{aligned}
 \frac{\partial^2 q_\kappa(\alpha)}{\partial R_{ss} \partial R_{kk}} &= \frac{1}{4} \kappa_s (\kappa_s - 1) \kappa_s (\kappa_s - 1) c_{\kappa-2e_s-2e_k}; \\
 \frac{\partial^2 q_\kappa(\alpha)}{\partial R_{ss} \partial R_{sl}} &= \frac{1}{2} \kappa_s (\kappa_s - 1) (\kappa_s - 2) \kappa_l c_{\kappa-3e_s-e_l}; \\
 \frac{\partial^2 q_\kappa(\alpha)}{\partial R_{ss} \partial R_{kl}} &= \frac{1}{2} \kappa_s (\kappa_s - 1) \kappa_k \kappa_l c_{\kappa-2e_s-e_k-e_l}, \\
 \frac{\partial^2 q_\kappa(\alpha)}{\partial R_{su} \partial R_{sl}} &= \kappa_s (\kappa_s - 1) \kappa_u \kappa_l c_{\kappa-2e_s-e_u-e_l}, \\
 \frac{\partial^2 q_\kappa(\alpha)}{\partial R_{su} \partial R_{kl}} &= \kappa_s \kappa_u \kappa_k \kappa_l c_{\kappa-e_s-e_u-e_k-e_l}; \\
 \frac{\partial^2 q_\kappa(\alpha)}{\partial \hat{X}_s \partial R_{ss}} &= \frac{1}{2} \kappa_s (\kappa_s - 1) (\kappa_s - 2) c_{\kappa-3e_s}; \\
 \frac{\partial^2 q_\kappa(\alpha)}{\partial \hat{X}_s \partial R_{sl}} &= \kappa_s (\kappa_s - 1) \kappa_l c_{\kappa-2e_s-e_l}, \\
 \frac{\partial^2 q_\kappa(\alpha)}{\partial \hat{X}_s \partial R_{kk}} &= \frac{1}{2} \kappa_s \kappa_k (\kappa_k - 1) c_{\kappa-e_s-2e_k}; \\
 \frac{\partial^2 q_\kappa(\alpha)}{\partial \hat{X}_s \partial R_{kl}} &= \kappa_s \kappa_k \kappa_l c_{\kappa-e_s-e_k-e_l}.
 \end{aligned}$$

Таким образом, имеем следующие утверждения разд. 4.

Теорема 4.1. Пусть МСтС (3), (4) — гауссовская ($\psi'' = 0$), выполнены условия существования и единственности решения, а матрица $\sigma_1 = \psi_1 \nu_0 \psi_1^T$ не вырождена. Тогда в основе алгоритма ОСОФ по МОР лежат уравнения (35), (41), и (42) при условии ограниченности функций (43).

Теорема 4.2. В условиях теоремы 4.1 алгоритм ОСОФ по МКМ определяется уравнениями (35), (41) и (42) при условии ограниченности функций (44).

4.3 Негауссовские системы. Ортогональный субоптимальный фильтр по методам ортогональных разложений и квазиоментов

Пользуясь формулами (7) и (8), устанавливаем, что наличие пуассоновского шума в (40) влияет только на функцию $f^{(2)}$. В результате заменим $f^{(2)}$ в (38) на $\bar{f}^{(2)}$ согласно (9) и придем к следующим утверждениям.

Теорема 4.3. Пусть для МСтС (3), (4) выполнены условия существования и единственности решения, а матрица $\sigma_1 = \psi_1 \nu_0 \psi_1^T$ не вырождена. Тогда алгоритм ОСОФ согласно МОР задается уравнениями (35), (41), и (42) при

условии ограниченности функций f , $f^{(1)}$, h , ρ_r , $\bar{f}^{(2)}$, F_κ , H_κ , определяемых (38)–(40), (9), (43).

Теорема 4.4. Пусть для МСтС (3), (4) выполнены условия существования и единственности решения, а матрица $\sigma_1 = \psi_1 \nu_0 \psi_1^T$ не вырождена. Тогда алгоритм ОСОФ согласно МКМ задается уравнениями (35), (41) и (42) при условии ограниченности функций f , $f^{(1)}$, $\bar{f}^{(2)}$, h , ρ_r , F_κ , H_κ , определяемых (38)–(40), (9) и (44).

5 Субоптимальные фильтры на основе модифицированного метода ортогональных разложений

Для ненормированных апостериорных одномерных распределений, обобщая [8] и используя (7), получаем следующие уравнения ОСОФ на базе ММОР:

$$\begin{aligned} \tilde{p}_t(x, \Theta) &\approx \tilde{p}_t^*(x; \Theta, \vartheta) = w(x; \Theta) \left[\mu_t + \sum_{k=1}^N \sum_{|\nu|=k} c_\nu p_\nu(x) \right]; \\ \tilde{c}_\nu &= M_{\Delta_x}^{p^*} [q_\nu(X)] = \left[q_\nu \left(\frac{\partial}{\partial i\lambda} \right) \tilde{g}_t(\lambda, \Theta) \right]_{\lambda=0}, \end{aligned} \quad (45)$$

где $q_\nu(\partial/\partial i\lambda)$ — результат замены в полиноме $q_\nu(x)$ всех одночленов $x_1^{r_1}, \dots, x_{n_x}^{r_{n_x}}$ соответствующими дифференциальными операторами $\partial^r/\partial(i\lambda)^{r_1}, \dots, \partial(i\lambda)^{r_{n_x}}$:

$$\begin{aligned} dc_\nu &= \left[\mu_t f_{\nu 0}(Y_t, \Theta, t) + \sum_{k=1}^N \sum_{|\nu|=k} \tilde{c}_\kappa f_{\nu \kappa}(Y_t, \Theta, t) \right] dt + \\ &+ \left[\mu_t h_{\nu 0}(Y_t, \Theta, t) + \sum_{k=1}^N \sum_{|\kappa|=k} \tilde{c}_\kappa h_{\nu \kappa}(Y_t, \Theta, t) \right] dY_t = \tilde{A}^{c\nu} dt + \tilde{B}^{c\nu} dY_t \\ &(\nu_1, \dots, \nu_{n_x} = 0, 1, \dots, N, |\nu| = 1, \dots, N), \end{aligned} \quad (46)$$

где

$$\left. \begin{aligned} f_{\nu 0}(Y_t, \Theta, t) &= M_{\Delta_x}^w \left\{ \varphi(Y_t, X, \Theta, t)^T q_{\nu x}(X) + \right. \\ &\quad \left. + \frac{1}{2} \text{tr} [q_{\nu xx}(X) \bar{\sigma}_0(Y_t, X, \Theta, t)] \right\}; \\ f_{\nu \kappa}(Y_t, \Theta, t) &= M_{\Delta_x}^{wp^*} \left\{ \varphi(Y_t, X, \Theta, t)^T q_{\nu x}(X) + \right. \\ &\quad \left. + \frac{1}{2} \text{tr} [q_{\nu xx}(X) \bar{\sigma}_0(Y_t, X, \Theta, t)] \right\}; \end{aligned} \right\} \quad (47)$$

$$\left. \begin{aligned} h_{\nu 0}(Y_t, \Theta, t) &= M_{\Delta x}^w \left\{ \varphi_1(Y_t, X, \Theta, t)^T q_\nu(X) + \right. \\ &\quad \left. + q_{\nu x}(X)^T \sigma_1(Y_t, X, \Theta, t) \right\} \sigma_2^{-1}(Y_t, \Theta, t); \\ h_{\nu \kappa}(Y_t, \Theta, t) &= M_{\Delta x}^{wp*} \left\{ \varphi_1(Y_t, X, \Theta, t)^T q_\nu(X) + \right. \\ &\quad \left. + q_{\nu x}(X)^T \sigma_1(Y_t, X, \Theta, t) \right\} \sigma_2(Y_t, \Theta, t)^{-1}, \end{aligned} \right\} (48)$$

$q_{\nu \kappa}(X)$ — матрица-столбец производных полинома $q_\nu(X)$ по компонентам вектора X ; $q_{\nu xx}(X)$ — матрица вторых производных $q_\nu(X)$ по компонентами x ;

$$d\mu_t = \left[\mu_t h_0(Y_t, \Theta, t) + \sum_{k=1}^N \sum_{|\nu|=k} \tilde{c}_\nu h_\nu(Y_t, \Theta, t) \right] dY_t = B^{\mu_t} dY_t, \quad (49)$$

где

$$\begin{aligned} h_0(Y_t, \Theta, t) &= M_{\Delta x}^w \left[\varphi(Y_t, X, \Theta, t)^T \right] \sigma_2(Y_t, \Theta, t)^{-1}; \\ h_\nu(Y_t, \Theta, t) &= M_{\Delta x}^w \left[\varphi_1(Y_t, X, \Theta, t)^T \right] \sigma_2(Y_t, \Theta, t)^{-1}. \end{aligned}$$

Начальные условия для (46) и (49) имеют вид:

$$\tilde{c}_\nu(t_0) = \tilde{c}_\nu^0, \quad \nu_1, \dots, \nu_n = 0, 1, \dots, N, \quad |\nu| = 1, \dots, N; \quad \mu(t_0) = 1, \quad (50)$$

где \tilde{c}_ν^0 — КОР (45) плотности $p_{t_0} = p_0(x | Y_0)$ начального значения X_0 процесса X_t относительно начального значения Y_0 процесса Y_t :

$$p_0(x, \Theta | Y_0) = w(x, \Theta) \left[1 + \sum_{k=1}^N \sum_{|\nu|=k} \tilde{c}_\nu^0 p_\nu(x) \right]. \quad (51)$$

Замечание 5.1. В частности, взяв за w нормальную плотность с параметрами \bar{x}_t и K_t , определим производные полинома $q_\nu(x) = G_\nu(x - \bar{x}_t)$ в (47) и (48) по известным формулам для производных функций [6, приложение I]:

$$\begin{aligned} \frac{\partial q_\nu(x)}{\partial x_l} &= \nu_l G_{\nu-e_l}(x - \bar{x}_t); \\ \frac{\partial^2 q_\nu(x)}{\partial x_l^2} &= \nu_l(\nu_l - 1) G_{\nu-2e_l}(x - \bar{x}_t); \\ \frac{\partial^2 q_\nu(x)}{\partial x_l \partial x_h} &= \nu_l \nu_h G_{\nu-e_l-e_h}(x - \bar{x}_t). \end{aligned}$$

После интегрирования уравнений (46) и (49) оптимальная оценка \hat{X}_t вектора X_t и апостериорная ковариационная матрица R_t ошибки фильтрации определяются приближенно по формулам:

$$\hat{X}_t = \bar{x}_t + \sum_{|\nu|=1} \frac{\tilde{c}_\nu}{\mu_t} M_{\Delta^x}^{wp_\nu} [X_t]; \quad (52)$$

$$R_t = M_{\Delta^x}^w \left[(X_t - \hat{X}_t) (X_t^T - \hat{X}_t^T) \right] + \sum_{k=1}^2 \sum_{|\nu|=k} \frac{\tilde{c}_\nu}{\mu_t} M_{\Delta^x}^{wp_\nu} \left[(X_t - \hat{X}_t) (X_t^T - \hat{X}_t^T) \right] \quad (53)$$

(каждый полином $p_\nu(x)$ ортогонален с весом w по всем одночленам $x_1^{r_1}, \dots, x_{n_x}^{r_{n_x}}$ при $|r| = r_1 + \dots + r_{n_x} < |\nu|$). Все интегралы в (52) и (53) выражаются через моменты распределения w до четвертого порядка включительно. В частности, при нормальной плотности w с параметрами \bar{x} и K все интегралы (52) и (53) выражаются через \bar{x}_t и K_t .

Замечание 5.2. Вектор \bar{x}_t и матрицу K_t в выражении плотности w , в частности нормальной, можно, в принципе, выбрать произвольно, например положить $\bar{x}_t = 0$ и $K_t = I_{n_x}$. Однако для повышения точности аппроксимации (45) при данном N или для уменьшения N при данной точности аппроксимации (45) следует выбирать \bar{x}_t и K_t как можно ближе к \hat{X}_t и R_t . Для достижения этой цели можно рекомендовать выбирать \bar{x}_t и K_t , равными в каждый момент времени t соответственно априорному условному математическому ожиданию $m_{x|y}$ и априорной условной ковариационной матрице $K_{x|y}$ вектора X_t относительно Y_t , полученным, например, МНА.

Замечание 5.3. Уравнения (46) и (49) линейны относительно c_ν и μ (точнее, являются линейными с параметрическими шумами). Однако число уравнений в (46) и (49) при данном N одинаково, а именно: $C_{N+n_x}^N$.

Итак, полученные результаты можно представить в виде следующих утверждений.

Теорема 5.1. Пусть выполнены условия теоремы 2.3. Тогда МОСОФ на базе ММОР определяются уравнениями (45), (46), (49)–(53) и (25).

Теорема 5.2. Пусть выполнены условия теоремы 2.4. Тогда уравнения МОСОФ определяются уравнениями МОСОФ теоремы 5.1 при $\sigma_0 = \psi\nu_0\psi^T$.

Аналогично формулируются соответствующие теоремы МОСОФ на основе ММКМ (см. замечание 5.1).

6 Точность и чувствительность ортогонального субоптимального фильтра

Точность СОФ на базе МНА (МСЛ) оценивается на основе ОСОФ по МОР или МКМ путем удержания конечного числа членов в разложении (35).

Применяя методы теории чувствительности [11, 12] для приближенного анализа фильтрационных уравнений разд. 3–5 и учитывая случайность параметров Θ , придем к следующим уравнениям для функций чувствительности первого порядка:

$$\begin{aligned} d\nabla^\Theta \hat{X}_s &= \nabla^\Theta A^{\hat{X}_s} dt + \nabla^\Theta B^{\hat{X}_s} dY_t, & \nabla^\Theta B^{\hat{X}_s}(t_0) &= 0; \\ d\nabla^\Theta R_{sq} &= \nabla^\Theta A^{R_{sq}} dt + \nabla^\Theta B^{R_{sq}} dY_t, & \nabla^\Theta R_{sq}(t_0) &= 0, \\ d\nabla^\Theta c_\kappa &= \nabla^\Theta A^{c_\kappa} dt + \nabla^\Theta B^{c_\kappa} dY_t, & \nabla^\Theta c_\kappa(t_0) &= 0. \end{aligned}$$

Здесь процедура взятия производных осуществляется по всем входящим переменным, а коэффициенты чувствительности вычисляются при $\Theta = m^\Theta$. При этом предполагается малость дисперсий по сравнению с их математическими ожиданиями. Очевидно, что при дифференцировании по Θ ($\nabla^\Theta = \partial/\partial\Theta$) порядок уравнений возрастает пропорционально числу производных. Аналогично составляются уравнения для элементов матриц вторых функций чувствительности.

Следуя [1, 2], для оценки качества ОСОФ, определяемых теоремами разд. 3–5 при гауссовских Θ с математическим ожиданием m^Θ и ковариационной матрицей K^Θ , введем условную функцию потерь, допускающую квадратичную аппроксимацию:

$$\rho^{\hat{X}_s} = \rho^{\hat{X}_s}(\Theta) = \rho(m^\Theta) + \sum_{i=1}^{n^\Theta} \rho'_i(m^\Theta) \Theta_s^0 + \sum_{i,j=1}^{n^\Theta} \rho''_{ij}(m^\Theta) \Theta_i^0 \Theta_j^0,$$

а также показатель ε

$$\varepsilon = \varepsilon_2^{1/4}.$$

Здесь

$$\begin{aligned} \varepsilon_2 &= M^N [\rho(\Theta)^2] - \rho(m^\Theta)^2; \\ M^N [\rho(\Theta)^2] &= \rho(m^\Theta)^2 + \rho'(m^\Theta)^T K^\Theta \rho'(m^\Theta) + \\ &+ 2\rho(m^\Theta) \text{tr} [\rho''(m^\Theta) K^\Theta] + \\ &+ \left\{ \text{tr} [\rho''(m^\Theta) K^\Theta] \right\}^2 + 2\text{tr} [\rho''(m^\Theta) K^\Theta]^2, \end{aligned}$$

а функции ρ' и ρ'' по известным формулам [11, 12] определяются на основе первых и вторых функций чувствительности.

Изложенные выше методы синтеза ОСОФ и МОСОФ дают принципиальную возможность получить фильтр, близкий к оптимальному по оценке с любой степенью точности. Чем выше максимальный порядок учитываемых моментов, КОР и квазимоментов, тем выше будет точность приближения к оптимальной оценке. Однако число уравнений, определяющих параметры апостериорного одномерного распределения, быстро растет с увеличением числа учитываемых параметров. Соответствующие оценки можно найти в [5–7].

7 Тестовые примеры

Рассмотрим нелинейную с мультипликативными шумами гауссовскую СтС [3, 6]:

$$\begin{aligned} \dot{X}_t &= -X_t^3 + X_t V_1(\Theta); & Z_t(\Theta) &= \dot{Y}_t = X_t + V_2(\Theta); \\ \hat{X}_{t_0} &= MX(t_0); & R_{t_0} &= DX(t_0). \end{aligned}$$

Здесь предполагается, что интенсивности гауссовских белых шумов ν_1 и ν_2 зависят от одного скалярного параметра Θ , т. е. $\nu_{1,2} = \nu_{1,2}(\Theta)$.

Обозначим $\nabla = \nabla^\Theta = (\partial/\partial\Theta)$. Тогда уравнения точности и чувствительности СОФ на основе МНА примут вид:

$$\left. \begin{aligned} \dot{\hat{X}}_t &= -\hat{X}_t \left(\hat{X}_t^2 + 3R_t \right) + \nu_2^{-1} R_t \left(Z_t - \hat{X}_t \right); \\ \dot{R}_t &= (\nu_1 - 6R_t) \left(\hat{X}_t^2 + R_t \right) - \nu_2^{-1} R_t^2; \end{aligned} \right\} \quad (54)$$

$$\left. \begin{aligned} \nabla \dot{\hat{X}}_t &= - \left[3\hat{X}_t + (3 + \nu_2^{-1}) \right] \nabla \hat{X}_t + \nu_2^{-1} R_t \nabla Z_t - \\ &\quad - \nu_2^{-2} R_t \left(Z_t - \hat{X}_t \right) \nabla \nu_2, \quad \nabla \hat{X}_{t_0} = 0, \\ \nabla \dot{R}_t &= 2(\nu_1 - 6R_t) \hat{X}_t \nabla \hat{X}_t + \left[(\nu_1 - 6R_t) - 6 \left(\hat{X}_t^2 + R_t \right) - \right. \\ &\quad \left. - 2\nu_2^{-1} R_t \right] \nabla R_t + \left(\hat{X}_t^2 + R_t \right) \nabla \nu_1 + \nu_2^{-2} R_t \nabla \nu_2, \quad \nabla R_{t_0} = 0. \end{aligned} \right\} \quad (55)$$

Уравнения (54) нелинейны относительно \hat{X}_t и R_t , причем фильтр существует только при наличии шума в наблюдениях $\nu_2 \neq 0$ и произвольном шуме интенсивности ν_1 в уравнении состояния. Уравнения (55) для первых функций чувствительности являются линейными неоднородными уравнениями вследствие зависимости $\nu_{1,2} = \nu_{1,2}(\Theta)$ и $Z_t = Z_t(\Theta)$.

Теперь рассмотрим ОСОФ. В соответствии с теоремой 3.1 (в отличие от [3], с точностью до вероятностных моментов четвертого порядка) имеем следующие уравнения ОСОФ:

$$\left. \begin{aligned}
 \dot{\hat{X}}_t &= -\hat{X}_t \left(\hat{X}_t^2 + 3R_t \right) + \nu_2^{-1} R_t \left(Z_t - \hat{X}_t \right) - c_3; \\
 \dot{R}_t &= \left[\nu_1 - 6 \left(\hat{X}_t^2 + R_t \right) \right] R_t - \nu_2^{-1} R_t^2 - 6\hat{X}_t c_3 + \\
 &\quad + \nu_2^{-1} c_3 \left(Z_t - \hat{X}_t \right) - 2c_4; \\
 \dot{c}_3 &= -18\hat{X}_t R_t^2 - 9 \left(\hat{X}_t^2 + 3R_t \right) c_3 - 9\hat{X}_t c_4 + 3\nu_1 \left(2\hat{X}_t R_t + c_3 \right) - \\
 &\quad - \frac{3}{2} \nu_2^{-1} c_3 + \nu_2^{-1} c_4 \left(Z_t - \hat{X}_t \right); \\
 \dot{c}_4 &= 12R_t^3 - 72\hat{X}_t R_t c_3 - 12 \left(\hat{X}_t^2 + 4R_t \right) c_4 + 4c_3^2 + \\
 &\quad + 6\nu_1 \left[\left(\hat{X}_t^2 + 2R_t \right) R_t + 2\hat{X}_t c_3 + c_4 \right] + \nu_2^{-1} \left(6R_t^3 - 3c_3^3 - 2R_t c_4 \right).
 \end{aligned} \right\} (56)$$

Уравнениям (56) отвечают следующие уравнения для первых функций чувствительности:

$$\begin{aligned}
 \nabla \dot{\hat{X}}_t &= - \left(3\hat{X}_t^2 + 3R_t + \nu_2^{-1} R_t \right) \nabla \hat{X}_t + \left[\nu_2^{-1} \left(Z_t - \hat{X}_t \right) - 3\hat{X}_t \right] \nabla R_t + \\
 &\quad + \nu_2^{-1} R_t \nabla Z_t - \nu_2^{-2} R_t \left(Z_t - \hat{X}_t \right) \nabla \nu_2 - \nabla c_3, \quad \nabla \hat{X}(t_0) = 0; \\
 \nabla \dot{R}_t &= - \left(12\hat{X}_t R_t + 6c_3 + \nu_2^{-1} c_3 \right) \nabla \hat{X}_t - \left(\nu_1 + \nu_2^{-1} R_t + 6\hat{X}_t^2 + 12R_t \right) \nabla R_t + \\
 &\quad + \nu_2^{-1} c_3 \nabla Z_t + \left[\nu_2^{-1} \left(Z_t - \hat{X}_t \right) - 6\hat{X}_t \right] \nabla c_3 - \\
 &\quad - 2\nabla c_4 + R_t \nabla \nu_1 + \nu_2^{-2} \left[R_t^2 - c_3 \left(Z_t - X_t \right) \right] \nabla \nu_2, \quad \nabla R(t_0) = 0; \\
 \nabla \dot{c}_3 &= -9 \left(c_3 + 2\hat{X}_t c_3 + 2R_t^2 + \nu_2^{-1} c_4 \right) \nabla \hat{X}_t + 3 \left(2\nu_1 \hat{X}_t - 9c_3 - 2\hat{X}_t R_t \right) \nabla R_t + \\
 &\quad + 3 \left(\nu_1 + \frac{9}{2} \nu_2^{-1} - 3\hat{X}_t - 9R_t - 3\hat{X}_t^2 \right) \nabla c_3 + \nu_2^{-1} c_4 \nabla Z_t + \\
 &\quad + \left[-9\hat{X}_t + \nu_2^{-1} \left(Z_t - X_t \right) \right] \nabla c_4 + 3 \left(2\hat{X}_t R_t + c_3 \right) \nabla \nu_1 + \\
 &\quad + \left[\frac{3}{2} - c_4 \left(Z_t - \hat{X}_t \right) \right] \nu_2^{-2} \nabla \nu_2, \quad \nabla c_3(t_0) = 0; \\
 \nabla \dot{c}_4 &= 12 \left(\nu_1 c_3 - 2\hat{X}_t c_4 - R_t c_3 \right) \nabla \hat{X}_t + \left[2\nu_2^{-1} \left(9R_t^2 - c_4 \right) + \right. \\
 &\quad \left. + 6\nu_1 \left(\hat{X}_t^2 + 3R_t \right) + 36R_t^2 - 72\hat{X}_t c_3 - 48c_4 \right] \nabla R_t + \\
 &\quad + 2 \left(4c_3 + 6\nu_1 \hat{X}_t - 3\nu_2^{-1} c_3 - 18\hat{X}_t R_t \right) \nabla c_3 +
 \end{aligned}$$

$$+ 2 \left[3\nu_1 - 6 \left(\hat{X}_t^2 + 4R_t \right) + \nu_2^{-1} R_t \right] \nabla c_4 + 6 \left[\left(\hat{X}_t^2 + 2R_t \right) R_t + \right. \\ \left. + 2\hat{X}_t c_3 + c_4 \right] \nabla \nu_1 + \nu_2^{-2} \left(3c_3^2 + 2R_t c_4 - 6R_t^3 \right) \nabla \nu_2, \quad \nabla c_4(t_0) = 0.$$

Уравнения точности этого ОСОФ нелинейны относительно \hat{X}_t , R_t и c_3 , c_4 и справедливы только при $\nu_2 \neq 0$ и произвольном ν_1 . Уравнения для первых функций чувствительности являются линейными неоднородными уравнениями вследствие зависимости $Z_t = Z_t(\Theta)$, $\nu_{1,2} = \nu_{1,2}(\Theta)$.

8 Заключение

Для нелинейных дифференциальных систем с винеровскими и пуассоновскими шумами в уравнениях состояния и винеровскими шумами в уравнениях наблюдения (в том числе и на гладких многообразиях) разработаны методы синтеза по среднеквадратическому критерию СОФ на основе аппроксимации апостериорного одномерного распределения МНА и МСЛ, а также МОР и МКМ. Особое внимание уделено модифицированным нормальным и ортогональным СОФ на основе ненормированных распределений. Получены уравнения точности и чувствительности нормальных и ортогональных СОФ. Приведены тестовые примеры для одномерной нелинейной системы с параметрическим шумом. Алгоритмическое обеспечение положено в основу инструментального программного обеспечения в библиотеке «StS-Filter».

В качестве дальнейших обобщений можно рассмотреть дискретные и непрерывно-дискретные МСтС. Развития требуют также методы экстраполяции и интерполяции в таких системах.

Литература

1. Сеницын И. Н. Аналитическое моделирование распределений на основе ортогональных разложений в нелинейных стохастических системах на многообразиях // Информатика и её применения, 2015. Т. 9. Вып. 3. С. 17–24.
2. Сеницын И. Н. Применение ортогональных разложений для аналитического моделирования многомерных распределений в нелинейных стохастических системах на многообразиях // Системы и средства информатики, 2015. Т. 25. № 3. С. 3–22.
3. Сеницын И. Н. Ортогональные субоптимальные фильтры для нелинейных стохастических систем на многообразиях // Информатика и её применения, 2016. Т. 10. Вып. 1. С. 34–44.
4. Ватанабэ С., Икэда Н. Стохастические дифференциальные уравнения и диффузионные процессы / Пер. с англ. — М.: Наука, 1986. 448 с. (*Watanabe S., Ikeda N. Stochastic differential equations and diffusion processes. — Amsterdam – Oxford – New York: North-Holland Publishing Co.; Tokyo: Kodansha Ltd., 1981. 476 p.*)
5. Справочник по теории вероятностей и математической статистике / Под ред. В. С. Королюка, Н. И. Портенко, А. В. Скорохода, А. Ф. Турбина. — М.: Наука, 1985. 640 с.
6. Пугачёв В. С., Сеницын И. Н. Стохастические дифференциальные системы. Анализ и фильтрация. — М.: Наука, 1990. 632 с.

7. Пугачёв В. С., Ситицын И. Н. Теория стохастических систем. — М.: Логос, 2000; 2004. 1000 с.
8. Ситицын И. Н. Фильтры Калмана и Пугачева. — 2-е изд. — М.: Логос, 2007. 776 с.
9. Wonham M. Some application of stochastic differential equations to optimal nonlinear filtering // J. Soc. Industr. Appl. Math. Control, 1965. Vol. 2. P. 347–369.
10. Zakai M. On the optimal filtering of diffusion processes // Ztschr. Wahrscheinlichkeitstheor. Verm. Geb., 1969. Bd. 11. S. 230–243.
11. Евланов А. Г., Константинов В. М. Системы со случайными параметрами. — М.: Наука, 1987. 568 с.
12. Справочник по теории автоматического управления / Под ред. А. А. Красовского. — М.: Наука, 1987. 712 с.

Поступила в редакцию 05.11.15

NORMAL AND ORTHOGONAL SUBOPTIMAL FILTERS FOR NONLINEAR STOCHASTIC SYSTEMS ON MANIFOLDS

I. N. Sinitsyn

Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilova Str., Moscow 119333, Russian Federation

Abstract: For nonlinear differential stochastic systems on manifolds (MStS) with Wiener and Poisson noises, the theory of synthesis of suboptimal filters based on the normal approximation method, the statistical linearization method, the orthogonal expansions method, and quasi-moment method is developed. Exact optimal (for mean square error criteria) equations for MStS with Gaussian noises in observation equations for one-dimensional *a posteriori* characteristic function are derived. Special attention is paid to modified filters based on unnormed distributions. Problems of approximate solving of exact equations are discussed. Accuracy and sensitivity equations are presented. A test example for nonlinear scalar differential equation with additive and multiplicative noises is given. Some generalizations are mentioned.

Keywords: *a posteriori* one-dimensional distribution; coefficient of orthogonal expansion; first sensitivity function; normal approximation method; normal suboptimal filter; modified NAM; modified OEM; orthogonal expansion method; orthogonal suboptimal filter; quasi-moment method; quasi-moment; statistical linearization method; stochastic system on manifolds; suboptimal filter; Wiener white noise

DOI: 10.14357/08696527160113

Acknowledgments

The research was supported by the Russian Foundation for Basic Research (project 15-07-002244).

References

1. Sinitsyn, I. N. 2015. Analiticheskoe modelirovanie raspredeleniy na osnove ortogonal'nykh razlozheniy v nelineynykh stokhasticheskikh sistemakh na mnogoobraznykh [Analytical modeling in stochastic systems on manifolds based on orthogonal expansions]. *Informatika i ee Primeneniya — Inform. Appl.* 9(2):17–24.
2. Sinitsyn, I. N. 2015. Primenenie ortogonal'nykh razlozheniy dlya analiticheskogo modelirovaniya mnogomernykh raspredeleniy v nelineynykh stokhasticheskikh sistemakh na mnogoobraznykh [Applications of orthogonal expansions for analytical modeling of multidimensional distributions in stochastic systems on manifolds]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 25(3):3–22.
3. Sinitsyn, I. N. 2016. Ortogonal'nye suboptimal'nye fil'try dlya nelineynykh stokhasticheskikh sistem na mnogoobraznykh [Orthogonal suboptimal filters for nonlinear stochastic systems on manifolds]. *Informatika i ee Primeneniya — Inform. Appl.* 10(1):34–44.
4. Watanabe, S., and N. Ikeda. 1981. *Stochastic differential equations and diffusion processes*. Amsterdam–Oxford–New York: North-Holland Publishing Co.; Tokyo: Kodansha Ltd. 476 p.
5. Korolyuk, V. S., N. I. Portenko, A. V. Skorokhod, and A. F. Turbin, eds. 1985. *Spravochnik po teorii veroyatnosti i matematicheskoy statistike* [Handbook: Probability theory and mathematical statistics]. Moscow: Nauka. 640 p.
6. Pugachev, V. S., and I. N. Sinitsyn. 1987. *Stochastic differential systems. Analysis and filtering*. — Chichester–New York, NY: John Wiley. 549 p.
7. Pugachev, V. S., and I. N. Sinitsyn. 2000, 2004. *Teoriya stokhasticheskikh sistem* [Stochastic systems. Theory and applications]. Moscow: Logos. 1000 p.
8. Sinitsyn, I. N. 2007. *Fil'try Kalmana i Pugacheva* [Kalman and Pugachev filters]. 2nd ed. Moscow: Logos. 776 p.
9. Wonham, M. 1965. Some application of stochastic differential equations to optimal nonlinear filtering. *J. Soc. Industr. Appl. Math. Control* 2:347–369.
10. Zakai, M. 1969. On the optimal filtering of diffusion processes. *Ztschr. Wahrscheinlichkeitstheor. Verm. Geb.* 11:230–243.
11. Evlanov, A. G., and V. M. Konstantinov. 1976. *Sistemy so slozhnymi parametrami* [Systems with random parameters]. Moscow: Nauka. 568 p.
12. Krasovskii, A. A., ed. 1987. *Spravochnik po teorii avtomaticheskogo upravleniya* [Handbook for automatic control]. Moscow: Nauka. 712 p.

Received November 5, 2015

Contributor

Sinitsyn Igor N. (b. 1940) — Doctor of Science in technology, professor, Honored scientist of RF, Head of Department, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; sinitsin@dol.ru

МАТЕМАТИЧЕСКОЕ ОБЕСПЕЧЕНИЕ АНАЛИТИЧЕСКОГО МОДЕЛИРОВАНИЯ НОРМАЛЬНЫХ ПРОЦЕССОВ В СТОХАСТИЧЕСКИХ СИСТЕМАХ СО СЛОЖНЫМИ ДРОБНО-РАЦИОНАЛЬНЫМИ НЕЛИНЕЙНОСТЯМИ¹

*И. Н. Синицын¹, В. И. Синицын², И. В. Сергеев³, Э. Р. Корепанов⁴,
В. В. Белоусов⁵, В. С. Шоргин⁶*

Аннотация: Рассмотрено математическое обеспечение аналитического моделирования нормальных (гауссовских) стохастических процессов (НСтП) в дифференциальных и разностных (дискретных) стохастических системах (СтС) со сложными дробно-рациональными нелинейностями (СДРН) скалярного и векторного аргумента. Даны типовые представления скалярных и векторных СДРН. Получены уравнения точности и чувствительности методов нормальной аппроксимации (МНА) и статистической линеаризации (МСЛ) для аналитического моделирования нестационарных и стационарных НСтП. Рассмотрены аналитические и численные методы расчета типовых интегралов МНА (МСЛ). Обсуждается состав модуля «StS-CFRN. Analysis» алгоритмического и инструментального программного обеспечения. В приложении даны примеры вычисления типовых интегралов и коэффициентов МНА (МСЛ) для СтС с СДРН. Рассмотрены обобщения полученных алгоритмов.

Ключевые слова: аналитическое моделирование; метод нормальной аппроксимации (МНА); метод статистической линеаризации (МСЛ); модуль «StS-CFRN. Analysis»; нормальный стохастический процесс (НСтП); сложные дробно-рациональные нелинейности (СДРН); стохастическая система (СтС)

DOI: 10.14357/08696527160114

*Работа выполнена при поддержке РФФИ (проект 15-07-02244).

¹Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, sinitsin@dol.ru

²Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, vsinitsin@ipiran.ru

³Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, isergeev@ipiran.ru

⁴Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, ekoerpanov@ipiran.ru

⁵Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, vbelousov@ipiran.ru

⁶Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, vshorgin@ipiran.ru

1 Введение

Методы и средства аналитического моделирования НСтП для дифференциальных СтС, в том числе и на многообразиях, описаны в [1–5]. В [6] для СтС с СДРН даны типовые представления дробно-рациональных функций (ДРФ) и соответствующих СДРН, получены уравнения МНА и МСЛ, представлено алгоритмическое обеспечение МНА (МСЛ) для СтС с СДРН и тестовые примеры.

Дадим развитие [6] на случай непрерывных и дискретных СтС. Особое внимание уделим алгоритмическому и программному обеспечению модуля «StS-CFRN. Analysis». В разд. 2 приведены краткие сведения из теории ДРФ и типовые представления СДРН. Рассмотрены примеры сложных нелинейностей, приводимых к СДРН. Алгоритмическое обеспечение на основе МНА (МСЛ) для дифференциальных СтС с винеровскими и пуассоновскими шумами представлено в разд. 3. Особое внимание уделено вопросам точности и чувствительности алгоритмов. В разд. 4 описано соответствующее алгоритмическое обеспечение для дискретных СтС. В приложении даны примеры вычисления типовых интегралов и коэффициентов МНА (МСЛ) для СтС с СДРН, а также некоторые применения.

2 Дробно-рациональные функции и нелинейности

Скалярной ДРФ называется любая функция $z = \varphi(y) = R^{\text{ДРФ}}(y)$, представляемая в виде отношения двух полиномов $P^+(y)$ и $P^-(y)$ [6–9]:

$$R^{\text{ДРФ}}(y) = \frac{P^-(y)}{P^+(y)}. \quad (1)$$

Для правильной ДРФ степень числителя $P^-(y)$ не больше степени знаменателя $P^+(y)$. В общем случае скалярную ДРФ можно однозначно представить в виде полинома $P(y)$ и правильной ДРФ:

$$\frac{P^-(y)}{P^+(y)} = P(y) + \frac{P_1^-(y)}{P^+(y)}. \quad (2)$$

Следуя [6–9], приведем основные свойства скалярных правильных ДРФ.

¹⁰ Правильная ДРФ, знаменатель $P^+(y)$ которой представим в виде произведения попарно взаимно простых множителей

$$P^+(y) = P_1^+(y)P_2^+(y) \cdots,$$

однозначно представима в виде суммы правильных ДРФ

$$\frac{P_1^-(y)}{P^+(y)} = \frac{P_{11}^-(y)}{P_1^+(y)} + \frac{P_{12}^-(y)}{P_2^+(y)} + \cdots.$$

2⁰ Пусть $P^-(y)/P^+(y)$ — правильная ДРФ, а знаменатель $P^+(y)$ имеет вид:

$$P^+(y) = a_0 (y - y_1)^{k_1} (y - y_2)^{k_2} \dots (y - y_m)^{k_m}.$$

Тогда имеет место следующее единственное представление:

$$\frac{P_1^-(y)}{P^+(y)} = \sum_{\nu=1}^m \left[\frac{A_{\nu 0}}{(y - y_\nu)^{k_\nu}} + \frac{A_{\nu 1}}{(y - y_\nu)^{k_{\nu-1}}} + \dots + \frac{A_{\nu, k_\nu-1}}{y - y_\nu} \right].$$

3⁰ Пусть $P_1^-(y)/P^+(y)$ является правильной ДРФ, а разложение знаменателя $P^+(y)$ на неприводимые множители имеет вид:

$$P^+(y) = a_0 (y - y_1)^{k_1} \dots (y - y_r)^{k_r} (y^2 + p_1 y + q_1)^{l_1} \dots (y^2 + p_s y + q_s)^{l_s}.$$

Тогда имеет место следующее единственное представление правильной ДРФ:

$$\frac{P_1^-(y)}{P^+(y)} = \sum_{\nu=1}^h \sum_{\mu=1}^{k_\nu} \frac{A_{\nu\mu}}{(y - y_\nu)^\mu} + \sum_{\nu=1}^s \sum_{\mu=1}^{l_\nu} \frac{B_{\nu\mu} y + C_{\nu\mu}}{(y^2 + p_\nu y + q_\nu)^\mu}.$$

4⁰ Дифференцирование ДРФ (1) проводится по формуле:

$$\frac{dR^{\text{ДРФ}}(y)}{dy} = \frac{(P^-(y))' P^+(y) - (P^+(y))' P^-(y)}{(P^+(y))^2}.$$

5⁰ Неопределенный интеграл от ДРФ всегда выражается через элементарные функции, так что

$$\begin{aligned} \int \frac{P^-(y)}{P^+(y)} dy &= \int P(y) dy + \int \frac{P_1^-(y)}{P^+(y)} dy = \\ &= \int P(y) dy + \sum_{\nu=1}^h \sum_{\mu=1}^{k_\nu} A_{\nu\mu} \int \frac{dy}{(y - y_\nu)^\mu} + \sum_{\nu=1}^s \sum_{\mu=1}^{l_\nu} \int \frac{B_{\nu\mu} y + C_{\nu\mu}}{(y^2 + p_\nu y + q_\nu)^\mu} dy, \end{aligned}$$

при этом интеграл от полинома $P(y)$ всегда выражается в конечном виде. Интегралы первой суммы равны:

$$\int \frac{dy}{(y - y_\nu)^\mu} = \frac{1}{1 - \mu} (y - y_\nu)^{1-\mu} + C \quad (\mu > 1).$$

Далее

$$\int \frac{By + C}{(y^2 + py + q)^\mu} = \frac{B}{2(1 - \mu)(y^2 + py + q)^{\mu-1}} + \frac{2C - Bp}{2} I_\mu(\xi).$$

Здесь $I_\mu(\xi)$ вычисляется по рекуррентной формуле:

$$I_\mu = \frac{\xi(\xi^2 + a^2)^{1-\mu}}{2a^2(\mu - 1)} + \frac{2\mu - 3}{a^2(2\mu - 2)} I_{\mu-1}(\xi),$$

где

$$\xi = y + \frac{p}{2}; \quad a^2 = q - \frac{p^2}{4}.$$

Таким образом, неопределенный интеграл от действительной ДРФ может быть записан в виде суммы, во-первых, некоторой рациональной функции с действительными коэффициентами, во-вторых, выражений вида

$$c_\nu^{(1)} (y - y_\nu)^{-\nu}, c_\nu^{(2)} \ln |y - y_\nu|, c_\nu^{(3)} \ln (y^2 + p_\nu y + q_\nu), c_\nu^{(4)} \operatorname{arctg} \left[\frac{2y + p_\nu}{\sqrt{4q_\nu - p_\nu^2}} \right]$$

и, в-третьих, произвольной постоянной.

6^0 Имеет место следующее правило Остроградского для выделения рациональной части $R_1(y)$ интеграла

$$\int \frac{P^-(y)}{P^+(y)} dy = R_1(y) + R_2(y),$$

где $R_2(y)$ — трансцендентная часть. Если предположить, что знаменатель $P^+(y)$ допускает представление вида:

$$P^+(y) = L(y)K(y),$$

где

$$\begin{aligned} L(y) &= \\ &= a_0 (y - y_1)^{k_1-1} \cdots (y - y_r)^{k_r-1} (y^2 + p_1 y + q_1)^{l_1-1} \cdots (y^2 + p_s y + q_s)^{l_s-1}; \\ K(y) &= (y - y_1) \cdots (y - y_r) (y^2 + p_1 y + q_1) \cdots (y^2 + p_s y + q_s), \end{aligned}$$

то имеет место единственное представление:

$$R_1(y) = \frac{M(y)}{L(y)}; \quad R_2(y) = \int \frac{N(y) dy}{K(y)}.$$

Коэффициенты полиномов $M(y)$ и $N(y)$ находятся методом неопределенных коэффициентов, поскольку $M(y)/L(y)$, $N(y)/K(y)$ и $L'(y)K(y)/L(y)$ являются правильными ДРФ.

Пример 2.1. Примерами СДРН, получаемых посредством отрезков сумм типовых ДРФ, могут служить следующие [6]:

$$\begin{aligned}\varphi^{\text{СДРН}}(Y, t) &= \sum_{r=1}^n l_{rt} \varphi_r^{\text{ДРН}}(Y); \\ \varphi^{\text{СДРН}}(Y, t) &= \sum_{r=1}^n l_{rt} \varphi_r^{\text{ДРН}}(Y) \varphi_r(Y); \\ \varphi^{\text{СДРН}}(Y, t) &= \frac{\sum_{r=1}^{n'} l'_{rt} \varphi_r^{\text{ДРН}}(Y)}{\sum_{r=1}^{n''} l''_{rt} \varphi_r^{\text{ДРН}}(Y)}; \\ \varphi^{\text{СДРН}}(Y, t) &= \frac{\sum_{r=1}^{n'} l'_{rt} \varphi_r^{\text{ДРН}}(Y) \varphi'_r(Y)}{\sum_{r=1}^{n''} l''_{rt} \varphi_r^{\text{ДРН}}(Y) \varphi''_r(Y)},\end{aligned}$$

где $\varphi_r^{\text{ДРН}}(Y)$, $\varphi_r^{\text{ДРН}}(Y)$ и $\varphi_r^{\text{ДРН}}(Y)$ — типовые ДРН; l_{rt} , l'_{rt} и l''_{rt} — коэффициенты, зависящие от времени t ; $\varphi_r(Y)$, $\varphi'_r(Y)$, $\varphi''_r(Y)$ — известные функции.

Другими примерами скалярных СДРН являются нелинейности, получаемые путем соответствующего преобразования аргумента [6]:

$$\varphi^{\text{СДРН}}(Y, t) = \varphi(\psi^{\text{ДРН}}(Y, t), t); \quad \varphi^{\text{СДРН}}(Y, t) = \varphi^{\text{ДРН}}(\psi(Y, t), t).$$

В приложении приведены примеры сложных нелинейностей, приводимых к СДРН.

Скалярная ДРФ векторного аргумента определяется как отношение конечных сумм [8]:

$$R^{\text{ДРФ}}(y_1, \dots, y_n) = \frac{\sum_{k_1, \dots, k_n} a_{k_1, \dots, k_n} (y_1)^{k_1} \dots (y_n)^{k_n}}{\sum_{i_1, \dots, i_n} b_{i_1, \dots, i_n} (y_1)^{i_1} \dots (y_n)^{i_n}}. \quad (3)$$

Пример 2.2. В качестве примеров скалярных СДРН векторного аргумента $Y = [Y_1 \dots Y_p]^T$ рассмотрим следующие [6]:

$$\begin{aligned}\varphi^{\text{СДРН}}(Y, t) &= \sum_{r=1}^n \prod_{h=1}^H l_{rh,t} \varphi_{rh}^{\text{ДРН}}(Y_h); \\ \varphi^{\text{СДРН}}(Y, t) &= \frac{\sum_{r=1}^{n'} \prod_{h=1}^{H'} l'_{rh,t} \varphi_{rh}^{\text{ДРН}}(Y_h)}{\sum_{r=1}^{n''} \prod_{h=1}^{H''} l''_{rh,t} \varphi_{rh}^{\text{ДРН}}(Y_h)}.\end{aligned}$$

В случае векторных и матричных СДРН имеют место соответствующие формулы (3) для компонент.

3 Дифференциальные стохастические системы со сложными дробно-рациональными нелинейностями. Алгоритмическое обеспечение

Следуя [6], уравнения конечномерных непрерывных нелинейных систем со стохастическими возмущениями путем расширения вектора состояния СтС могут быть записаны в виде следующего векторного стохастического дифференциального уравнения Ито:

$$dY_t = a(Y_t, \Theta, t) dt + b(Y_t, \Theta, t) dW_0 + \int_{R_0} c(Y_t, \Theta, t, v) P^0(\Theta, dt, dv),$$

$$Y(t_0) = Y_0. \quad (4)$$

Здесь Y_t — $(p \times 1)$ -мерный вектор состояния, $Y_t \in \Delta^y$ (Δ^y — гладкое многообразие состояний); $a = a(Y_t, \Theta, t)$ и $b = b(Y_t, \Theta, t)$ — известные $(p \times 1)$ -мерная и $(p \times m)$ -мерная функции Y_t и t ; $W_0 = W_0(\Theta, t)$ — $(r \times 1)$ -мерный винеровский СтП интенсивности $\nu_0 = \nu_0(\Theta, t)$; $c(Y_t, \Theta, t, v)$ — $(p \times 1)$ -мерная функция Y_t, t и вспомогательного $(q \times 1)$ -мерного параметра v ; $\int_{\Delta} dP^0(t, A)$ — центрированная пуассоновская мера, определяемая соотношением:

$$\int_{\Delta} dP^0(\Theta, t, A) = \int_{\Delta} dP(\Theta, t, A) = \int_{\Delta} \nu_P(\Theta, t, A) dt.$$

В (4) принято: \int_{Δ} — число скачков пуассоновского СтП в интервале времени $\Delta = (t_1, t_2]$; $\nu_P(\Theta, t, A)$ — интенсивность пуассоновского СтП $P(\Theta, t, A)$; A — некоторое борелевское множество пространства R_0^q с выколотым началом. Начальное значение Y_0 представляет собой случайную величину (с.в.), не зависящую от приращений $W_0(\Theta, t)$ и $P(\Theta, t, A)$ на интервалах времени, следующих за t_0 , $t_0 \leq t_1 \leq t_2$, для любого множества A , Θ — $(p^\Theta \times 1)$ -мерный вектор параметров системы.

В случае аддитивных нормальных (гауссовских) и обобщенных пуассоновских возмущений уравнение (4) имеет вид:

$$\dot{Y}_t = a(Y_t, \Theta, t) + b_0(\Theta, t)V, \quad V = \dot{W}, \quad Y(t_0) = Y_0. \quad (5)$$

Здесь W — СтП с независимыми приращениями, представляющий собой смесь нормального и обобщенного пуассоновского СтП.

Для компонент $\varphi(Y_t, \Theta, t) = \{a_h, b_{kj}, c_h\}$ функций a , b и c , являющихся СДРН, примем представления (1)–(3). Если предположить существование конечных вероятностных моментов второго порядка для моментов времени t_1 и t_2 , то уравнения МНА примут следующий вид [10, 11]:

– для характеристических функций:

$$\left. \begin{aligned} g_1^N(\lambda; t) &= \exp \left[i\lambda^T m_t - \frac{1}{2} \lambda^T K_t \lambda \right]; \\ g_{t_1, t_2}^N(\lambda_1, \lambda_2; t_1, t_2) &= \exp \left[i\bar{\lambda}^T \bar{m}_2 - \frac{1}{2} \bar{\lambda}^T \bar{K}_2 \lambda \right], \end{aligned} \right\} \quad (6)$$

где

$$\bar{\lambda} = \left[\lambda_1^T \lambda_2^T \right]^T; \quad \bar{m}_2 = \left[m_{t_1}^T m_{t_2}^T \right]^T; \quad \bar{K}_2 = \begin{bmatrix} K(t_1, t_1) & K(t_1, t_2) \\ K(t_2, t_1) & K(t_2, t_2) \end{bmatrix};$$

– для математических ожиданий m_t , ковариационной матрицы K_t и матрицы ковариационных функций $K(t_1, t_2)$:

$$\left. \begin{aligned} \dot{m}_t &= A^m(m_t, K_t, \Theta, t), \quad m_0 = m(\Theta, t_0); \\ \dot{K}_t &= A^K(m_t, K_t, \Theta, t), \quad K_0 = K(\Theta, t_0); \\ \frac{\partial K(t_1, t_2)}{\partial t_2} &= A^{K_{t_1, t_2}}(m_t, K_t, \Theta, t) = K(t_1, t_2) a_{21}(m_{t_2}, K_{t_2}, \Theta, t_2)^T, \\ &K(\Theta, t_1, t_1) = K_{t_1}. \end{aligned} \right\} \quad (7)$$

Здесь приняты следующие обозначения:

$$\left. \begin{aligned} m_t &= M_{\Delta y}^N Y_t; \quad Y_t^0 = Y_t - m_t; \\ K_t &= M_{\Delta y}^N [Y_t^0 Y_t^{0T}]; \quad K(t_1, t_2) = M_{\Delta y}^N [Y_{t_1}^0 Y_{t_2}^{0T}]; \\ A^m &= A^m(m_t, K_t, \Theta, t) = M_{\Delta y}^N [a(Y_t, \Theta, t)]; \\ A^K &= A^K(m_t, K_t, \Theta, t) = \\ &= a_{21}(m_t, K_t, \Theta, t) + a_{21}(m_t, K_t, \Theta, t)^T + a_{22}(m_t, K_t, \Theta, t); \\ a_{21} &= a_{21}(m_t, K_t, \Theta, t) = M_{\Delta y}^N [a(Y_t, \Theta, t) Y_t^{0T}]; \\ a_{22} &= a_{22}(m_t, K_t, \Theta, t) = M_{\Delta y}^N [\sigma(Y_t, \Theta, t)]; \\ \sigma(Y_t, \Theta, t) &= b(Y_t, \Theta, t) \nu_0(t) b(Y_t, \Theta, t)^T + \\ &+ \int_{R_0^g} c(Y_t, \Theta, t, v) c(Y_t, \Theta, t, v)^T \nu_P(\Theta, t, dv), \end{aligned} \right\} \quad (8)$$

где $M_{\Delta^y}^N$ — символ вычисления математического ожидания для нормальных распределений (6) на многообразии Δ^y .

Для стационарных СтС стационарные НСтП — если они существуют, то $m_t = m^*$, $K_t = K^*$, $K(\Theta, t_1, t_2) = k(\Theta, \tau)$ ($\tau = t_1 - t_2$), — определяются уравнениями [10, 11]:

$$\left. \begin{aligned} A^m(m^*, K^*, \Theta) &= 0; & A^K(m^*, K^*, \Theta) &= 0; \\ \dot{k}_\tau(\Theta, \tau) &= A^k(m^*, K^*, \Theta) K^{*-1} k(\Theta, \tau), & k(0) &= K^* \quad (\forall \tau > 0); \\ k(\tau) &= k(-\tau)^T \quad (\forall \tau < 0). \end{aligned} \right\} \quad (9)$$

При этом необходимо, чтобы матрица $A^k = a_{21}(m^*, K^*, \Theta) = a_{21}^*$ была асимптотически устойчивой.

Уравнения МНА в случае СтС (5) переходят в уравнения МСЛ [10, 11]:

$$\left. \begin{aligned} \dot{m}_t &= A^m(m_t, K_t, \Theta, t) = a_1(m_t, K_t, \Theta, t), & m_0 &= m(\Theta, t_0); \\ \dot{K}_t &= A^K(m_t, K_t, \Theta, t) = k_1^a(m_t, K_t, \Theta, t) K_t + \\ &+ K_t k_1^a(m_t, K_t, \Theta, t)^T + \sigma_0(t), & K_0 &= K(\Theta, t_0); \\ \frac{\partial K(\Theta, t_1, t_2)}{\partial t_2} &= A^{K_{t_1, t_2}}(m_{t_2}, K_{t_2}, K(\Theta, t_1, t_2), \Theta, t_2) = \\ &= K(\Theta, t_1, t_2) K_{t_2} k_1^a(m_{t_2}, K_{t_2}, \Theta, t_2)^T, & K(\Theta, t_1, t_1) &= K_{t_1}, \end{aligned} \right\} \quad (10)$$

если принять

$$\left. \begin{aligned} a(Y_t, \Theta, t) &= a_1(m_t, K_t, \Theta) + k_1^a(m_t, K_t, \Theta) Y_t^0; \\ b(Y_t, \Theta, t); & \sigma(Y_t, \Theta, t) = b_0(\Theta, t) \nu(\Theta, t) b_0(\Theta, t)^T = \sigma_0(\Theta, t); \\ k_1^a(m_t, K_t, \Theta, t) &= \left[\left(\frac{\partial}{\partial m_t} \right) A^m(m_t, K_t, \Theta, t)^T \right]^T. \end{aligned} \right\} \quad (11)$$

Для стационарных СтС (5) при условии асимптотической устойчивости матрицы $k_1^a(m^*, K^*)$ в основе МСЛ лежат уравнения (9), записанные в виде:

$$\left. \begin{aligned} A^m(m^*, K^*, \Theta) &= 0; \\ A^K &= k_1^a(m^*, K^*, \Theta) K^* + K^* k_1^a(m^*, K^*, \Theta)^T + \sigma_0^* = 0; \\ \dot{k}_\tau(\Theta, \tau) &= A^{k(\tau)} = k_1^a(m^*, K^*, \Theta) k(\tau), & k(0) &= K^* \quad (\forall \tau > 0); \\ k(\tau) &= k(-\tau)^T \quad (\forall \tau < 0). \end{aligned} \right\} \quad (12)$$

При фиксированном векторе параметров Θ уравнения (6) и (7) лежат в основе алгоритмического обеспечения на базе МНА для СтС на многообразиях

(МСтС) (4), а уравнения (10) и (11) — на базе МСЛ для СтС (5). Для определения стационарных стохастических процессов согласно МНА служат соотношения (9), а МСЛ — (12). При этом требуется существование конечных интегралов (8) и (11).

Для МНА необходимо уметь вычислять следующие интегралы:

$$\begin{aligned} I_0^a &= I_0^a(m_t, K_t, \Theta, t) = A^m(m_t, K_t, \Theta, t) = M_{\Delta y}^N[a(Y_t, \Theta, t)]; \\ I_1^a &= I_1^a(m_t, K_t, \Theta, t) = a_{21}(m_t, K_t, \Theta, t) = M_{\Delta y}^N[a(Y_t, \Theta, t) Y_t^{0T}]; \\ I_0^{\bar{\sigma}} &= I_0^{\bar{\sigma}}(m_t, K_t, \Theta, t) = a_{22}(m_t, K_t, \Theta, t) = M_N[\bar{\sigma}(Y_t, \Theta, t)]. \end{aligned} \quad (13)$$

Для МСЛ достаточно вычислить интеграл (13), причем интеграл I_1^a вычисляется по формуле [10, 11]:

$$k_1^a = k_1^a(m_t, K_t, \Theta, t) = \left[\left(\frac{\partial}{\partial m_t} \right) I_0^a(m_t, K_t, \Theta, t) \right]^T.$$

Как следует из свойства ДРФ 5^0 , базовыми интегралами при вычислении (13) будут следующие интегралы:

$$M_{\Delta y}^N \left[\frac{1}{(Y - Y_\nu)^\mu} \right]; M_{\Delta y}^N \left[\frac{1}{(Y^2 + p_\nu Y + q_\nu)^\mu} \right]; M_{\Delta y}^N \left[\frac{Y}{(Y^2 + p_\nu Y + q_\nu)^\mu} \right]. \quad (14)$$

Подробно численные методы вычисления интегралов (14) описаны в [12–14]. Для дифференциальных СтС с СДРН можно рекомендовать следующие:

- степенные разложения;
- многочленные приближения, в первую очередь на основе многочленов Эрмита;
- дробно-рациональные приближения;
- разложения в цепные дроби;
- асимптотические приближения;
- итеративные процессы различных порядков.

В модуле «StS-CFRN. Analysis» реализованы степенные разложения и многочленные приближения на основе многочленов Эрмита.

В приложении рассмотрены примеры вычисления типовых интегралов (14). В случае $\Delta y = (-\infty, \infty)$ можно рекомендовать квадратурные формулы на основе ортогональных полиномов Эрмита $H_n(x)$. В этом случае используется следующий алгоритм [6]:

$$\int_{-\infty}^{\infty} e^{-x^2} \varphi(x, \pi) dx = \sum_{i=1}^n w_i \varphi(x_i, \pi) + \rho_n, \quad (15)$$

где π — вектор параметров; x_i — нуль $H_n(x)$, w_i — весовой коэффициент:

$$w_i = \frac{2^{n-1} n! \sqrt{\pi}}{n^2 [H_{n-1}(x_i)]^2};$$

ρ_n — остаточный член:

$$\rho_n = \frac{n! \sqrt{\pi}}{2^n (2n)!} \varphi^{(2n)}(\xi) \quad (-\infty < \xi < \infty).$$

Входящий в (15) вектор параметров обычно включает в себя дисперсию D_y и отношение сигнал/шум $\zeta_y = m_y / \sqrt{D_y}$.

Уравнения точности и чувствительности к вектору параметров Θ по методу функций чувствительности являются частным случаем уравнений [15, 16] для функций чувствительности ∇m_t , ∇K_t , $\nabla K(t_1, t_2)$ ($\nabla = \nabla^\Theta$):

$$\begin{aligned} \nabla \dot{m}_t &= \nabla A^m, & \nabla m(t_0) &= 0; \\ \nabla \dot{K}_t &= \nabla A^K, & \nabla K(t_0) &= 0; \\ \frac{\partial \nabla K(t_1, t_2)}{\partial t_2} &= \nabla A^{K(t_1, t_2)}, & \nabla K(t_1, t_1) &= 0. \end{aligned}$$

Уравнения МНА (МСЛ) содержат интегралы I_0^a , I_1^a и I_0^σ в виде соответствующих коэффициентов, поэтому процедура вычисления интегралов должна быть согласована с методом численного решения обыкновенных дифференциальных уравнений для m_t , K_t и $K(t_1, t_2)$. Эти коэффициенты допускают дифференцирование по m_t и K_t , так как под интегралом стоит сглаживающая нормальная плотность.

В [14] изложены алгоритмы аналитического и статистического моделирования распределений (в том числе нормальных) в нелинейных МСтС. Алгоритмы аналитического и статистического моделирования для МСтС с СДРН, а также смешанные алгоритмы различной степени точности относительно шага интегрирования также представлены в [13].

4 Дискретная стохастическая система со сложными дробно-рациональными нелинейностями. Алгоритмическое обеспечение

Рассмотрим дискретную СтС с СДРН, описываемую уравнениями вида [10, 11]:

$$Y_{k+1} = a_k(\Theta, Y_k) + b_k(\Theta, Y_k) V_k^d \quad (k = 1, 2, \dots). \quad (16)$$

Здесь Y_k — $(p \times 1)$ -мерный вектор состояния, $Y_k \in \Delta^y$ (Δ^y — многообразие состояний); функции $a_k(\Theta, Y_k)$ и $b_k(\Theta, Y_k)$ имеют размерности $(p \times 1)$ и $(p \times m)$ соответственно; через V_k^d обозначен векторный дискретный шум, обладающий интенсивностью ν_k^d . В случае аддитивного шума, когда $b_k(\Theta, Y_k) = b_{0k}$, уравнение (16) примет вид:

$$Y_{k+1} = a_k Y_k + b_{0k} V_k^d.$$

В основе МНА лежат следующие соотношения и уравнения [10, 11]:

$$\begin{aligned} g_{1k}^N(\lambda) &= \exp \left\{ i\lambda m_k - \frac{1}{2} \lambda^T K_k \lambda \right\}; \\ g_{k_1 k_2}^N &= \exp \left\{ i\bar{\lambda}^T \bar{m}_2 - \frac{1}{2} \bar{\lambda}^T \bar{K}_2 \bar{\lambda} \right\}; \\ m_{k+1} &= a_{1k} = M_{\Delta^y}^N a_k, \quad m_1 = M_{\Delta^y}^N Y_1; \\ K_{k+1} &= a_{2k} = M_{\Delta^y}^N \left[a_k a_k^T \right] - \left[M_{\Delta^y}^N a_k \right] \left[M_{\Delta^y}^N a_k^T \right] + M_{\Delta^y}^N \left[b_k \nu_k^d b_k^T \right]; \\ K_1 &= M_{\Delta^y}^N \left[Y_1^0 Y_1^{0T} \right]; \\ K(\Theta, l, h) &= a_{3k} = M_{\Delta^y}^N \left[Y_l^0 a_h(Y_h)^T \right], \quad K(\Theta, l, l) = K_l \quad (l < h); \\ K(\Theta, l, h) &= K(\Theta, h, l)^T \quad (l > h). \end{aligned}$$

В основе МСЛ после статистической линейаризации функции $a_k(\Theta, Y_k)$ согласно

$$a_k(\Theta, Y_k) = a_{0k}(\Theta, m_k, K_k) + k_{1k}^a(\Theta, m_k, K_k) Y_k^0$$

будут лежать уравнения [10, 11]:

$$\begin{aligned} m_{k+1} &= a_{0k}, \quad m(1) = m_1; \\ K_{k+1} &= k_{1k}^a K_k (k_{1k}^a)^T + b_{0k} \nu_k^d b_{0k}^T, \quad K(1) = K_1; \\ K(\Theta, l, h+1) &= K(\Theta, l, h) (k_{1h}^a)^T, \quad K(\Theta, l, l) = K_l \quad \text{при } l < h, \\ K(\Theta, l, h) &= K(\Theta, h, l)^T \quad \text{при } l > h. \end{aligned}$$

Для определения стационарных СтП согласно МНА и МСЛ с характеристиками

$$m_k = m^*; \quad K_k = K^*; \quad K(\Theta, l, h) = k^*(r) \quad (r = h - l)$$

используется уравнение

$$m^* = a_{1k}(\Theta, m^*, K^*),$$

а также уравнения

$$\begin{aligned} K^* &= a_{2k}(\Theta, m^*, K^*); \\ K^* &= k_1^a K^* (k_1^a)^T + b_0 \nu_k^d b_0^T; \\ k^*(\Theta, r+1) &= k(r) (k_1^a)^T, \quad k^*(0) = K^*. \end{aligned}$$

Как следует из уравнений МНА, необходимо уметь вычислять следующие интегралы:

$$I_{0k}^a = I_{0k}^a(\Theta, m_k, K_k) = M_{\Delta y}^N [a_k(\Theta, Y_k)]; \quad (17)$$

$$I_{1k}^a = I_{1k}^a(\Theta, m_k, K_k) = M_{\Delta y}^N [a_k(\Theta, Y_k) Y_k^{0T}]; \quad (18)$$

$$I_{0k}^\sigma = I_{0k}^\sigma(\Theta, m_k, K_k) = M_{\Delta y}^N [\sigma(\Theta, Y_k)] \quad (\sigma(\Theta, Y_k) = b_k \nu_k^d b_k^T).$$

Для МСЛ достаточно вычислить интеграл (17), причем интеграл (18) вычисляется по формуле:

$$k_{1k}^a = k_{1k}^a(\Theta, m_k, K_k) = \left[\frac{\partial}{\partial m_k} I_{0k}^a(\Theta, m_k, K_k)^T \right]^T.$$

5 Заключение

Рассмотрены непрерывные и дискретные СтС (в том числе и на гладких многообразиях) с винеровскими и пуассоновскими шумами и с СДРН. Такие модели описывают поведение ряда современных нано- и квантовооптических средств информатики. Приводятся уравнения МНА и МСЛ для аналитического моделирования нестационарных и стационарных нормальных процессов. Рассматриваются методы вычисления типовых интегралов для одно- и многомерных СДРН скалярного и векторного аргумента, получающихся из суперпозиции типовых ДРН.

Методическое и алгоритмическое обеспечение положено в основу модуля «StS-CFRN», входящего в библиотеку «StS-Analysis» [17]. Тестирование модуля проведено с помощью примеров П.1–П.3 и П.6.

Результаты допускают обобщение на случай дискретных СтС, а также интегродифференциальных и операторных СтС с СДРН, в том числе с автокоррелированными шумами.

Приложения

П.1. Для типовой ДРН вида $R^{\text{ДРН}}(y) = 1/(y^2 + b^2)$, учитывая соотношение [3]

$$\int_0^\infty \frac{e^{-h^2 \eta^2}}{\eta^2 + b^2} d\eta = \frac{\pi}{2b} e^{h^2 b^2} \operatorname{erfc}(hb) \quad (h^2 > 0, \quad b^2 > 0),$$

где

$$\operatorname{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^{\infty} e^{-t^2} dt = 1 - \operatorname{erf} x; \quad \operatorname{erf} x = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt,$$

имеем следующее выражение для интеграла (13) при $h^2 = D_y/2$:

$$I_0 = M^N \left[\frac{1}{Y^2 + b^2} \right] = \frac{1}{\sqrt{2\pi D_y}} \int_{-\infty}^{\infty} \frac{e^{-\eta^2/(2D_y)} d\eta}{\eta^2 + b^2} = \frac{\pi}{b} e^{b^2/(2D_y)} \operatorname{erfc} \left(\frac{b}{\sqrt{2D_y}} \right).$$

Учитывая известное неравенство [12, 13]

$$\frac{\xi}{\xi^2 + 1} < \frac{\sqrt{\pi}}{2} e^{\xi^2} \operatorname{erfc}(\xi) < \frac{1}{2\xi},$$

получим следующее неравенство для интеграла $I_0(\xi)$:

$$I_0^-(\xi) < I_0(\xi) < I_0^+(\xi) \quad \left(\xi^2 = \frac{b^2}{2D_y} \right);$$

$$I_0^+ = \frac{\pi}{b} \left(e^{\xi^2} - \frac{2}{\sqrt{\pi}} \frac{\xi}{2\xi^2 + 1} \right), \quad I_0^- = \frac{\pi}{b} \left(e^{\xi^2} - \frac{1}{\sqrt{\pi}\xi} \right).$$

Таким образом, для вычисления этого интеграла используются таблицы функций $\operatorname{erf}(\xi)$ или $\operatorname{erfc}(\xi)$ [12, 13] либо полученные неравенства.

П.2. Для СДРН $R^{\text{СДРН}}(y) = y^2/(y^2 + b^2)$, учитывая соотношение [18]

$$\int_0^{\infty} \frac{\eta^2 e^{-h^2 \eta^2}}{\eta^2 + b^2} d\eta = \frac{\sqrt{\pi}}{2h} - \frac{\pi b}{2} e^{h^2 b^2} [1 - \operatorname{erf}(bh)],$$

получаем

$$I_0^{\text{ДРН}} = \frac{\sqrt{\pi}}{h} - \pi b e^{h^2 b^2} [1 - \operatorname{erf}(bh)] = \sqrt{2\pi D_y} - \pi b e^{b^2/(2D_y)} \left[1 - \operatorname{erf} \left(\frac{b}{\sqrt{2D_y}} \right) \right].$$

П.3. Для СДРН вида $R^{\text{СДРН}}(y) = (y + b)^{-1} \mathbf{1}(y)$ при $h^2 > 0$, $b > 0$, учитывая соотношение [3]

$$\int_0^{\infty} \frac{e^{-h^2 \eta^2}}{\eta + b} d\eta = e^{-h^2 b^2} \left[\sqrt{\pi} \int_0^{h\sqrt{b}} e^{\eta^2} d\eta - \frac{1}{2} \operatorname{Ei}(h^2 b^2) \right]$$

(E_i — интегральная показательная функция [7, 12, 13]), имеем следующее выражение для интеграла (13) при $h^2 = D_y/2$:

$$I_0^{\text{ДРН}} = M^N \left[\frac{1(y)}{Y+b} \right] = \frac{1}{\sqrt{2\pi D_y}} \int_{-\infty}^{\infty} \frac{\mathbf{1}(\eta) e^{-\eta^2/(2D_y)}}{\eta+b} d\eta =$$

$$= \frac{1}{\sqrt{2\pi D_y}} e^{-b^2/(2D_y)} \left[\sqrt{\pi} \int_0^{\sqrt{b/(2D_y)}} e^{\eta^2} d\eta - \frac{1}{2} \text{Ei} \left(\frac{b^2}{2D_y} \right) \right].$$

П.4. Сложные нелинейности, приводимые к СДРН.

1. Если скалярная сложная нелинейность (СН) является ДРФ от $\sin y$ и $\cos y$, то полагают $u = \text{tg } y/2$, так что

$$\sin y = \frac{2u}{1+u^2}; \quad \cos y = \frac{1-u^2}{1+u^2}.$$

2. Если скалярная СН является ДРФ от $\text{sh } y$ и $\text{ch } y$, то полагают $u = \text{th } y/2$, так что

$$\text{sh } y = \frac{2u}{1-u^2}; \quad \text{ch } y = \frac{1+u^2}{1-u^2}.$$

3. Если скалярная СН является ДРФ от y и либо $(1-y^2)^{1/2}$, либо $(y^2-1)^{1/2}$, либо $(y^2+1)^{1/2}$, то задача сводится к случаям 1–2, если сделать соответствующие подстановки $y = \cos v$, или $y = \text{ch } v$, или $y = \text{sh } v$.

4. Если скалярная СН является ДРФ от y и либо $(y^2+1)^{1/2}$, либо $(y^2-1)^{1/2}$, то можно положить соответственно $u = y + (y^2 \pm 1)^{1/2}$. Тогда следует сделать подстановку

$$y = \frac{1}{2} \left(u \pm \frac{1}{u} \right); \quad (y^2 \pm 1)^{1/2} = \frac{1}{2} \left(u \pm \frac{1}{u} \right).$$

5. Если скалярная СН является ДРФ от y и от $(ay^2 + by + c)^{1/2}$, то задача сводится к случаям 3–4, если сделать подстановку

$$v = \frac{2ay + b}{\sqrt{|4ac - b^2|}}; \quad y = \frac{v\sqrt{|4ac - b^2|} - b}{2a}.$$

6. Если скалярная СН является ДРФ от y и $u = \sqrt{(ay + b)/(cy + d)}$, причем $|ad - bc| \neq 0$, то в качестве новой переменной берут u .

П.5. В основе моделей современных средств сейсмоударозащиты лежат упруго-пластические амортизирующие элементы [17, 19]. Как правило, их характеристики описываются формулами П.1–П.4.

П.6. В основе модели лазерного интерферометра лежит следующее уравнение [20]:

$$\ddot{X}_t + \omega_0^2 X_t + \frac{c_0}{X_t^2 + c^2} = a - 2\varepsilon\omega_0 \dot{X}_t + bV. \quad (19)$$

Здесь ε , ω_0 , c_0 , c , a и b — постоянные параметры; V — гауссовский белый шум интенсивности ν .

Полагая $X_t = Y_1$ и $\dot{Y}_1 = Y_2$ и учитывая $\dot{Y}_2 = Y_2(dY_2/dY_1)$, получим следующие дифференциальные соотношения для фазового портрета рассматриваемой динамической системы при $a = 0$, $b = 0$ и $\varepsilon = 0$:

$$\frac{dY_2}{dY_1} = -\frac{\omega_0^2 Y_1(Y_1^2 + c^2) + c_0}{Y_2(Y_1^2 + c^2)}; \quad \frac{Y_2^2}{2} + \Pi(Y_1) = \frac{Y_{20}^2}{2} + \Pi(Y_{10}),$$

где

$$\Pi(Y_1) = \int \left(\frac{\omega_0^2 Y_1}{2} + \frac{c_0}{Y_1^2 + c^2} \right) dY_1 = \frac{1}{2} \omega_0^2 Y_1^2 + \frac{c_0}{c} \operatorname{arctg} \frac{Y_1}{c} + \operatorname{const}. \quad (20)$$

Уравнения (19) при $a = 0$, $b = 1$ и $c_0 > 0$ допускают режим стационарных стохастических колебаний с одномерной плотностью, вычисляемой по формуле Гиббса [10]:

$$f_1^*(y_1, y_2) = \operatorname{const} \cdot e^{-\alpha \mathcal{H}(y_1, y_2)}.$$

Здесь $\mathcal{H} = \mathcal{H}(y_1, y_2)$ — функция Гамильтона $\mathcal{H} = (Y_2^2/2) + \Pi(Y_1)$, где $\Pi(Y_1)$ определена в (20). Отсюда видно, что распределение Y_2 гауссовское, а распределение Y_1 негауссовское.

Уравнения МСЛ для определения вероятностных моментов первого и второго порядка имеют следующий вид:

$$\left. \begin{aligned} \dot{m}_1 &= m_2; & \dot{m}_2 &= -\omega_0^2 m_1 - 2\varepsilon \omega_0 m_2 + R_0(m_1, D_1); \\ \dot{D}_1 &= 2K_{12}; & \dot{D}_2 &= \nu - 2\lambda K_{12} - 4\varepsilon \omega_0 D_2; \\ \dot{K}_{12} &= D_2 - \lambda D_1 - 2\varepsilon \omega_0 K_{12}. \end{aligned} \right\} \quad (21)$$

Здесь

$$\begin{aligned} \lambda &= \lambda(m_1, D_1) = \omega_0^2 \left[1 - \frac{R_1(m_1, D_1)}{\omega_0^2} \right]; \\ R_0(m_1, D_1) &= \frac{1}{\sqrt{2\pi D_1}} \int_{-\infty}^{\infty} \frac{c_0}{\eta^2 + c^2} e^{-(\eta - m_1)^2 / (2D_1)} d\eta = \\ &= \frac{2c_0}{\sqrt{2\pi D_1}} e^{-m_1^2 / (2D_1)} \int_0^{\infty} e^{-\eta^2 / (2D_1) + \eta(m_1 / D_1)} \frac{d\eta}{\eta^2 + c^2}; \\ R_1(m_1, D_1) &= \frac{\partial R_0(m_1, D_1)}{\partial m_1}. \end{aligned}$$

Из (21) при $\varepsilon > 0$ для устойчивого стационарного режима имеем соотношения:

$$m_2^* = 0; \quad m_1^* = R_0(m_1^*, D_1^*) \omega_0^{-2}; \quad D_2^* = \alpha = \frac{\nu^*}{4\varepsilon \omega_0}; \quad K_{12}^* = 0; \quad D_1^* \lambda(m_1^*, D_1^*) = \alpha.$$

При аналитическом моделировании процессов в (21) использовались отрезки следующих степенных представлений функций $R_0 = R_0(m_1, D_1)$, $R_1 = R_1(m_1, D_1)$ и $\lambda = \lambda(m_1, D_1)$:

$$\begin{aligned} R_0 &= \sum_{r=1}^{\infty} \gamma_r^{R_0}(D_1) R_{0r}(\xi_1^2); \\ R_1 &= \sum_{r=1}^{\infty} \gamma_r^{R_1}(D_1) R_{1r}(\xi_1); \\ \lambda &= \lambda(m_1, D_1) = \omega_0^2 \left[1 - \frac{R_1(m_1, D_1)}{\omega_0^2} \right]. \end{aligned}$$

Здесь введены следующие обозначения:

$$\begin{aligned} \gamma_r^{R_0}(D_1) &= \frac{c_0}{2^{2(r+1)}} D_1^r; \quad \gamma_r^{R_1}(D_1) = D_1^{r-1/2}; \\ R_{0r}(\xi_1^2) &= e^{-\xi_1^2/2} \sum_{\rho=1}^{\infty} \frac{[2(r+\rho)-1]!!}{(2\rho)!} \xi_1^{2\rho}; \quad R_{1r}(\xi_1) = 2\xi_1 \frac{\partial R_{0r}(\xi_1^2)}{\partial \xi_1^2}. \end{aligned}$$

При их выводе были использованы следующие интегралы и ряды [7, 10, 13]:

$$\int_0^{\infty} t^{2n+1} e^{-at^2/2} dt = \frac{n!}{2^n a^{n+1}}; \quad \int_0^{\infty} t^{2n} e^{-at^2/2} dt = \frac{(2n-1)!!}{2a^n} \sqrt{\frac{2\pi}{a}};$$

$$\int_{-\infty}^{\infty} t^{2p+1} e^{-(at^2/2)+bt} dt = \int_0^{\infty} t^{2p+1} e^{-at^2/2} (e^{bt} + e^{-bt}) dt;$$

$$e^x + e^{-x} = 2 \sum_{\gamma=1}^{\infty} \frac{x^{2\gamma}}{(2\gamma)!}; \quad (1+x)^{-1} = \sum_{r=0}^{\infty} (-1)^r x^r.$$

Как показано в [17], точность и чувствительность алгоритмов в стационарном режиме при малых отношениях сигнал/шум ξ_1 составляют 0,1%–1%, а при больших — 10%–20%. Для типовых технических применений этого вполне достаточно.

Литература

1. *Сеницын И. Н., Сеницын В. И.* Аналитическое моделирование нормальных процессов в стохастических системах со сложными нелинейностями // Информатика и её применения, 2014. Т. 8. Вып. 3. С. 2–4.

2. *Синицын И. Н., Синицын В. И., Сергеев И. В., Белоусов В. В., Шоргин В. С.* Математическое обеспечение аналитического моделирования стохастических систем со сложными нелинейностями // Системы и средства информатики, 2014. Т. 24. № 3. С. 4–29.
3. *Синицын И. Н., Синицын В. И., Корепанов Э. Р.* Моделирование нормальных процессов в стохастических системах со сложными иррациональными нелинейностями // Информатика и её применения, 2015. Т. 9. Вып. 1. С. 2–8.
4. *Синицын И. Н., Синицын В. И., Сергеев И. В., Корепанов Э. Р., Белоусов В. В., Шоргин В. С.* Математическое обеспечение моделирования нормальных процессов в стохастических системах со сложными иррациональными нелинейностями // Системы и средства информатики, 2015. Т. 25. № 2. С. 3–19.
5. *Синицын И. Н., Синицын В. И., Корепанов Э. Р.* Моделирование нормальных процессов в стохастических системах со сложными трансцендентными нелинейностями // Информатика и её применения, 2015. Т. 9. Вып. 2. С. 23–29.
6. *Синицын И. Н., Синицын В. И., Корепанов Э. Р.* Аналитическое моделирование нормальных процессов в стохастических системах со сложными дробно-рациональными нелинейностями // Информатика и её применения, 2015. Т. 9. Вып. 4. С. 3–17.
7. *Корн Г., Корн Т.* Справочник по математике / Пер. с англ. — М.: Наука, 1984. 831 с. (*Korn G. A., Korn T. M.* Mathematical handbook. — New-York, NY, USA: McGraw-Hill, 1968. 943 p.)
8. *Зверович Э. И.* Вещественный и комплексный анализ: В 6 ч. Кн. 2. Ч. 2. Интегральное исчисление функций скалярного аргумента. — Минск: Высшая школа, 2008. 319 с.
9. *Куликов В.* Рациональная функция // Математическая энциклопедия / Под ред. И. М. Виноградова. — М.: Советская энциклопедия, 1978. Т. 4. С. 918–919.
10. *Пугачёв В. С., Синицын И. Н.* Теория стохастических систем. — М.: Логос, 2000; 2004. 1000 с.
11. *Синицын И. Н., Синицын В. И.* Лекции по нормальной и эллипсоидальной аппроксимации распределений в стохастических системах. — М.: ТОРУС ПРЕСС, 2013. 488 с.
12. Справочник по специальным функциям / Под ред. М. Абрамовича, И. Стигана. — М.: Наука, 1979. 832 с.
13. *Попов Б. А., Теслер Г. С.* Вычисление функций на ЭВМ: Справочник. — Киев: Наукова Думка, 1984. 599 с.
14. *Синицын И. Н.* Параметрическое статистическое и аналитическое моделирование распределений в нелинейных стохастических системах на многообразиях // Информатика и её применения, 2013. Т. 7. Вып. 2. С. 4–16.
15. *Синицын И. Н.* Аналитическое моделирование распределений методом ортогональных разложений в нелинейных стохастических системах на многообразиях // Информатика и её применения, 2015. Т. 9. Вып. 3. С. 17–24.
16. *Синицын И. Н.* Применение ортогональных разложений для аналитического моделирования многомерных распределений в нелинейных стохастических системах на многообразиях // Системы и средства информатики, 2015. Т. 25. № 3. С. 3–22.

17. Синицын И. Н., Сергеев И. В., Синицын В. И., Корепанов Э. Р., Белоусов В. В., Шоргин В. С., Горшенин А. К. Руководство пользователя библиотеки «StS-ANALYSIS» (версия 2.0) / Под ред. И. Н. Синицына. — М.: ИПИ ФИЦ ИУ РАН, 2016 (в печати). Препринт.
18. Градштейн И. С., Рыжик И. М. Таблицы интегралов, сумм, рядов и произведений. — М.: ГИФМЛ, 1963. 1100 с.
19. Синицын И. Н., Сергеев И. В., Синицын В. И., Корепанов Э. Р., Белоусов В. В. Математическое обеспечение параметрического моделирования распределений в интегродифференциальных стохастических системах // Системы и средства информатики, 2014. Т. 24. № 1. С. 4–45.
20. Морозов А. Н., Назолин А. Л. Динамические системы с флуктуирующим временем. — М.: МГТУ им. Н. Э. Баумана, 2001. 200 с.

Поступила в редакцию 23.12.15

MATHEMATICAL SOFTWARE FOR ANALYTICAL MODELING OF NORMAL PROCESSES IN STOCHASTIC SYSTEMS WITH COMPLEX FRACTION-RATIONAL NONLINEARITIES

*I. N. Sinitsyn, V. I. Sinitsyn, I. V. Sergeev, E. R. Korepanov, V. V. Belousov,
and V. S. ShorGIN*

Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilova Str., Moscow 119333, Russian Federation

Abstract: Mathematical software for analytical modeling of normal stochastic processes in differential and discrete stochastic systems (StS) with complex fraction-rational nonlinearities (CFRN) is considered. Typical presentation of CFRN is given. Accuracy and sensitivity equations based on the normal approximation (NAM) and the statistical linearization methods (SLM) are given. Basic analytical and numerical methods for computing NAM (SLM) integrals are discussed. Special attention is paid to the software tool «StS-CFRN. Analysis». The Appendix contains test examples and some applications. Some generalizations are mentioned.

Keywords: analytical modeling; complex fraction-rational nonlinearities (CFRN); module “StS-CFRN. Analysis;” method of statistical linearization (MSL); normal approximation method (NAM); stochastic systems (StS)

DOI: 10.14357/08696527160114

Acknowledgments

The work was supported by the Russian Foundation for Basic Research (project 15-07-02244).

References

1. Sinitsyn, I. N., and V. I. Sinitsyn. 2014. Analiticheskoe modelirovanie normal'nykh protsessov v stokhasticheskikh sistemakh so slozhnymi nelineynostyami [Analytical modeling of normal processes in stochastic systems with complex nonlinearities]. *Informatika i ee Primeneniya — Inform. Appl.* 8(3):2–4.
2. Sinitsyn, I. N., V. I. Sinitsyn, I. V. Sergeev, V. V. Belousov, and V. S. Shorgin. 2014. Matematicheskoe obespechenie analiticheskogo modelirovaniya stokhasticheskikh sistem so slozhnymi nelineynostyami [Mathematical software for analytical modeling of stochastic systems with complex nonlinearities]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 24(3):4–29.
3. Sinitsyn, I. N., V. I. Sinitsyn, and E. R. Korepanov. 2015. Modelirovanie normal'nykh protsessov v stokhasticheskikh sistemakh so slozhnymi irratsional'nymi nelineynostyami [Modeling of normal processes in stochastic systems with complex irrational nonlinearities]. *Informatika i ee Primeneniya — Inform. Appl.* 9(1):2–8.
4. Sinitsyn, I. N., V. I. Sinitsyn, I. V. Sergeev, E. R. Korepanov, V. V. Belousov, and V. S. Shorgin. 2015. Matematicheskoe obespechenie modelirovaniya normal'nykh protsessov v stokhasticheskikh sistemakh so slozhnymi irratsional'nymi nelineynostyami [Mathematical software for modeling of normal processes in stochastic systems with complex irrational nonlinearities]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 25(2): 3–19.
5. Sinitsyn, I. N., V. I. Sinitsyn, and E. R. Korepanov. 2015. Modelirovanie normal'nykh protsessov v stokhasticheskikh sistemakh so slozhnymi transtsendentnymi nelineynostyami [Modeling of normal processes in stochastic systems with complex transcendental nonlinearities]. *Informatika i ee Primeneniya — Inform. Appl.* 9(2):23–29.
6. Sinitsyn, I. N., V. I. Sinitsyn, and E. R. Korepanov. 2015. Analiticheskoe modelirovanie normal'nykh protsessov v stokhasticheskikh sistemakh so slozhnymi drobnoratsional'nymi nelineynostyami [Analytical modeling of normal processes in stochastic systems with complex fractional-rational nonlinearities]. *Informatika i ee Primeneniya — Inform. Appl.* 9(4):3–17.
7. Korn, G. A., and T. M. Korn. 1968. *Mathematical handbook*. New-York, NY: McGraw-Hill. 943 p.
8. Zverovich, E. I. 2008. *Veshchestvennyy i kompleksnyy analiz: V 6 ch. Kn. 2. Ch. 2. Integral'noe ischislenie funktsiy skalyarnogo argumenta* [Real and complex analysis: In 6 parts. Bk. 2. Pt. 2. Integral calculus of scalar argument]. Minsk: Vysshaya Shkola. 319 p.
9. Kulikov, V. 1978. Ratsional'naya funktsiya [Rational function]. *Matematicheskaya entsiklopediya* [Mathematical encyclopedia]. Ed. I. M. Vinogradov. Moscow: Sovetskaya Entsiklopediya. 4:918–919.
10. Pugachev, V. S., and I. N. Sinitsyn. 2001. *Stochastic systems. Theory and applications*. Singapore: World Scientific. 908 p.
11. Sinitsyn, I. N., and V. I. Sinitsyn. 2013. *Lektsii po normal'noy i ellipsoidal'noy approksimatsii raspredeleniy v stokhasticheskikh sistemakh* [Lectures on normal and ellipsoidal approximation of distributions in stochastic systems]. Moscow: TORUS PRESS. 488 p.
12. Abramovich, M., and I. Stigan, eds. 1979. *Spravochnik po spetsial'nym funktsiyam* [Handbook on special functions]. Moscow: Nauka. 832 p.

13. Popov, B. A., and G. S. Tesler. 1984. *Vychislenie funktsiy na EVM: Spravochnik* [Calculation of functions on a computer: Handbook]. Kiev: Naukova Dumka. 599 p.
14. Sinitsyn, I. N. 2013. Parametricheskoe statisticheskoe i analiticheskoe modelirovanie raspredeleniy v nelineynykh stokhasticheskikh sistemakh na mnogoobraznykh [Parametric statistical and analytical modeling of distributions in stochastic systems on manifolds]. *Informatika i ee Primeneniya — Inform. Appl.* 7(2):4–16.
15. Sinitsyn, I. N. 2015. Analiticheskoe modelirovanie raspredeleniy metodom ortogonal'nykh razlozheniy v nelineynykh stokhasticheskikh sistemakh na mnogoobraznykh [Analytical modeling of distributions by orthogonal expansions method]. *Informatika i ee Primeneniya — Inform. Appl.* 9(3):17–24.
16. Sinitsyn, I. N. 2015. Primenenie ortogonal'nykh razlozheniy dlya analiticheskogo modelirovaniya mnogomernykh raspredeleniy v nelineynykh stokhasticheskikh sistemakh na mnogoobraznykh [Application of orthogonal expansions for analytical modeling of multidimensional distributions in nonlinear stochastic systems on manifolds]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 25(3):3–22.
17. Sinitsyn, I. N., I. V. Sergeev, V. I. Sinitsyn, E. R. Korepanov, V. V. Belousov, V. S. Shorgin, and A. K. Gorshenin. 2016 (in press). *Rukovodstvo biblioteki “StS-ANALYSIS” (versiya 2.0)* [“StS-ANALYSIS” (version 2.0 Manual)]. Ed. I. N. Sinitsyn. Moscow: IPI FITs IU RAN. Preprint.
18. Gradshteyn, I. S., and I. M. Ryzhik. 1963. *Tablitsy integralov, summ, ryadov i proizvedeniy* [Tables of integrals, sums, series, and products]. Moscow: GIFML. 1100 p.
19. Sinitsyn, I. N., I. V. Sergeev, V. I. Sinitsyn, E. R. Korepanov, and V. V. Belousov. 2014. Matematicheskoe obespechenie parametricheskogo modelirovaniya raspredeleniy v integrodifferentsial'nykh stokhasticheskikh sistemakh [Mathematical software for parametric modeling of distributions in integrodifferential stochastic systems]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 24(1):4–45.
20. Morozov, A. N., and A. L. Nazolin. 2001. *Dinamicheskie sistemy s fluktuiruyushchim vremenem* [Dynamical systems with time fluctuating]. Moscow: Izd-vo MGU im. N. E. Bauman. 200 p.

Received December 23, 2015

Contributors

Sinitsyn Igor N. (b. 1940) — Doctor of Science in technology, professor, Honored scientist of RF, Head of Department, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; sinitsin@dol.ru

Sinitsyn Vladimir I. (b. 1968) — Doctor of Science in physics and mathematics, associate professor, Head of Department, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; VSinitsyn@ipiran.ru

Sergeev Igor V. (b. 1965) — Candidate of Science (PhD) in technology, Deputy Director, Federal Research Center “Computer Science and Control” of the Russian

Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; ISergeev@ipiran.ru

Korepanov Eduard R. (b. 1966) — Candidate of Science (PhD) in technology, Head of Laboratory, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; Ekorepanov@ipiran.ru

Belousov Vasiliy V. (b. 1977) — Candidate of Science (PhD) in technology, Head of Laboratory, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; VBelousov@ipiran.ru

Shorgin Vsevolod S. (b. 1978) — Candidate of Science (PhD) in technology, senior scientist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; VShorgin@ipiran.ru

Агаев Фаик Бахадур оглы (р. 1986) — заведующий отделом Правового просвещения, научно-аналитического, информационного и международных отношений Аппарата Уполномоченного по правам человека (Омбудсмана) Азербайджанской Республики

Адамович Игорь Михайлович (р. 1934) — кандидат технических наук, заведующий отделом Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Азанов Виталий Геннадьевич (р. 1983) — начальник отдела АСУ Краевого государственного бюджетного учреждения здравоохранения «Краевая клиническая больница»

Андреев Сергей Евгеньевич (р. 1990) — младший научный сотрудник Института общей физики Российской академии наук; аспирант Московского технологического университета (МИРЭА)

Аникин Валерий Иванович (р. 1949) — доктор технических наук, профессор Поволжского государственного университета сервиса

Аникина Оксана Владимировна (р. 1978) — кандидат технических наук, доцент кафедры прикладной математики и информатики Тольяттинского государственного университета

Анисимов Владимир Иванович (р. 1926) — доктор технических наук, профессор Санкт-Петербургского государственного электротехнического университета «ЛЭТИ»; главный научный сотрудник Центра информационных технологий в проектировании Российской академии наук (ЦИТП РАН)

Ахмад Алтаиб Давод (р. 1982) — аспирант Санкт-Петербургского государственного электротехнического университета «ЛЭТИ»

Белоусов Василий Владимирович (р. 1977) — кандидат технических наук, заведующий сектором Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Горшенин Андрей Константинович (р. 1986) — кандидат физико-математических наук, старший научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук; доцент Московского технологического университета (МИРЭА)

Гоюшов Айдын Ислам оглы (р. 1978) — докторант Института систем управления Национальной академии наук Азербайджана

Гридин Владимир Николаевич (р. 1944) — доктор технических наук, научный руководитель Центра информационных технологий в проектировании Российской академии наук (ЦИТП РАН)

Джамалов Зейнал Рамазан оглы (р. 1959) — кандидат физико-математических наук, доцент кафедры Бакинского государственного университета

Дулин Сергей Константинович (р. 1950) — доктор технических наук, профессор; ведущий научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Дулина Наталья Георгиевна (р. 1947) — кандидат технических наук, старший научный сотрудник Вычислительного центра им. А. А. Дородницына Федерального исследовательского центра «Информатика и управление» Российской академии наук

Егоров Владимир Борисович (р. 1948) — кандидат технических наук, ведущий научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Зацаринный Александр Алексеевич (р. 1951) — доктор технических наук, заместитель директора Федерального исследовательского центра «Информатика и управление» Российской академии наук

Земсков Дмитрий Вячеславович (р. 1969) — ведущий программист Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Калинин Юрий Павлович (р. 1948) — доктор технических наук, профессор, заведующий лабораторией Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Карманова Александра Александровна (р. 1986) — инженер-программист ООО «НетКрэкер»

Коновалов Григорий Михайлович (р. 1980) — кандидат физико-математических наук, ведущий научный сотрудник Частного учреждения Государственной корпорации по атомной энергии «Росатом» «Проектный центр ИТЭР»

Корепанов Эдуард Рудольфович (р. 1966) — кандидат технических наук, заведующий сектором Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Королёв Вадим Иванович (р. 1943) — доктор технических наук, профессор, ведущий научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук; профессор Национального исследовательского ядерного университета «МИФИ»

Кучин Александр Сергеевич (р. 1993) — студент Московского технологического университета (МИРЭА)

Никишин Дмитрий Александрович (р. 1976) — кандидат технических наук, заведующий сектором Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Рзаев Рамин Рза оглы (р. 1961) — доктор технических наук, профессор, главный научный сотрудник Института систем управления Национальной академии наук Азербайджана

Сергеев Игорь Викторович (р. 1965) — кандидат технических наук, заместитель директора Федерального исследовательского центра «Информатика и управление» Российской академии наук

Синицын Владимир Игоревич (р. 1968) — доктор физико-математических наук, доцент, заведующий отделом Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Синицын Игорь Николаевич (р. 1940) — доктор технических наук, профессор, заслуженный деятель науки РФ, заведующий отделом Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Хорошилов Александр Алексеевич (р. 1952) — доктор технических наук, профессор, ведущий научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Хорошилов Алексей Алексеевич (р. 1958) — кандидат технических наук, научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Шоргин Всеволод Сергеевич (р. 1978) — кандидат технических наук, старший научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Правила подготовки рукописей статей для публикации в журнале «Системы и средства информатики»

Журнал «Системы и средства информатики» публикует теоретические, обзорные и дискуссионные статьи, посвященные научным исследованиям и разработкам в области информационных технологий.

Журнал издается на русском языке. По специальному решению редколлегии отдельные статьи могут печататься на английском языке.

Тематика журнала охватывает следующие направления:

- информационно-телекоммуникационные системы и средства их построения;
- архитектура и программное обеспечение вычислительных машин, комплексов и сетей;
- методы и средства защиты информации.

1. В журнале печатаются статьи, содержащие результаты, ранее не опубликованные и не предназначенные к одновременной публикации в других изданиях.

Публикация не должна нарушать закон об авторских правах.

Направляя рукопись в редакцию, авторы сохраняют все права собственников данной рукописи и при этом передают учредителям и редколлегии неисключительные права на издание статьи на русском языке (или на языке статьи, если он отличен от русского) и на ее распространение в России и за рубежом. Авторы должны предоставить в редакцию письмо в следующей форме:

Соглашение о передаче права на публикацию:

«Мы, нижеподписавшиеся, авторы рукописи «. . .», передаем учредителям и редколлегии журнала «Системы и средства информатики» неисключительное право опубликовать данную рукопись статьи на русском языке как в печатной, так и в электронной версиях журнала. Мы подтверждаем, что данная публикация не нарушает авторских прав других лиц или организаций, а также не содержит сведений, запрещенных к опубликованию в открытой печати.

Подписи авторов: (ф. и. о., дата, адрес)».

Это соглашение может быть предоставлено в бумажном виде или в виде отсканированной копии (с подписями авторов).

Редколлегия вправе запросить у авторов экспертное заключение о возможности публикации представленной статьи в открытой печати.

2. К статье прилагаются данные автора (авторов) (см. п. 8). При наличии нескольких авторов указывается фамилия автора, ответственного за переписку с редакцией.
3. Редакция журнала осуществляет экспертизу присланных статей в соответствии с принятой в журнале процедурой рецензирования.

Возвращение рукописи на доработку не означает ее принятия к печати.

Доработанный вариант с ответом на замечания рецензента необходимо прислать в редакцию.

4. Решение редколлегии о публикации статьи или ее отклонении сообщается авторам. Редколлегия может также направить авторам текст рецензии на их статью. Дискуссия по поводу отклоненных статей не ведется.

5. Редактура статей высылается авторам для просмотра. Замечания к редакции должны быть присланы авторами в кратчайшие сроки.
6. Рукопись предоставляется в электронном виде в форматах MS WORD (.doc или .docx) или L^AT_EX (.tex), дополнительно — в формате .pdf, на дискете, лазерном диске или электронной почтой. Предоставление бумажной рукописи необязательно.
7. При подготовке рукописи в MS Word рекомендуется использовать следующие настройки.

Параметры страницы: формат — А4; ориентация — книжная; поля (см): внутри — 2,5, снаружи — 1,5, сверху и снизу — 2, от края до нижнего колонтитула — 1,3.

Основной текст: стиль — «Обычный», шрифт — Times New Roman, размер — 14 пунктов, абзацный отступ — 0,5 см, 1,5 интервала, выравнивание — по ширине.

Рекомендуемый объем рукописи — не свыше 20 страниц указанного формата.

Сокращения слов, помимо стандартных, не допускаются. Допускается минимальное количество аббревиатур.

Все страницы рукописи нумеруются.

Шаблоны примеров оформления представлены в Интернете:

<http://www.ipiran.ru/publications/collected/template.doc>

8. Статья должна содержать следующую информацию на **русском и английском языках**:

- название статьи;
- Ф.И.О. авторов, на английском можно только имя и фамилию;
- место работы, с указанием города и страны и электронного адреса каждого автора;
- сведения об авторах, в соответствии с форматом, образцы которого представлены на страницах:
http://www.ipiran.ru/journal/collected/2012_22_02_rus/authors.asp и
http://www.ipiran.ru/journal/collected/2012_22_02_eng/authors.asp;
- аннотация (не менее 100 слов на каждом из языков). Аннотация — это краткое резюме работы, которое может публиковаться отдельно. Она является основным источником информации в информационных системах и базах данных. Английская аннотация должна быть оригинальной, может не быть дословным переводом русского текста и должна быть написана хорошим английским языком. В аннотации не должно быть ссылок на литературу и, по возможности, формул;
- ключевые слова — желательно из принятых в мировой научно-технической литературе тематических тезаурусов. Предложения не могут быть ключевыми словами.
- источники финансирования работы (ссылки на гранты, проекты, поддерживающие организации и т. п.).

9. Требования к спискам литературы.

Ссылки на литературу в тексте статьи нумеруются (в квадратных скобках) и располагаются в каждом из списков литературы в порядке первых упоминаний.

Списки литературы представляются в двух вариантах:

- (1) **Список литературы к русскоязычной части.** Русские и английские работы — на языке и в алфавите оригинала.
- (2) **References.** Русские работы и работы на других языках — в латинской транслитерации с переводом на английский язык; английские работы и работы на других языках — на языке оригинала.

Необходимо для составления списка “References” пользоваться размещенной на сайте <http://www.translit.ru/> бесплатной программой транслитерации русского текста в латиницу, при этом в закладке «варианты. . . » следует выбрать опцию BNG.

Список литературы “References” приводится полностью отдельным блоком, повторяя все позиции из списка литературы к русскоязычной части, независимо от того, имеются или нет в нем иностранные источники. Если в списке литературы к русскоязычной части есть ссылки на иностранные публикации, набранные латиницей, они полностью повторяются в списке “References”.

Примеры ссылок на различные виды публикаций в списке “References”:

Описание статьи из журнала:

Zhang, Z., and D. Zhu. 2008. Experimental research on the localized electrochemical micromachining. *Rus. J. Electrochem.* 44(8):926–930. doi:10.1134/S1023193508080077.

Описание статьи из электронного журнала:

Swaminathan, V., E. Lepkoswka-White, and B. P. Rao. 1999. Browsers or buyers in cyberspace? An investigation of electronic factors influencing electronic exchange. *JCMC* 5(2). Available at: <http://www.ascusc.org/jcmc/vol5/issue2/> (accessed April 28, 2011).

Описание материалов конференций:

Usmanov, T. S., A. A. Gusmanov, I. Z. Mullagalin, R. Ju. Muhametshina, A. N. Chervyakova, and A. V. Sveshnikov. 2007. Osobennosti proektirovaniya razrobotki mestorozhdeniy s primeneniem gidrorazryva plasta [Features of the design of field development with the use of hydraulic fracturing]. *Trudy 6-go Mezhdunarodnogo Simpoziuma “Novye resursosberegayushchie tekhnologii nedropol’zovaniya i povyshe-niya neftegazootdachi”* [6th Symposium (International) “New Energy Saving Subsoil Technologies and the Increasing of the Oil and Gas Impact” Proceedings]. Moscow. 267–272.

Описание книги (монографии, сборники):

Lindorf, L. S., and L. G. Mamikonians, eds. 1972. *Ekspluatatsiya turbogeneratorov s neposredstvennym okhlazhdeniem* [Operation of turbine generators with direct cooling]. Moscow: Energy Publs. 352 p.

Описание переводной книги (в списке литературы к русскоязычной части необходимо указать: / Пер. с англ. — после названия книги, а в конце ссылки указать оригинал книги в круглых скобках):

1. В русскоязычной части:

Тимошенко С. П., Янг Д. Х., Уивер У. Колебания в инженерном деле / Пер. с англ. — М.: Машиностроение, 1985. 472 с. (*Timoshenko S. P., Young D. H., Weaver W.* Vibration problems in engineering. — 4th ed. — N.Y.: Wiley, 1974. 521 p.)

2. В англоязычной части:

Timoshenko, S. P., D. H. Young, and W. Weaver. 1974. *Vibration problems in engineering*. 4th ed. N.Y.: Wiley. 521 p.

Описание неопубликованного документа:

Latypov, A. R., M. M. Khasanov, and V. A. Baikov. 2004. Geology and production (NGT GiD). Certificate on official registration of the computer program No. 2004611198. (In Russian, unpubl.)

Описание интернет-ресурса:

Pravila tsitirovaniya istochnikov [Rules for the citing of sources]. Available at: <http://www.scribd.com/doc/1034528/> (accessed February 7, 2011).

Описание диссертации или автореферата диссертации:

Semenov, V. I. 2003. Matematicheskoe modelirovanie plazmy v sisteme kompaktnyy tor [Mathematical modeling of the plasma in the compact torus]. D.Sc. Diss. Moscow. 272 p.

Kozhunova, O. S. 2009. Tekhnologiya razrabotki semanticheskogo slovary informatzionnogo monitoringa [Technology of development of semantic dictionary of information monitoring system]. PhD Thesis. Moscow: IPI RAN. 23 p.

Описание ГОСТа:

GOST 8.586.5-2005. 2007. Metodika vypolneniya izmereniy. Izmerenie raskhoda i kolichestva zhidkostey i gazov s pomoshch'yu standartnykh suzhayushchikh ustroystv [Method of measurement. Measurement of flow rate and volume of liquids and gases by means of orifice devices]. Moscow: Standardinform Publ. 10 p.

Описание патента:

Bolshakov, M. V., A. V. Kulakov, A. N. Lavrenov, and M. V. Palkin. 2006. Sposob orientirovaniya po krenu letatel'nogo apparata s opticheskoy golovkoy samonavedeniya [The way to orient on the roll of aircraft with optical homing head]. Patent RF No. 2280590.

10. Присланные в редакцию материалы авторам не возвращаются.

11. При отправке файлов по электронной почте просим придерживаться следующих правил:

- указывать в поле subject (тема) название журнала и фамилию автора;
- использовать attach (присоединение);
- в состав электронной версии статьи должны входить: файл, содержащий текст статьи, и файл(ы), содержащий(е) иллюстрации.

12. Журнал «Системы и средства информатики» является некоммерческим изданием. Плата за публикацию не взимается, гонорар авторам не выплачивается.

Адрес редакции журнала «Системы и средства информатики»:

Москва 119333, ул. Вавилова, д. 44, корп. 2, ФИЦ ИУ РАН

Тел.: +7 (499) 135-86-92 Факс: +7 (495) 930-45-05

e-mail: rust@ipiran.ru (Сейфуль-Мулюков Рустем Бадриевич)

<http://www.ipiran.ru/journal/collected>

Requirements for manuscripts submitted to Journal “Systems and Means of Informatics”

Journal “Systems and Means of Informatics” publishes theoretical, review, and discussion articles on the research and development in the field of information technology.

The journal is published in Russian. By a special decision of the editorial board, some articles can be published in English.

Topics covered include the following areas:

- information and communication systems and tools of their design;
- architecture and software of computational complexes and networks; and
- methods and tools of information protection.

1. The Journal publishes original articles which have not been published before and are not intended for simultaneous publication in other editions. An article submitted to the Journal must not violate the Copyright law. Sending the manuscript to the Editorial Board, the authors retain all rights of the owners of the manuscript and transfer the nonexclusive rights to publish the article in Russian (or the language of the article, if not Russian) and its distribution in Russia and abroad to the Founders and the Editorial Board. Authors should submit a letter to the Editorial Board in the following form:

Agreement on the transfer of rights to publish:

“We, the undersigned authors of the manuscript “. . . ,” pass to the Founder and the Editorial Board of the Journal “Systems and Means of Informatics” the nonexclusive right to publish the manuscript of the article in Russian (or in English) in both print and electronic versions of the Journal. We affirm that this publication does not violate the Copyright of other persons or organizations.

Author(s) signature(s): (name(s), address(es), date).”

This agreement should be submitted in paper form or in the form of a scanned copy (signed by the authors).

The Editorial Board has the right to request from the authors an official expert conclusion that the submitted article has no classified data prohibited for publication.

2. A submitted article should be attached with **the data on the author(s)** (see item 8). If there are several authors, the contact person should be indicated who is responsible for correspondence with the Editorial Board and other authors about revisions and final approval of the proofs.
3. The Editorial Board of the Journal examines the article according to the established reviewing procedure. If authors receive their article for correction after reviewing, it does not mean that the article is approved to be published. The corrected article should be sent to the Editorial Board for the subsequent review and approval.
4. The decision on the article publication or its rejection is communicated to the authors. The Editorial Board may also send the reviews on the submitted articles to the authors. Any discussion upon the rejected articles is not possible.
5. The edited articles will be sent to the authors for proofread. The comments of the authors to the edited text of the article should be sent to the Editorial Board as soon as possible.

6. The manuscript of the article should be presented electronically in the MS WORD (.doc or .docx) or L^AT_EX (.tex) formats, and additionally in the .pdf format. All documents may be sent by e-mail or provided on a CD or diskette. A hard copy submission is not necessary.
7. The recommended typesetting instructions for manuscript.

Pages parameters: format A4, portrait orientation, document margins (cm): left — 2.5, right — 1.5, above — 2.0, below — 2.0, footer 1.3.

Text: font — Times New Roman, font size — 14, paragraph indent — 0.5, line spacing — 1.5, justified alignment.

The recommended manuscript size: not more than 20 pages of the specified format.

Use only standard abbreviations. Avoid abbreviations in the title and abstract. The full term for which an abbreviation stands should precede its first use in the text unless it is a standard unit of measurement.

All pages of the manuscript should be numbered.

The templates for the manuscript typesetting are presented on site:
<http://www.ipiran.ru/publication/collected/template.doc>
8. Articles should enclose data both in **Russian and English**:
 - title;
 - author's name and surname;
 - affiliation — organization, its address with ZIP code, city, country, and official e-mail address;
 - data on authors according to the format (see site):
http://www.ipiran.ru/journal/collected/2012_22_02_rus/authors.asp and
http://www.ipiran.ru/journal/collected/2012_22_02_eng/authors.asp;
 - abstract (not less than 100 words) both in Russian and in English. Abstract is a short summary of the article that can be published separately. The abstract is the main source of information on the article and it could be included in leading information systems and data bases. The abstract in English has to be an original text and should not be an exact translation of the Russian one. Good English is required. In abstracts, avoid references and formulae.
 - Indexing is performed on the basis of keywords. The use of keywords from the internationally accepted thematic Thesauri is recommended.
Important! Keywords must not be sentences.
 - Acknowledgments.
9. References. Russian references have to be presented both in English translation and in Latin transliteration (refer <http://www.translit.ru>, option BGN).

Please take into account the following examples of Russian references appearance:

Article in journal:
Zhang, Z., and D. Zhu. 2008. Experimental research on the localized electrochemical micromachining. *Rus. J. Electrochem.* 44(8):926–930. doi:10.1134/S1023193508080077.

Journal article in electronic format:
Swaminathan, V., E. Lepkoswka-White, and B. P. Rao. 1999. Browsers or buyers in cyberspace? An investigation of electronic factors influencing electronic

exchange. *JCMC* 5(2). Available at: <http://www.ascusc.org/jcmc/vol5/issue2/> (accessed April 28, 2011).

Conference proceedings:

Usmanov, T. S., A. A. Gusmanov, I. Z. Mullagalin, R. Ju. Muhametshina, A. N. Chervyakova, and A. V. Sveshnikov. 2007. Osobennosti proektirovaniya razrabotki mestorozhdeniy s primeneniem gidrorazryva plasta [Features of the design of field development with the use of hydraulic fracturing]. *Trudy 6-go Mezhdunarodnogo Simpoziuma "Novye resursosberegayushchie tekhnologii nedropol'zovaniya i povyshe-niya neftegazootdachi"* [6th Symposium (International) "New Energy Saving Subsoil Technologies and the Increasing of the Oil and Gas Impact" Proceedings]. Moscow. 267–272.

Books and other monographs:

Lindorf, L. S., and L. G. Mamikonians, eds. 1972. *Ekspluatatsiya turbogenera-torov s neposredstvennym okhlazhdeniem* [Operation of turbine generators with direct cooling]. Moscow: Energy Publs. 352 p.

Dissertation and Thesis:

Kozhunova, O. S. 2009. Tekhnologiya razrabotki semanticheskogo slovarya informat-sionnogo monitoringa [Technology of development of semantic dictionary of information monitoring system]. PhD Thesis. Moscow: IPI RAN. 23 p.

State standards and patents:

GOST 8.586.5-2005. 2007. Metodika vypolneniya izmereniy. Izmerenie raskhoda i kolichestva zhidkostey i gazov s pomoshch'yu standartnykh suzhayushchikh ustroystv [Method of measurement. Measurement of flow rate and volume of liquids and gases by means of orifice devices]. M.: Standardinform Publs. 10 p.

Bolshakov, M. V., A. V. Kulakov, A. N. Lavrenov, and M. V. Palkin. 2006. Sposob orientirovaniya po krenu letatel'nogo apparata s opticheskoy golovkoy samonavedeniya [The way to orient on the roll of aircraft with optical homing head]. Patent RF No. 2280590.

References in Latin transcription are presented in the original language.

References in the text are numbered according to the order of their first appearance; the number is placed in square brackets. All items from the reference list should be cited.

10. Manuscripts and additional materials are not returned to Authors by the Editorial Board.
11. Submissions of files by e-mail must include:
 - the journal title and author's name in the "Subject" field;
 - an article and additional materials have to be attached using the "attach" function;
 - an electronic version of the article should contain the file with the text and a separate file with figures.
12. "System and Means of Informatics" journal is not a profit publication. There are no charges for the authors as well as there are no royalties.

Editorial Board address:

FRC CSC RAS, 44, block 2, Vavilov Str., Moscow 119333, Russia

Ph.: +7 (499) 135 86 92, Fax: +7 (495) 930 45 05

e-mail: rust@ipiran.ru (to Prof. Rustem Seyful-Mulyukov)

http://www.ipiran.ru/english/journal_systems.asp

SYSTEMS AND MEANS OF INFORMATICS (SISTEMY I SREDSTVA INFORMATIKI)

SCIENTIFIC JOURNAL

Volume 26 No.1 Year 2016

Editor-in-Chief and Chair of Editorial Council
Academician I. A. Sokolov

IN THIS ISSUE:

SPARSE BUFFERS — USING THE VIRTUAL MEMORY MECHANISM TO REDUCE PHYSICAL MEMORY AND CPU TIME USAGE <i>I. M. Adamovich and D. V. Zemskov</i>	4
STRUCTURAL AND FUNCTIONAL MODEL OF PATIENT FLOW CONTROL IN A REGIONAL CONSULTATION CLINIC <i>V. G. Azanov</i>	13
ALGORITHM AND PROGRAM COMPLEX FOR ANALYSIS OF THE CHARACTERISTICS OF PLASMA RELATIVISTIC ULTRAHIGH FREQUENCY GENERATOR OUTPUT RADIATION <i>S. E. Andreev</i>	30
THE EXPERIMENTAL ANALYSIS OF THE METHOD OF CLUSTERING AND RANKING OF MULTIDIMENSIONAL DATA USING THE KOHONEN NEURAL NETWORK <i>V. I. Anikin, O. V. Anikina, and A. A. Karmanova</i>	44
ON A REALIZATION OF AN AUTOMATED TESTING SERVICE <i>A. K. Gorshenin and A. S. Kuchin</i>	62
BUILDING .NET-CLIENT APPLICATIONS IN DISTRIBUTED CIRCUIT COMPUTER-AIDED DESIGN <i>V. N. Gridin, V. I. Anisimov, and A. D. Ahmad</i>	76
PROBLEMS OF MAINTENANCE OF SEMANTIC GEOINTEROPERABILITY AND COORDINATION OF UNDERSTANDING OF GEODATA SEMANTICS <i>S. K. Dulin, N. G. Dulina, and D. A. Nikishin</i>	86
SOME ISSUES OF THE SDN CONCEPT PRACTICAL IMPLEMENTATION <i>V. B. Egorov</i>	109