

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ

10(158)
2009

ТЕОРЕТИЧЕСКИЙ И ПРИКЛАДНОЙ НАУЧНО-ТЕХНИЧЕСКИЙ ЖУРНАЛ

Издается с ноября 1995 г.

УЧРЕДИТЕЛЬ
Издательство "Новые технологии"

СОДЕРЖАНИЕ

БЕЗОПАСНОСТЬ ИНФОРМАЦИИ

- Васенин В. А. Критическая энергетическая инфраструктура: средства противодействия кибертерроризму 2
Ремизов А. В., Филиппов М. В., Чичварин Н. В. Методы защиты информации в звуковых файлах 10
Шапченко К. А. Способ проверки свойств безопасности в моделях логического разграничения доступа с древовидной иерархией объектов доступа 13

СЕТИ И СИСТЕМЫ СВЯЗИ

- Мочалов В. А. Алгоритмы размещения транзитных узлов в сенсорной сети . . . 18
Огнев В. А., Иванов С. Р. Математическая модель приемовычислителя систем ГЛОНАСС/GPS 24

СИСТЕМЫ АВТОМАТИЗИРОВАННОГО ПРОЕКТИРОВАНИЯ

- Петунин А. А. Автоматизация нерегулярного фигурного раскроя: основные подходы к разработке систем проектирования раскроя и управляющих программ для машин резки листовых материалов 29
Немтинов В. А., Егоров С. Я., Пахомов П. И. Применение теории нечетких множеств и экспертных систем при автоматизированном выборе элемента технической системы 34

ИНТЕЛЛЕКТУАЛЬНЫЕ СИСТЕМЫ

- Альгулиев Р. М. Реферирование документов через ранжирование предложений . . 39
Данчул А. Н., Новиков А. П. Представление знаний о предметной области на основе гносеологической модели 44

МОДЕЛИРОВАНИЕ И ОПТИМИЗАЦИЯ

- Мухачёва Э. А., Заминова А. Р. Исследование эффективности эволюционных алгоритмов для решения задач дискретной оптимизации на примере задачи коммивояжера 51
Воловикова Е. В., Увайсов С. У. Метод диагностирования радиоэлектронных устройств на основе комплексного электротеплового моделирования 57
Секаев В. Г. Использование алгоритмов комбинирования эвристик при построении оптимальных расписаний 61

ПРИКЛАДНЫЕ ИНФОРМАЦИОННЫЕ СИСТЕМЫ

- Дорожко В. М. Информационная технология идентификации начального этапа разгона морского судна 65
Агаев Н. Б. Диагностирование предаварийного состояния систем газоснабжения на основе теории нечетких множеств 71

ИНФОРМАЦИЯ

- Contents 77
Приложение. Штрик А. А. Электронные технологии в деятельности органов государственной власти России: анализ и перспективы развития.

Главный редактор
НОРЕНКОВ И. П.

Зам. гл. редактора
ФИЛИМОНОВ Н. Б.

Редакционная
коллегия:

АВДОШИН С. М.
АНТОНОВ Б. И.
БАТИЩЕВ Д. И.
БАРСКИЙ А. Б.
БОЖКО А. Н.
ВАСЕНИН В. А.
ГАЛУШКИН А. И.
ГЛОРИОЗОВ Е. Л.
ГОРБАТОВ В. А.
ДОМРАЧЕВ В. Г.
ЗАГИДУЛЛИН Р. Ш.
ЗАРУБИН В. С.
ИВАННИКОВ А. Д.
ИСАЕНКО Р. О.
КОЛИН К. К.
КУЛАГИН В. П.
КУРЕЙЧИК В. М.
ЛЬВОВИЧ Я. Е.
МАЛЬЦЕВ П. П.
МЕДВЕДЕВ Н. В.
МИХАЙЛОВ Б. М.
НАРИНЬЯНИ А. С.
НЕЧАЕВ В. В.
ПАВЛОВ В. В.
ПУЗАНКОВ Д. В.
РЯБОВ Г. Г.
СОКОЛОВ Б. В.
СТЕМПКОВСКИЙ А. Л.
УСКОВ В. Л.
ЧЕРМОШЕНЦЕВ С. Ф.
ШИЛОВ В. В.

Редакция:

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

Информация о журнале доступна по сети Internet по адресу <http://www.informika.ru/text/magaz/it/> или <http://novtex.ru/IT>.

Журнал включен в систему Российского индекса научного цитирования.

Журнал входит в Перечень научных журналов, в которых по рекомендации ВАК РФ должны быть опубликованы научные результаты диссертаций на соискание ученой степени доктора и кандидата наук.

УДК 620.9:004.056

В. А. Васенин, д-р физ.-мат. наук, проф.,
Институт проблем информационной
безопасности, Московский государственный
университет им. М. В. Ломоносова,
e-mail:vasenin@msu.ru

Критическая энергетическая инфраструктура: средства противодействия кибертерроризму

Настоящая публикация продолжит тему, затронутую в статье "Критическая энергетическая инфраструктура: кибертеррористическая угроза". Она посвящена вопросам защиты объектов критически важной энергетической инфраструктуры от деструктивных компьютерных воздействий с террористическими целями. Основное внимание в данной статье уделено методам и средствам защиты таких объектов на основе комплексного подхода к обеспечению безопасности информационных технологий.

Ключевые слова: энергетическая инфраструктура, топливно- и гидроэнергетический комплекс, атомная энергетика, деструктивное воздействие, критически важный объект, методы и средства защиты, информационная безопасность, комплексный подход.

Настоящая публикация продолжит тему, затронутую в предыдущей статье*, и посвящена вопросам защиты объектов критически важной энергетической инфраструктуры от деструктивных компьютерных воздействий с террористическими целями. Основное внимание в данной статье уделено методам защиты таких объектов на основе комплексного подхода к обеспечению безопасности информационных технологий.

Для того, чтобы сделать изложение более понятным, напомним базовые определения, которые используются далее.

Под *энергетической инфраструктурой* будем понимать совокупность составляющих ее элементов — объектов, включая здания и сооружения, технические средства и технологии, персонал, обслуживающий их с целью получения энергоресурсов.

Под *энергоресурсами* будем понимать продукты первичной переработки сырья, полученного в ре-

зультате эксплуатации нефтяных и газовых месторождений, а также генерации электрической и тепловой энергии на гидроэлектростанциях, тепловых и атомных станциях. К числу задач, определенных назначением отдельных секторов энергетической инфраструктуры национального хозяйственного комплекса любой страны, относится организация производственно-хозяйственной и финансово-экономической деятельности, направленных на добычу (производство) и переработку, хранение и транспортировку ресурсов:

- нефтегазовых (топливно-энергетических);
- электроэнергетики;
- атомной энергетики.

В некоторых странах, которые именуются добывающими, может решаться весь комплекс перечисленных задач, включая добычу ресурсов и их первичную переработку. В других странах решается только часть таких задач, например, переработка, хранение и транспортировка. В разных странах эти задачи близки по целям, ради достижения которых энергетические инфраструктуры создаются. Они похожи по способам их решения и, как следствие, по проблемам, которые при этом возникают. Однако в каждой из стран, имеющей такие инфраструктуры, целевые установки на организацию их деятельности в той или иной форме диктует, а условия и способы их достижения, как правило, контролирует государство в лице уполномоченных на эту деятельность структур и отдельных лиц. Причина подобного внимания к энергетической инфраструктуре в том, что в отличие от целого ряда других инфраструктур национального хозяйственного комплекса, эта инфраструктура является критически важной для государства, а основные (ключевые) системообразующие ее объекты относятся к категории критически важных.

К числу критически важных [1, 2] будем относить следующие инфраструктуры и их элементы, которые в той или иной мере использованы в данном сообщении.

Критически важная инфраструктура — набор взаимодействующих сегментов и составляющих их (входящих в их состав) объектов национального хозяйственного комплекса, поддерживающих сферы жизнедеятельности, частичная деградация или полная потеря функциональности которых способна прямо и в течение относительно короткого интервала времени влиять на состояние тех или иных составляющих национальной безопас-

* **Васенин В. А.** Критическая энергетическая инфраструктура: кибертеррористическая угроза // Информационные технологии. 2009. № 9. С. 2—8.

ности, приводить к чрезвычайным ситуациям определенного уровня и масштаба.

Критически важная информационная инфраструктура — совокупность инфраструктурных элементов, включая информационно-вычислительные и коммуникационные ресурсы, обеспечивающие управление объектами критически важных инфраструктур.

Критически важные составляющие национальной информационно-телекоммуникационной инфраструктуры — такие элементы (сегменты, объекты) национальной информационно-телекоммуникационной инфраструктуры, которые обеспечивают управление отдельными частями (компонентами) критически важных для государства инфраструктур.

Критически важный объект — объект критически важной инфраструктуры, который в случае частичной деградации или полной потери функциональности способен прямо и в течение относительно короткого интервала времени влиять на состояние национальной безопасности, тех или иных ее составляющих, приводить к чрезвычайным ситуациям определенного масштаба и уровня.

Критически важный сегмент — совокупность критически важных объектов, объединенных по одному или нескольким квалификационным признакам, таким, например, как использование в одном технологическом процессе, ведомственная принадлежность, требования к безопасности подерживаемой ими информации, и другие.

Отметим, что в предыдущей статье рассматривались основные принципы, архитектурные и технологические особенности организации деловых процессов, обеспечивающих решение задач энергетической инфраструктуры при кибертеррористических угрозах и уязвимостях.

Подходы к организации противодействия

Рассматривая подходы к организации противодействия кибератакам на объекты критически важного энергетического комплекса, будем принимать во внимание отмеченную выше принадлежность объектов всех трех его секторов к одному классу с позиций обеспечения их информационной безопасности. Следует также учитывать то обстоятельство, что многие из характеризующих эти объекты атрибутов, включая угрозы, способы их реализации и средства противодействия им, являются общими для большинства критически важных объектов других инфраструктур. Принимая во внимание эти соображения, а также для краткости изложения будем далее опускать упоминание принадлежности критически важного объекта к энергетической инфраструктуре в тех местах по тексту, где этого не требуют дополнительные обстоятельства.

Как уже отмечалось ранее, для организации эффективного противодействия проявлениям кибертеррористических угроз объектам критически важных инфраструктур необходим подробный и всесторонний анализ этой предметной области, ее систематизация и формальное описание [1]. Такое описание должно содержать:

- способы идентификации, систематизации и категоризации объектов защиты как элементов той или иной критической инфраструктуры, имеющей свои особенности;
- способы определения уровня защищенности сложно организованного критически важного объекта и методы оценки рисков реализации против них деструктивных информационных воздействий;
- подходы (механизмы и модели, средства и системы) к организации противодействия проявлениям кибертеррористических угроз на всех уровнях комплексного подхода к обеспечению информационной безопасности критически важных объектов.

Не затрагивая способы идентификации, систематизации и категоризации, вопросы определения уровня защищенности и управления рисками для критически важных объектов, которые не являются предметом настоящей публикации, остановимся на организации комплексного подхода к обеспечению их информационной безопасности. Такой подход предполагает сочетание скоординированных и дополняющих друг друга мер и мероприятий, механизмов, моделей и инструментальных средств на нескольких уровнях обеспечения информационной безопасности подконтрольного объекта. К их числу относятся нормативно-правовой (законодательный), административно-организационный, процедурный и программно-технический уровни. Рассмотрим содержание каждого из этих уровней.

Нормативно-правовой уровень обеспечения информационной безопасности

Нормативно-правовой уровень [3] обеспечения информационной безопасности любого, в том числе объекта критически важной (включая энергетическую) инфраструктуры, основывается на использовании:

- правовых норм действующего в стране законодательства;
- положений документов, развивающих правовые нормы и регламентирующих деятельность организаций различных форм собственности, государственных органов, уполномоченных на эту деятельность;
- стандартов и рекомендаций, как отечественных, так и зарубежных.

Законодательные акты и нормативно-регулирующие документы создают основу для действий

на всех других уровнях обеспечения информационной безопасности.

Анализ действующих в стране правовых норм [1] показывает, что в настоящее время законодательное поле не соответствует (не удовлетворяет) современным потребностям в защите даже менее сложных в архитектурно-технологическом плане и менее ответственных с позиций категории их значимости для государства объектов. Не останавливаясь на деталях этого несоответствия, которые не являются предметом данной публикации, отметим, что специфика объектов критически важных инфраструктур в законодательстве отражена только с самых общих позиций. Она заключается только в "грубом" (в самом общем) делении информации, и, соответственно, методов ее защиты по категориям доступа к ней на общедоступную и ограниченного доступа (конфиденциальную), включая персональные данные, служебную информацию и государственную тайну. На проблемную область, которая определяется подходами к защите объектов критически важных инфраструктур от деструктивных информационных воздействий, такое деление отображается (влияет) очень опосредованно. Как объективное следствие такого положения дел, отсутствие документов, регламентирующих деятельность по обеспечению информационной безопасности критически важных объектов.

В связи с изложенным следует отметить тот факт, что в последние годы, в том числе при инициативном участии Совета Безопасности Российской Федерации, внимание к вопросам обеспечения информационной безопасности объектов критически важных инфраструктур и к кибертеррористической угрозе значительно повысилось. Оно находит отражение в обсуждениях на разных уровнях государственной власти. Эти вопросы серьезно изучаются и обсуждаются в среде ученых на различных форумах. В качестве результатов такого обсуждения можно привести материалы (в том числе переведенные на английский язык) трех международных научных конференций по проблемам безопасности и противодействию терроризму, которые в 2005—2007 годах [4—6] проходили в Москве, в МГУ имени М. В. Ломоносова. В них принимали участие представители США, Германии, Великобритании, Китая и ряда других стран. В конце октября 2008 года состоялась четвертая конференция из этого цикла. По инициативе законодателей в Государственную думу внесен и находится в стадии обсуждения проект Федерального закона "Об особенностях обеспечения информационной безопасности критически важных объектов информационно-телекоммуникационной инфраструктуры". Заметим, что рассматриваемые выше автоматизированные системы управления технологическими процессами,

производственно-хозяйственной и финансово-экономической деятельностью в сфере добычи (получения), переработки, хранения и транспортировки энергоресурсов представляют собой именно такие объекты.

Принимая во внимание изложенные выше соображения относительно стандартов и нормативно-регламентирующей базы безопасности информационных технологий с позиций их применения к автоматизированным информационным системам в составе критически важных объектов, необходимо отметить высокую важность решения вытекающих отсюда задач. Следует подчеркнуть, что перечисленные недостатки нормативно-правового обеспечения информационной безопасности критически важных объектов в полной мере относятся к объектам национальной энергетической инфраструктуры, входящим с позиций обеспечения информационной безопасности в один из классов критически важных объектов. Отмеченное обстоятельство с полным правом можно отнести к еще одному источнику уязвимостей этой инфраструктуры.

Административный уровень обеспечения информационной безопасности

Действия на административном уровне обеспечения информационной безопасности объекта критически важной инфраструктуры должны быть направлены:

- на формирование политики безопасного использования его ресурсов;
- на формулировку требований к среде окружения и описание профиля защиты объекта;
- на разработку (с позиций информационной безопасности) спецификаций средств, используемых в подконтрольном объекте.

Принимая во внимание отмеченное ранее отсутствие надлежащей законодательной базы и нормативных документов, регламентирующих деятельность по обеспечению информационной безопасности объектов критически важных инфраструктур, можно констатировать, что систематизированных требований и даже рекомендаций к организации деятельности на этом уровне на настоящее время не существует. Аналогичное положение в силу тех же причин, судя по материалам анализа нормативно-правовых документов, складывается и в других странах. Это обстоятельство безусловно является отрицательным фактором, снижающим эффективность работ, направленных на создание автоматизированных информационных систем в составе объектов критически важных инфраструктур. В полной мере эти недостатки относятся и к национальной энергетической инфраструктуре.

В связи с изложенным, важным направлением является разработка рекомендаций к организации

действий, на административном уровне, обеспечения информационной безопасности объектов критически важных инфраструктур. В развитие этих работ необходимы рекомендации по применению отдельным инфраструктурам национального хозяйственного комплекса, в том числе энергетической, которые учитывали бы их специфику.

Процедурный уровень обеспечения информационной безопасности

Меры процедурного уровня ориентированы на защиту критически важных объектов от деструктивных информационных воздействий через совокупность действий, предпринимаемых администрирующим объект персоналом и пользователями его активов. Такие меры должны быть направлены:

- на управление персоналом;
- на обеспечение физической защиты;
- на поддержание работоспособности;
- на реагирование в случае нарушения режима безопасности;
- на планирование восстановительных работ.

Механизмы поддержки взаимодействия отдельных информационно-вычислительных систем в составе сложных критически важных объектов должны быть четко документированы в целях обеспечения единого подхода персонала, администрирующего такие системы, к вопросам обеспечения их корректной и надежной работы. Документированные регламенты данного уровня должны быть направлены на обеспечение согласованной работы персонала в ходе разработки, наладки, сопровождения и тестирования систем в интерактивном режиме.

Действия на процедурном уровне обеспечения информационной безопасности регламентированы положениями ранее упомянутого ГОСТ Р ИСО/МЭК 17799—2005 и рядом других специальных требований и рекомендаций по технической защите конфиденциальной информации. Однако в случае сложно организованных автоматизированных информационных систем в составе критически важных объектов, к которым относятся объекты национальной энергетической инфраструктуры, реализация этих рекомендаций сталкивается с рядом объективных трудностей. Например, обеспечение высокого уровня их отказоустойчивости путем перманентного мониторинга состояния отдельных элементов таких объектов на предмет работоспособности без применения специальных систем с высоким уровнем автономности не представляется возможным. Еще большие сложности связаны с "on-line" анализом "сбойных" ситуаций и выработкой мер оперативного реагирования на них. Эти и ряд других трудностей реализации мер на процедурном уровне обеспечения информационной безопасности объектов критически важных инфраструктур настоятельно требуют раз-

работки специальных систем с высоким уровнем автоматизации, обладающих должным интеллектом и способных к автономным действиям в случае появления "сбойных" ситуаций.

Программно-технический уровень обеспечения информационной безопасности

Программно-технический уровень обеспечения информационной безопасности в традиционном понимании представляет собой комплекс средств защиты информационных активов и других ресурсов подконтрольного объекта. Такой комплекс включает математическое и программное обеспечение, средства вычислительной техники и коммуникационное оборудование. Современный рынок информационных технологий располагает богатым набором технических средств защиты информации, рассчитанных, в первую очередь, на обеспечение конфиденциальности и целостности данных. В меньшей степени представлены средства защиты от атак на отказ в обслуживании, еще меньше — от распределенных DDoS-атак. Вместе с тем, как уже отмечалось ранее, архитектурно и технологически сложно организованные автоматизированные информационные системы в составе объектов критически важных инфраструктур имеют высокие категории значимости и, соответственно, повышенные требования к обеспечению безопасности их ресурсов.

Анализ представленных на мировом IT-рынке средств вычислительной техники и автоматизированных систем с позиций подходов и критериев, изложенных в ГОСТ Р ИСО/МЭК 15408—2002 [7—9], и с позиций Руководящих документов ФСТЭК [10, 11] и других специальных требований и рекомендаций, показывает, что лишь незначительная их часть может отвечать таким требованиям. Причина в том, что эти средства в силу законов рынка, как правило, рассчитаны на массового потребителя. Они имеют широкую функциональность и, как следствие, несут в себе дополнительные источники уязвимости и не предоставляются пользователю с открытым исходным кодом. Это обстоятельство является отражением того объективного факта, что "нетиповые" критически важные объекты требуют соответствующих "нетиповых" подходов к обеспечению их безопасности. Реализация этих подходов по сути своей не является задачей бизнеса и IT-рынка. Как следствие, к числу важнейших для государства относится задача разработки таких средств, включая математическое, алгоритмическое и программное обеспечение, средства вычислений и телекоммуникаций, которые учитывали бы специфику критически важных объектов как с позиций их потребительских качеств, так и в плане повышенных требований к обеспечению безопасности. В разработке подобных средств, которая должна проводиться по

заказу уполномоченных на это государственных органов, активное участие должны принимать бизнес-структуры. К их числу относятся компании и корпорации, обслуживающие энергетический комплекс страны. Выступая в качестве потенциальных потребителей средств и систем обеспечения информационной безопасности подведомственных им объектов и инвестируя эти работы, такие бизнес-структуры должны оказывать влияние на сроки выполнения и результаты работ.

К числу базовых средств на рассматриваемом направлении можно отнести операционные системы и механизмы их расширения (организации взаимодействия) на ключевые сервисы обеспечения информационной безопасности – идентификацию/аутентификацию, логическое разграничение доступа к ресурсам, шифрование данных, аудит состояния системы на предмет наличия деструктивных воздействий и оперативного реагирования, а также на ряд других. Рассмотрим эти средства и направления их совершенствования более подробно.

Дистрибутивы операционных систем. Операционная система является одним из базовых элементов, обеспечивающих функционирование современных компьютерных комплексов. Она предназначена для управления аппаратными ресурсами компьютера и организации интерфейса с пользователем, обеспечивая запуск и выполнение пользовательских программ, прикладных и системных сервисов. Операционная система включает в себя большинство базовых механизмов и служб, поддерживающих безопасность компьютерных комплексов. Учет повышенных требований к защищенности, предъявляемых к компьютерным системам управления критически важными объектами на ранних этапах их проектирования и реализации, позволяет эффективно использовать необходимые для их выполнения механизмы операционной системы. Надежность функционирования операционной системы и механизмов безопасности в ее составе во многом определяет уровень защищенности автоматизированных систем управления критически важными объектами.

Традиционные подходы к созданию операционных систем зачастую приводят к избыточности программной реализации. Это обстоятельство, в свою очередь, затрудняет проведение аудита автоматизированных систем, использующих такие операционные системы, на предмет выполнения требований по безопасности. Важным аспектом подобной избыточности являются излишне усложненные механизмы управления средствами обеспечения безопасности, например подсистемой логического разграничения доступа. В качестве одного из эффективных подходов к разрешению этого вопроса может быть предложена многопрофильная архитектура набора дистрибутивов

операционных систем. Такой подход позволяет контролировать избыточность, не приводя к появлению большого числа независимо разрабатываемых программных комплексов. Каждый дистрибутив из разрабатываемого набора должен быть ориентирован на свой профиль, подразумевающий поддержку специализированных сервисов, включаемых в его состав.

Важную роль в операционных системах с позиций обеспечения безопасности играют такие базовые сервисы безопасности, как идентификация/аутентификация и логическое разграничение доступа. Реализация первого из перечисленных сервисов может быть усовершенствована по сравнению с традиционными подходами за счет внедрения строгих протоколов взаимодействия участников процессов аутентификации и реализации алгоритмов, использующих стойкие аутентификационные данные. Эффективное функционирование сервиса логического разграничения доступа в контексте обеспечения безопасности критически важных объектов требует использования современных моделей разграничения доступа, для которых возможна автоматизированная проверка их корректности относительно задаваемого набора правил (требований по безопасности). Такие усовершенствования позволяют создать средства надежной изоляции отдельных программных компонентов в операционной системе и программно-аппаратных комплексов, в составе которых она используется.

В качестве перспективного подхода к созданию операционных систем для объектов критически важных инфраструктур предлагается разработка Unix-подобных операционных систем на основе ядра Linux и программного обеспечения с открытым исходным кодом [12]. Поэтапная разработка таких операционных систем в обязательном порядке должна включать проведение всестороннего аудита программных средств на предмет наличия программных ошибок и уязвимостей с использованием как статического анализа кода (например, на основе методов формальной верификации), так и тестирования программных средств в процессе их функционирования. Проведение таких мероприятий в процессе разработки, а также создание инструментальных средств, позволяющих автоматизировать проверку выполнения требований по безопасности, дает возможность существенно повысить уровень надежности сложно организованных компьютерных комплексов.

Мониторинг состояния работоспособности. Одной из определяющих в комплексе мер по обеспечению информационной безопасности компьютерных систем управления критически важными объектами является поддержка штатных, предусмотренных регламентом, режимов их функционирования. Аппаратно-программные компоненты

критически важных объектов, включая программное обеспечение, реализующие системные и общецелевые сервисы, а также коммуникационные функции, потенциально подвержены внешним и внутренним деструктивным воздействиям. Выявление параметров, описывающих состояние работоспособности базовых аппаратно-программных средств, которые поддерживают функционирование объекта в штатном (согласно принятому регламенту) режиме, постоянный мониторинг и анализ значений этих параметров позволяют своевременно выявлять аномальные ситуации, оперативно реагировать на них и добиваться стабильной, высокоэффективной работы как отдельных компонентов, так и всего критически важного объекта в целом.

Для решения указанных задач в состав системы управления критически важным объектом необходимо включать отдельную службу — подсистему мониторинга работоспособности ключевых элементов, поддерживающих работу объекта в штатном режиме [13]. Основное назначение такой службы состоит в автоматическом обнаружении неисправностей аппаратно-программных компонентов подлежащего защите объекта, в подготовке предложений по системе мер и, возможно, реализации действий, направленных на их локализацию и устранение.

Подсистема активного аудита. Одним из основных требований в рамках комплексного подхода к обеспечению безопасности автоматизированных систем управления критически важными объектами является эшелонированный характер размещения программно-технических средств защиты информации. Необходимость такого требования обусловлена целым рядом причин, в том числе:

- отсутствием адекватного современным потребностям набора механизмов обеспечения безопасности в операционных системах;
- наличием ошибок в реализации программного обеспечения и в администрировании системы;
- постоянным появлением новых ошибок и уязвимостей, неучтенных при формировании политики информационной безопасности и при оценке рисков угроз безопасности функционирования компьютерных систем.

Подобные причины обусловлены как человеческим фактором, так и объективными недостатками механизмов безопасности. Определяющими факторами являются сложный характер объектов, которые подлежат защите, и отсутствие необходимых для решения таких задач научно-методической, технологической базы и технических средств.

С учетом изложенного разумным дополнительным эшелоном защиты в сложившейся ситуации представляется система активного аудита. Отдельные элементы таких инструментальных комплексов в настоящее время принято именовать

системами обнаружения вторжений (IDS — Intrusion Detection System). Комплексы активного аудита предназначены:

- для раннего выявления аномального поведения компьютерных систем, вызванного действиями злоумышленников, ошибками легальных пользователей или целым рядом других возможных причин;
- для оперативного реагирования на нештатные ситуации и предотвращения нанесения значительного ущерба защищаемой компьютерной системе.

Следует отметить, что система активного аудита должна быть модульной, интегрированной и легко настраиваемой, что означает выполнение следующих требований:

- возможность оперативного добавления и удаления новых алгоритмов выявления аномального поведения компьютерной системы, позволяющая реагировать на появление новых видов угроз или уязвимостей;
- возможность сбора и централизованного анализа информации со всех компонентов распределенной системы, что позволит принимать корректные решения в случае распределенной атаки на защищаемую систему.

Дополнительным требованием является возможность принимать решения в режиме реального времени работы защищаемой автоматизированной системы, чтобы оперативно реагировать на возникновение нештатных ситуаций.

Средства анализа исходного кода программ на наличие уязвимостей. Одним из важнейших требований к подсистеме обеспечения информационной безопасности автоматизированных систем управления критическими важными объектами является наличие в ее составе средств анализа исходного кода программ на наличие уязвимостей. Повышенное внимание к этой проблеме обусловлено высокой вероятностью наличия в программном обеспечении компьютерных систем уязвимостей, причинами которых могут быть как ошибки разработчиков, так и намеренные действия некоторых из них. Использование таких уязвимостей злоумышленником может привести к нарушению штатного режима функционирования подконтрольного критически важного объекта, а также к нерегламентированному политикой безопасности повышению привилегий злоумышленника на каком-то из узлов системы, что является одним из этапов компьютерной атаки на объект.

Существующие в настоящее время подходы к решению данной задачи предусматривают, как правило, анализ исходного кода программ, проводимый вручную специально подготовленной группой экспертов. Как следствие, существует высокая вероятность возникновения ошибки при работе экспертов, которая с увеличением объема

анализируемого кода увеличивается. Указанное обстоятельство недопустимо при построении автоматизированных систем управления критически важными объектами, поэтому необходимы более совершенные подходы к решению рассмотренной задачи. Одним из таких подходов является использование средств автоматизированного анализа исходного кода программ в целях выявления в нем возможных уязвимостей [14]. При использовании данного подхода эксперты должны анализировать лишь отдельные участки кода, выявленные автоматическим "анализатором". Следует отметить, что в настоящее время эффективных и широко распространенных автоматизированных средств анализа исходного кода не существует. Тем не менее, существует ряд теоретических подходов к построению подобных средств, практическая реализация которых затруднена по ряду объективных причин. Одной из них является значительная сложность используемых языков программирования, таких как С или С++. Вместе с тем, следует отметить, что проведенные предварительные эксперименты показали возможность построения эффективных средств автоматического анализа программ на предмет наличия в них ряда типовых уязвимостей, характерных для операций с памятью. Такие средства прошли апробацию на некоторых, в том числе отдельных крупных, элементах дистрибутива операционной системы с открытым исходным кодом, включая ее ядро, базовые приложения и библиотеки для работы с исполняемым бинарным кодом, основные сетевые приложения (веб-сервер, веб-клиент), графическую подсистему.

Автоматизированная система тематического анализа информации. Важную роль в процессе информационной безопасности объектов критически важных инфраструктур играет осведомленность администрирующего их персонала о появлении новых угроз и уязвимостей, об атаках, реализованных на близкие по назначению и архитектурно-технологическому каркасу объекты, а также о способах противодействия им. Для поддержания в актуальном состоянии информации, имеющей отношение к обеспечению безопасности подконтрольных объектов, необходимо проведение перманентного поиска и тематического анализа больших объемов данных. В указанном контексте представляется целесообразным использовать автоматизированные системы для тематического анализа информации [15]. Их функцией является настраиваемый мониторинг выбранных пользователем информационных источников, аннотирование, визуализация результатов поиска, хранение текстовой информации в различных форматах, осуществление поиска, тематического анализа, классификации, фильтрации и ранжирования текстовой информации. Ав-

томатизация перечисленных процессов позволит эффективно использовать актуальную информацию в процессах разработки и модернизации программного обеспечения, при формировании правил политики информационной безопасности, в ходе проведения аудита аппаратно-программных комплексов на предмет выполнения ими требований по безопасности. Наличие такой автоматизированной системы в каждой из критически важных инфраструктур, учитывающей ее специфику и потребности, может существенно повлиять на уровень обеспечения информационной безопасности ее объектов.

Выводы

Результаты исследований, связанных с анализом защищенности критически важных объектов энергетической инфраструктуры от деструктивных информационных воздействий с террористическими целями, позволяют сделать следующие выводы.

1. Проблемная область, отражающая (включая) вопросы защиты объектов критически важных инфраструктур вообще, в том числе — инфраструктуры энергетической, на настоящее время недостаточно хорошо исследована и систематизирована, что не позволяет сделать надежные выводы о потенциальных угрозах, о способах их реализации и о подходах к выбору эффективных мер противодействия.

2. К числу объектов, потенциально уязвимых для кибертеррористической атаки на элементы критически важной энергетической инфраструктуры, подлежащих защите, в первую очередь, относятся:

- автоматизированные системы управления технологическими процессами на нижнем уровне их реализации и их компоненты (серверы, в первую очередь, серверы SCADA, автоматизированные рабочие места, микропроцессорные контроллеры, средства телемеханики);
- информационно-телекоммуникационные сети, поддерживающие автоматизированные системы управления технологическими процессами;
- объекты информатизации процессов добычи (получения), переработки (преобразования) и транспортировки энергоресурсов (объекты, поддерживающие компрессорные системы, перекачку топлива, электроснабжение и подобные им).

3. Угрозы кибертеррористического воздействия на объекты критически важной энергетической инфраструктуры могут преследовать цели нарушения:

- конфиденциальности информации АСУ ТП;
- целостности информации АСУ ТП (несанкционированная модификация информации);
- работоспособности (доступности) информационных ресурсов АСУ ТП, успешная реализация

которых по отдельности или в каком-либо сочетании может привести к чрезвычайным ситуациям, к другим потерям, способным повлиять на состояние национальной безопасности.

4. Анализ способов реализации деструктивных информационных воздействий с террористическими целями на производственные объекты энергетической инфраструктуры позволяет отметить следующие их особенности.

- К числу вероятных сценариев кибертеррористической атаки на объекты энергетической инфраструктуры следует отнести те, которые обеспечивают не только временную или полную потерю их функциональности, но и, как следствие (вторичный эффект), создают широкомасштабную чрезвычайную ситуацию с высоким уровнем потерь (материальных, человеческих и других) и/или угрозу национальной безопасности.
- Реализация таких сценариев с высокой вероятностью будет осуществляться группой координирующих свои действия агентов из разных точек сетевой среды, которые обеспечивают доступ к ресурсам объекта атаки, в том числе, из точек, расположенных вне страны, объекты которой подвергаются атаке.
- Наиболее значимыми с позиции размера потенциального ущерба (социального, материального, политического) являются:
 - ◆ сложно предотвращаемые в оперативном порядке распределенные атаки на отказ в обслуживании (DDoS- Distributed Denial of Service);
 - ◆ комплексные атаки, результатом которых является получение контроля над управляемым производственным объектом или важными технологическими процессами, обеспечивающими его функционирование по назначению.

5. При разработке системы мер противодействия кибертеррористическим атакам на объекты энергетической инфраструктуры необходимо:

- проведение системных научных исследований и прикладных работ на данном направлении под эгидой государства при активном участии бизнес-структур, работающих в данной области;
- тесное взаимодействие государственных органов, организаций и бизнес-структур на национальном уровне и консолидация усилий стран на уровне международном;
- комплексный подход к обеспечению информационной безопасности подконтрольных объектов, предполагающий скоординированную систему мер и мероприятий, моделей, механизмов и инструментальных средств на нормативно-правовом, административном, операционном и программно-техническом уровнях его реализации.

Автор выражает благодарность Шапченко Кириллу Александровичу, аспиранту Механико-математического факультета МГУ им. М. В. Ломоно-

сова, за помощь в подготовке материалов к настоящей публикации.

Список литературы

1. **Критически** важные объекты и кибертерроризм. Ч. 1. Системный подход к организации и противодействию // О. О. Андреев и др. Под ред. В. А. Васенина. М.: МЦНМО, 2008. 398 с.
2. **Васенин В. А.** Научные проблемы противодействия кибертерроризму // Проблемы безопасности и противодействия терроризму. Материалы международной научной конференции. Интеллектуальный центр МГУ, 2—3 ноября 2005 г. — М.: МЦНМО, 2006. С. 49—63.
3. **Стрельцов А. А.** Цель, структура и методы правового обеспечения информационной безопасности Российской Федерации // Сб. "Научные и методологические проблемы информационной безопасности". Под ред. В. П. Шерстюка. М.: Издво МГУ, 2004. С. 67—83.
4. **Проблемы** безопасности и противодействия терроризму. Материалы международной научной конференции. Интеллектуальный центр МГУ, 2—3 ноября 2005 г. М.: МЦНМО, 2006. 480 с.
5. **Проблемы** безопасности и противодействия терроризму. Материалы второй международной научной конференции. Интеллектуальный центр МГУ, 25—26 октября 2006 г. — М.: МЦНМО, 2007. — 664 с.
6. **Проблемы** безопасности и противодействия терроризму. Материалы третьей международной научной конференции. Интеллектуальный центр МГУ, 25—27 октября 2007 г. — М.: МЦНМО, 2008. — 501 с.
7. **ГОСТ Р ИСО/МЭК 15408-1—2002.** Информационная технология. Методы и средства обеспечения безопасности. Критерии оценки безопасности информационных технологий. Часть 1. Введение и общая модель. — М.: ИПК Издательство стандартов, 2002.
8. **ГОСТ Р ИСО/МЭК 15408-2—2002.** Информационная технология. Методы и средства обеспечения безопасности. Критерии оценки безопасности информационных технологий. Часть 2. Функциональные требования безопасности. — М.: ИПК Издательство стандартов, 2002.
9. **ГОСТ Р ИСО/МЭК 15408-3—2002.** Информационная технология. Методы и средства обеспечения безопасности. Критерии оценки безопасности информационных технологий. Часть 3. Требования доверия к безопасности. — М.: ИПК Издательство стандартов, 2002.
10. **Средства** вычислительной техники. Защита от несанкционированного доступа к информации. Показатели защищенности от несанкционированного доступа к информации. Руководящий документ ФСТЭК от 30 марта 1992 года / ФСТЭК. — 1992.
11. **Автоматизированные** системы. Защита от несанкционированного доступа к информации. Классификация автоматизированных систем и требования по защите информации. Руководящий документ ФСТЭК от 30 марта 1992 года / ФСТЭК. — 1992.
12. **Васенин В. А., Шапченко К. А., Андреев О. О.** Математические модели и механизмы логического разграничения доступа в операционной системе Linux: текущее состояние и перспективы развития // Материалы второй международной научной конференции по проблемам безопасности и противодействия терроризму. Московский государственный университет имени М. В. Ломоносова. М.: МЦНМО, 2007. С. 159—172.
13. **Коршунов А. А., Дзыба М. С., Астапов И. С.** Подсистема мониторинга состояния функциональности элементов компьютерных систем управления критически важными объектами // Критически важные объекты и кибертерроризм. Часть 2. Аспекты программной реализации / О. О. Андреев и др. Под ред. В. А. Васенина. — М.: МЦНМО, 2008. С. 329—375.
14. **Пучков Ф. М.** Средства аудита программного обеспечения на предмет обнаружения уязвимостей. // Критически важные объекты и кибертерроризм. Часть 2. Аспекты программной реализации. / О. О. Андреев и др. Под ред. В. А. Васенина. — М.: МЦНМО, 2008. С. 375—454.
15. **Козицын А. С., Афонин С. А., Титов А. С.** Автоматизированная система для тематического анализа информации // Критически важные объекты и кибертерроризм. Часть 2. Аспекты программной реализации. / О. О. Андреев и др. Под ред. В. А. Васенина. — М.: МЦНМО, 2008. С. 497—578.

А. В. Ремизов, аспирант,
М. В. Филиппов, канд. техн. наук, доц.,
Н. В. Чичварин, канд. техн. наук, доц.,
 МВТУ им. Н. Э. Баумана

Методы защиты информации в звуковых файлах

Рассматриваются методы сокрытия информации в звуковых файлах формата MP3. Показана структура MP3-кодирования. Предлагаются новые методы сокрытия, как для сокрытия больших объемов информации, так и для сокрытия с высокой степенью устойчивости к обнаружению. Приводятся сравнительные результаты работы существующих и предлагаемых методов с точки зрения емкости, быстродействия и устойчивости к обнаружению.

Ключевые слова: стеганография, стеганографический анализ, стеганографический ключ, сообщение, сжатие с потерями, битрейт, квантование, энтропийное сжатие, Хаффман-кодирование, СЛАУ.

Анализ текущих публикаций, посвященных тематике информационной безопасности [1, 2], показывает, что актуальность проблемы защиты информации, передаваемой по различным каналам связи и хранимой на цифровых носителях, непрерывно возрастает. Очевидность этого утверждения подтверждается еще и тем, что действия и квалификация злоумышленников совершенствуются по мере совершенствования методов и средств защиты информации.

Известно, что основными направлениями развития программных и программно-аппаратных средств обеспечения информационной безопасности в настоящее время являются:

- разработка криптографических алгоритмов;
- разработка средств, реализующих методы стеганографии.

Сопоставление этих направлений показывает, что возможностей скрыть от несанкционированного пользователя сам факт наличия информации гораздо больше при реализации методов стеганографии.

В отличие от методов криптографии, где злоумышленник с высокой вероятностью может определить, является ли передаваемое сообщение зашифрованным текстом, методы стеганографии позволяют встраивать секретные сообщения в безобидные сообщения так, что обнаружить факт существования сокрытого сообщения требует значительных затрат.

В данной статье рассматриваются стеганографические методы сокрытия информации в звуко-

вых файлах. Принимается во внимание то, что стандартным форматом звуковых файлов в текущий момент является формат MP3.

1. Анализ формата кодирования MP3

MP3 (MPEG 1 Layer 3) — формат кодирования звуковых файлов, входящий в стандарт кодирования видеoinформации MPEG 1 [3]. Принципиальной особенностью формата является сжатие с потерями, т. е. после упаковки и распаковки звукового файла с помощью MP3 результат не является побитной копией оригинала. Напротив, при кодировании несущественные компоненты целенаправленно исключаются из упаковываемого сигнала. Многолетняя практика показывает, что при сохранении заданного наперед качества MP3 позволяет сжать звуковые данные в 10 и более раз.

Рассматриваемый в настоящей работе метод учитывает то, что MP3 основан на нелинейных свойствах слухового анализатора человека. Имеется в виду использование эффекта маскирования слабого сигнала одного диапазона частот более мощным сигналом соседнего диапазона, когда он имеет место, или мощным сигналом предыдущего фрейма, вызывающего временное понижение чувствительности уха к сигналу текущего фрейма (удаляются второстепенные звуки, которые не слышатся человеческим ухом из-за наличия в данный или предыдущий момент другого, более громкого звука). Учитывается также неспособность большинства людей различать сигналы, по мощности лежащие ниже порогового уровня, различного для разных частотных диапазонов.

Известно, что данный процесс называется адаптивным кодированием и позволяет экономить на наименее значимых с точки зрения восприятия человеком деталях звучания. Степень сжатия (следовательно и качество) определяется задаваемой при кодировании шириной потока данных — *битрейтом*.

2. Анализ стеганографических методов

На основании обзора существующих публикаций можно сделать вывод, что большая часть существующих стеганографических методов, реализующих сокрытие звуковой информации, встраивает данные в несжатый звуковой поток непосредственно перед MP3-кодированием. К ним относятся следующие методы:

- *широкополосное кодирование*, когда в сигнал добавляется модулированный сообщением шум с амплитудой чуть выше предела маскирования;

- *фазовое кодирование*, основанное на том, что сигнал разбивается на участки; значения фаз на первом участке используются для кодирования скрываемого сообщения, значения фаз остальных участков меняются с тем, чтобы разность фаз между участками осталась неизменной;
- *замена шума*, когда знаки коэффициентов косинусного преобразования в "шумных" частотных полосах модулируются скрываемым сообщением.

Анализ показывает, что общим недостатком данных методов является значительное искажение сигнала-контейнера, снижение пропускной способности канала (около 40 бит/с) и значительная вычислительная сложность. Кроме того, внесенные искажения с высокой вероятностью обнаруживаются специализированными методами стеганографического анализа.

Из методов, непосредственно работающих с форматом MP3, наиболее тщательно разработанным является *mp3stego*[3]. Данный метод скрывает биты сообщения в служебных полях MP3 в ходе MP3-кодирования. Недостатками метода также является то, что его реализация приводит к снижению пропускной способности канала и, как будет показано далее, слабо защищена от стеганографического анализа, поскольку изменения служебных полей не соответствуют стандартной логике MP3-кодирования.

3. Описание предлагаемых стеганографических методов

Как известно, требования к скорости работы, пропускной способности и надежности зависят от конкретного сценария использования стеганографического алгоритма. Исходя из этого, предлагается два различных метода сокрытия, которые, как представляется, могут быть использованы в широком перечне сценариев.

Сокрытие информации между MP3-фреймами. При разработке метода учтено следующее:

1. Поток данных MP3 состоит из кадров — участков данных, кодирующих 26 мс звукового сигнала.
2. Каждый кадр содержит 4-байтный заголовок, содержащий синхронизирующее слово и служебную информацию.
3. Служебная информация содержит, в частности, версию формата, признак стереорежима и битрейт. Эти значения однозначно определяют размер кадра.

Учтено также, что число байтов, необходимых для хранения сжатого звукового сообщения с заданным качеством, зависит от характеристик конкретного звукового сигнала. Таким образом, на сложные для сжатия участки сигнала с большим числом деталей требуется больше байтов, чем на простые. Исходя из этого MP3 предусматривает

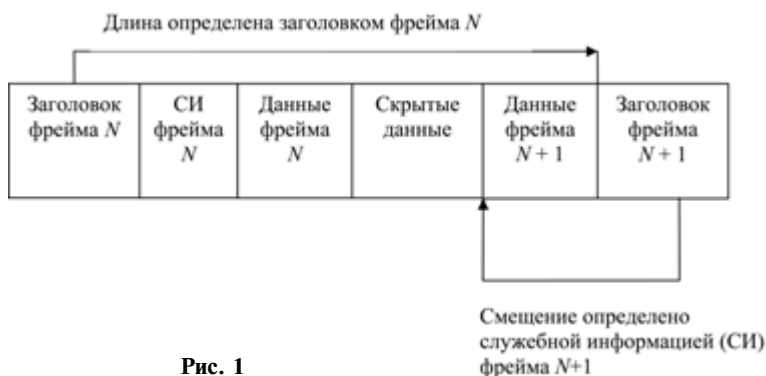


Рис. 1

возможность хранения упакованных данных кадра в оставшихся незанятыми байтах предыдущего кадра. Чтобы правильно позиционироваться, в заголовке кадра указывается смещение на начало данных. Признака конца данных как такового нет, декодирование останавливается при завершении распаковки требуемого числа коэффициентов косинусного преобразования.

Любая информация между концом данных одного кадра и началом данных следующего кадра игнорируется декодером, соответственно, именно там можно разместить скрываемые данные (рис. 1). Для этого увеличивается размер кадра путем увеличения его битрейта на одну ступень (например, с 128 кбит/с до 160 кбит/с).

Данный метод не вносит искажений в звуковой сигнал-контейнер, прост, быстр, имеет значительную пропускную способность (порядка 20 % от объема контейнера). Однако зная алгоритм сокрытия, легко обнаружить наличие скрытой информации исходя из того, что в стандартном MP3-потоке нет "лишних" байтов. Таким образом, метод предлагается использовать в тех случаях, когда требуется быстрая передача значительного объема информации и не ожидается серьезного стеганографического анализа.

Сокрытие в MP3-коэффициентах. При разработке метода учтено, что MP3-кодирование построено по классической схеме сжатия с потерями, состоящей из трех шагов:

- преобразование сигнала с тем, чтобы большая часть информации о сигнале была сосредоточена в небольшом числе коэффициентов. В MP3 использовано модифицированное дискретное косинусное преобразование;
- квантование, т. е. деление полученных коэффициентов на определенные значения с последующим округлением результата до целого числа. Именно на этом этапе происходит потеря информации. Значения делителей при квантовании в MP3 вычисляются исходя из заданного битрейта и психоакустической модели, определяющей максимально допустимый уровень шума;

- энтропийное сжатие без потерь. В MP3 используется сжатие с помощью статических кодов Хаффмана.

Полученные на втором этапе целочисленные коэффициенты могут быть использованы для сокрытия данных, например, размещая скрываемый бит в младшем бите коэффициента. Однако это вносит значительный шум в выходной сигнал и, как следствие, с высокой вероятностью может быть обнаружено существующими алгоритмами стеганоанализа.

Для минимизации вносимого шума предлагается использовать для сокрытия путем изменения младшего бита лишь те коэффициенты, у которых дробная часть близка к 0,5, — их абсолютные погрешности при округлении в большую и меньшую сторону близки. В этом случае проблемой является декодирование данных на стороне получателя, так как декодер получает лишь целочисленные значения и не имеет никакой информации о том, в каких коэффициентах размещены скрытые данные.

Данная проблема в работе [5] сводится к решению систем линейных алгебраических уравнений (СЛАУ) $\mathbf{D}\mathbf{b} = \mathbf{m}$, где \mathbf{D} — секретная двоичная матрица размером $m \times n$, общая для отправителя и получателя; \mathbf{b} — вектор размера n младших битов всех коэффициентов; \mathbf{m} — вектор размером m битов сообщения. Не подлежащие модификации коэффициенты задают значения некоторых элементов \mathbf{b} , таким образом, имеем систему из m уравнений с k переменными, где k — число подлежащих модификации коэффициентов.

Матрица \mathbf{D} заполняется с помощью генератора псевдослучайных чисел, инициализируемого стеганографическим ключом. Показано, что максимально возможный размер передаваемого сообщения, при котором система имеет решение, стремится к k при увеличении k .

Структура энтропийного кодирования в MP3 накладывает дополнительные ограничения на допустимые для модификации коэффициенты. Коэффициенты в рамках одного блока разбиты на три группы, каждая из которых упакована с помощью отдельной Хаффман-таблицы. Первые две группы упакованы парами следующим образом: в потоке записан код из таблицы, соответствующий паре, после чего записаны знаковые биты каждого отсчета. Числа, больше максимального, для таблицы кодируются как максимально возможный отсчет плюс определяемое таблицей число бит, записанных после кода, которые содержат прибавляемое к значению число.

Последний блок коэффициентов закодирован четверками с возможными значениями коэффициентов — 1, 0, 1. Длина блоков содержится в служебной информации кадра, что позволяет вычислить количество нулевых значений отсчетов в конце кадра.

Внесение изменений в коэффициенты последнего блока, т. е. замена -1 и 0 , 0 и 1 приведет к изменению размера сжатого коэффициента (вследствие появления или исчезновения знакового бита), что, в свою очередь, может привести к нарушению структуры MP3-файла (информация окажется больше или меньше отведенного под нее места в кадре).

Из коэффициентов первых двух блоков модификации подлежит не более одного коэффициента в паре, при этом необходимо, чтобы размер кода модифицированной пары совпадал с размером исходного кода. Как правило, это условие выполняется для кодов больших чисел. Процент доступных коэффициентов резко падает с уменьшением битрейта, так как большее количество коэффициентов кодируются нулями или близкими к нулю значениями.

Данные ограничения приводят к отсеиванию около 85 % подходящих коэффициентов.

Предельно допустимое отклонение дробной части коэффициента от 0,5 задается исходя из размера сообщения, которое надо скрыть в данном контейнере, с тем, чтобы число доступных коэффициентов было максимально близким к размеру сообщения. На рис. 2 приведены зависимости объема скрываемой информации от предельно допустимого отклонения от 0,5 для различных значений битрейта.

Для достижения допустимого уровня быстродействия метода MP3-поток разбивается на группы кадров одинакового размера. Число кадров в группах выбирается так, чтобы число модифицируемых коэффициентов в них было приблизительно одинаковым. В каждой группе СЛАУ сокрытия решается независимо. Таким образом, несмотря на то, что время работы метода Гаусса в каждой группе пропорционально кубу от числа уравнений, общее время работы растет линейно с увеличением числа используемых коэффициентов. Однако каждая группа должна содержать определенное количество служебных битов, в которых записана длина сооб-

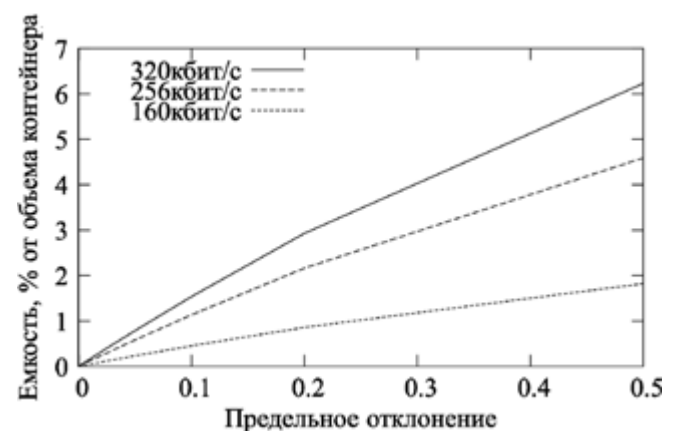


Рис. 2

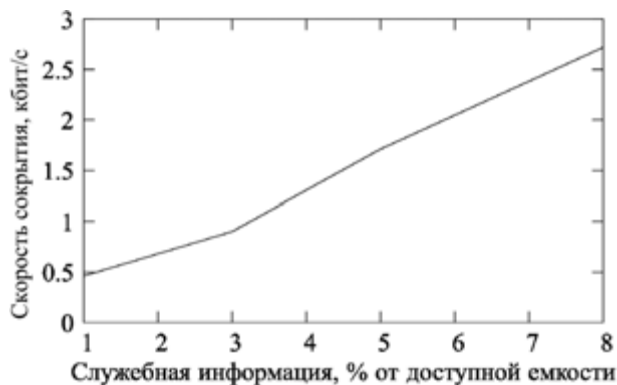


Рис. 3

Параметры	Алгоритм <i>mp3stego</i>	Межкадровое сокрытие	Сокрытие в MP3-коэффициентах
Емкость, %	0,015	33	0,55
Скорость сокрытия, кбит/с	0,07	27 833	1,02
Скорость извлечения, кбит/с	0,15	55 109	4,43
Степень защиты от обнаружения	Низкая	Низкая	Высокая

шения в этой группе. Следовательно, чем меньше группы (и быстрее работает метод), тем больше битов расходуется на служебную информацию. Указанная тенденция отражена на рис. 3.

Данный метод при высокой защите от обнаружения предоставляет пропускную способность порядка 0,5 % от объема контейнера. Его предлагается использовать в тех ситуациях, когда ожи-

дается серьезное противодействие информированного противника.

4. Экспериментальные результаты

Сравнение алгоритмов проводилось со следующими параметрами: битрейт контейнера 256 кбит/с, размер сообщения — максимально возможный для каждого алгоритма, предельное отклонение при сокрытии в MP3-коэффициентах 0,05, в группе кадров 256 коэффициентов. Данные по скорости работы учитывают время, потраченное на MP3-кодирование для требующих этого алгоритмов. Результаты сравнения алгоритмов приведены в таблице.

Отметим, что алгоритм *mp3stego* модифицирует наименьший значащий бит длины сжатых данных в кадре. Увеличение этого значения невозможно ввиду ограничения на общую длину кадра при заданном битрейте. Таким образом, файлы со скрытой информацией в среднем имеют меньшую длину данных в кадрах, чем стандартные.

Список литературы

1. Грибунин В. Г., Оков И. Н., Туринцев И. В. Цифровая стеганография. М.: СОЛОН-Пресс, 2002.
2. Сычев А. В., Александров Э. В. О возможности скрытной передачи информации внутри цифровых звуковых сигналов на основе учета эффектов маскирования. Труды 3-й Международной конференции DSPA-2000. Москва, 2000.
3. Brandenburg K., Stoll G. ISO-MPEG-1 Audio: A General Standard for Coding of High-Quantity Digital Audio // J. Audio Eng. Soc. 1994. N 42 (10). P. 780—792.
4. <http://www.petitcolas.net/fabien/steganography/mp3stego>
5. Fridrich J., Goljan M., Soukal D. Efficient wet paper codes // Proc. Information Hiding Workshop, Springer LNCS. 2005. С. 204—218.

УДК 004.056.52, 004.056.53

К. А. Шапченко, ст. науч. сотр.,

Институт проблем информационной безопасности, МГУ им. М. В. Ломоносова,
e-mail: shapchenko@iisi.msu.ru

Способ проверки свойств безопасности в моделях логического разграничения доступа с древовидной иерархией объектов доступа

Рассматривается задача о спецификации и проверке свойств безопасности в моделях логического разграничения доступа. Для проверки класса таких свойств, а именно — свойств информационных потоков, предлагается подход, в основу которого положены методы верификации на модели. Описывается применение таких методов в процессе анализа моделей логического разграничения доступа в Unix-подобных операционных системах с учетом древовидной иерархической структуры, заданной на объектах доступа.

Ключевые слова: информационная безопасность, модели логического разграничения доступа, верификация на модели.

Введение

Одним из требований, которые предъявляются к автоматизированным системам с повышенным уровнем защищенности, является строгое фор-

мальное доказательство корректности работы программных компонентов таких систем. На этом направлении представляется уместным выделение класса задач формальной проверки коррект-

ности конфигурации — набора настроек, подаваемых на вход программ. В настоящей работе рассматривается задача из такого класса, а именно — задача проверки свойств безопасности в моделях логического разграничения доступа (ЛРД).

Модели ЛРД служат основой для механизмов логического разграничения доступа, в том числе — реализуемых в ядрах операционных систем (ОС). Задача механизмов ЛРД состоит в проверке, является ли доступ некоторого субъекта к некоторому объекту разрешенным в заданной системе правил разграничения доступа. Выделим две категории правил разграничения доступа. К первой из них относятся правила, действующие в некотором фиксированном состоянии автоматизированной системы. Такие правила составляют статическую часть правил логического разграничения доступа. Ко второй категории — динамической части правил ЛРД — относятся правила, описывающие логику принятия решения о выполнении административных действий, а именно — действий по изменению правил ЛРД.

В рамках типовых подходов к управлению настройками механизмов ЛРД, как правило, реализуется только "локальное" изменение правил ЛРД — для одного или нескольких объектов доступа. Такая особенность приводит к фрагментарному управлению правилами ЛРД без оценки всех таких правил в их объединении. Указанное обстоятельство может привести к ошибкам настройки, в результате которых становится возможным несанкционированный доступ к контролируемым объектам. Проверка "вручную" всех правил и оценка соответствия их совместной работы заданным требованиям, как правило, затруднены для администратора безопасности. Причина — в большом объеме конфигурационных данных механизмов ЛРД в современных автоматизированных системах. Перечисленные особенности процесса настройки механизмов логического разграничения доступа приводят к задачам автоматизации процедуры комплексного управления настройками механизмов ЛРД. К таким задачам относится автоматизированная проверка выполнения заданных требований в конфигурации таких механизмов. Одним из классов подходов к решению задач проверки выполнения свойств в моделях ЛРД является использование алгоритмов и средств верификации на модели ("model checking"). Ряд современных способов использования такого инструментария при решении рассматриваемой задачи для нескольких видов упрощенных моделей ЛРД представлен в работах [1—4]. В настоящей статье предлагается подход к проверке свойств в модели ЛРД с древовидной иерархией объектов доступа, близкой к иерархии файлов и каталогов во многих современных файловых системах. Излагается также ряд замечаний о сокращении размера задачи в

целях обеспечения возможности проведения автоматизированной проверки с помощью универсальных средств верификации.

1. Общая схема подхода к проверке свойств

Предлагаемый подход реализует статическую проверку требований в модели ЛРД. Общая схема подхода представляет определенную последовательность действий:

- сбор данных о конфигурации механизмов ЛРД;
- представление полученных данных в виде формальной модели;
- задание проверяемых свойств на используемом языке спецификации;
- проверка выполнения свойств в модели;
- интерпретация результатов с точки зрения настроек механизмов ЛРД.

В основе предлагаемого подхода, а именно — на шаге проведения проверки, используются алгоритмы верификации на модели. Методы верификации на модели ("model checking") основаны на задании модели как системы переходов между состояниями, для каждого из которых задан набор истинных в нем атомарных высказываний пропозициональной логики. Проверяемое свойство специфицируется в виде формулы временной логики, накладывающей ограничения на последовательность состояний в путях в заданной системе переходов [5]. Для проведения автоматической проверки подобных свойств существует ряд инструментальных средств, в том числе с открытым исходным кодом [6].

Отметим ряд аспектов предлагаемых подходов в контексте проверки свойств моделей ЛРД, которые реализуются механизмами операционных систем. Сбор данных о конфигурации механизмов ЛРД проводится штатными средствами ОС, позволяющими получить действующие правила разграничения доступа. Собранная таким образом структурированная информация о настройках механизмов ЛРД преобразуется к теоретико-множественной модели логического разграничения доступа, на основе которой строится модель для проверки. В рассматриваемом подходе к верификации в качестве такой модели для проверки используется система переходов [5]. Проверяемые свойства задаются как формулы на входном языке спецификации свойств в средстве верификации. Проверка свойств осуществляется в автоматизированном режиме. В случае ложности проверяемого свойства средство проверки предоставляет контрпример, а именно — последовательность переходов между состояниями, на которой заданное свойство не выполняется. Интерпретация контрпримера в терминах конфигурации механизма ЛРД и заключение об истинности проверяемых свойств являются решением задачи проверки вы-

полнения заданных требований в исследуемой модели ЛРД.

2. Анализируемая модель и проверяемые свойства

Предлагаемая модель ЛРД описывает правила разграничения доступа, близкие к традиционно используемым в Unix-подобных ОС. Отличиями предлагаемой модели являются:

- отсутствие описания административных действий (т. е. рассматривается только статическая часть модели ЛРД);
- отсутствие описания специальных атрибутов доступа, например SUID-битов.

Предлагаемая модель не претендует на полноту в смысле описания всех возможностей по разграничению доступа в ОС. Она поддерживает лишь описание части механизмов ЛРД к объектам файловых систем.

Для описания правил, по которым определяется, разрешен ли доступ в предлагаемой модели ЛРД, используется теоретико-множественный подход. Определяется набор базовых множеств, описывающих основные сущности, используемые в модели. Для этих множеств описывается ряд связывающих их отношений и формулируется предикат предоставления доступа, аргументами которого являются субъект, объект и тип доступа. Значение предиката определяет, разрешен ли такой доступ в рамках заданной статической части правил ЛРД. В качестве базовых в предлагаемой модели определяются следующие конечные множества:

- U и G — множества пользователей и групп пользователей (задаваемых, например, уникальными алфавитно-цифровыми идентификаторами);
- O и S — множества объектов доступа (объектов файловой системы) и субъектов доступа (процессов в ОС);
- A — множество типов доступа, в рассматриваемой модели принимается, что $A = \{r, w, x\}$, где r — доступ на чтение, w — доступ на запись, x — доступ на выполнение файла или на чтение списка элементов каталога.

Определяются следующие функции (отношения): $user: S \cup O \rightarrow U$ и $group: S \cup O \rightarrow G$ — определение идентификатора пользователя-владельца и группы-владельца;

$groups: U \rightarrow 2^G$ — принадлежность пользователя множеству групп;

$type: O \rightarrow \{File, Dir\}$ — функция разделения объектов файловых систем на файлы и каталоги (другие типы подобных объектов в целях упрощения примера не рассматриваются);

$parent: O \rightarrow \{o \in O \mid type(o) = Dir\} \cup \{None\}$ — функция, определяющая родительский каталог файлового объекта, $None$ в случае, если исходный объект является корневым каталогом; на функцию $parent$ устанавливается ограничение — она

задает древовидную иерархию с единым корнем на рассматриваемых объектах файловой системы — файлах и каталогах;

$u_perm: O \rightarrow 2^A$, $g_perm: O \rightarrow 2^A$, $o_perm: O \rightarrow 2^A$ — функции, задающие множества прав доступа для пользователя-владельца объекта, для пользователей из группы-владельца объекта и для других пользователей соответственно.

Предикат предоставления доступа формулируется следующим образом:

$\Delta(s, o, a) = \text{ложь}$, если $o = None$;

$(parent(o) = None \vee \Delta(s, parent(o), x)) \wedge$
 $\wedge (user(s) = user(o) \wedge$

$\wedge a \in u_perm(o)) \vee (group(o) \in$
 $\in groups(user(s)) \wedge a \in g_perm(o)) \vee$

$\vee (a \in o_perm(o))$, иначе.

Укажем класс актуальных свойств информационной безопасности, проверка которых реализуется предлагаемым методом, а именно — ограничений на информационные потоки, разрешенные согласно исследуемой модели. В первую очередь определяется понятие элементарного информационного потока, а на его основе строится определение для общего вида информационного потока. Пусть в рассматриваемой модели ЛРД рассматривается субъект доступа s и объект доступа o . Будем считать, что элементарный информационный поток от s к o разрешен, если правилами ЛРД допускается доступ "на запись" от s к o . Аналогичным образом, элементарный информационный поток от o к s разрешен, если правилами допускается доступ "на чтение" от s к o . Под типами доступа "на чтение" и "на запись" здесь понимаются такие, что при реализации соответствующего доступа информация передается от объекта к субъекту (для типа доступа "на чтение") и от субъекта к объекту (для типа доступа "на запись"). Из нескольких таких элементарных информационных потоков складываются составные допустимые информационные потоки. Будем считать, что между сущностями e_1 и e_2 , каждая из которых может быть субъектом или объектом доступа в модели ЛРД, разрешен информационный поток, если существует конечная последовательность элементарных информационных потоков такая, что первый элементарный поток начинается в e_1 , последний элементарный поток заканчивается в e_2 , и конец каждого элементарного потока кроме последнего совпадает с началом следующего элементарного потока.

Общий вид одного из типовых свойств информационных потоков формулируется следующим образом. Пусть задано ограничение на информационный поток, в частности, — указаны требования к начальной и конечной сущностям в потоке, а также к последовательности сущностей и типов доступа в потоке, например, требование того, что одна сущность e_1 всегда идет после сущности e_2 и

между ними всегда есть сущность e_3 . Необходимо проверить, выполняется ли это свойство для всех допустимых информационных потоков в заданной модели ЛРД. В качестве основного свойства, на котором демонстрируется применение методов верификации, выберем следующее: при заданных классах сущностей модели E_1 , E_2 и E_3 любой информационный поток от сущности из E_1 к сущности из E_2 проходит через одну из сущностей в классе E_3 .

3. Преобразование модели ЛРД в систему переходов и проведение проверки

Опишем построение системы переходов по заданной модели ЛРД. Каждое состояние такой системы задается значениями нескольких переменных. В предлагаемом способе построения используются следующие переменные состояния:

- атрибуты субъекта доступа, определяемые в рассматриваемой модели идентификатором пользователя $sv_{user} \in U$;
- субъект или объект доступа $sv_{obj} \in O \cup \{User\}$, где уникальная константа $User$ соответствует субъектам доступа, причем доопределяется функция $type$ таким образом, что $type(User) = User$;
- тип доступа $sv_{action} \in A \cup \{x_0, n\}$, дополнительный тип доступа x_0 дублирует тип доступа x и используется вместе с ним для учета иерархии объектов доступа в модели ЛРД, а дополнительный фиктивный тип доступа n предназначен для обозначения отсутствия выполнения доступа в состоянии.

Для обеспечения свойства полноты системы переходов — отсутствия в ней тупиковых состояний — вводится дополнительная булева переменная состояния sv_{good} , а множества значений других переменных состояния расширяются уникальной константой $BadState$. Свойство полноты в системе достигается за счет возможности перехода из любого состояния в состояние ($sv_{user} = BadState$, $sv_{obj} = BadState$, $sv_{action} = BadState$, $sv_{good} = \text{ложь}$). Начальные состояния в такой системе переходов удовлетворяют условию $sv_{user} \neq BadState \wedge sv_{obj} \neq BadState \wedge sv_{action} \neq BadState \wedge sv_{good}$.

Предикат возможности перехода из одного состояния в другое формулируется следующим образом (переменные нового состояния отмечены префиксом " $n_$ "):

$$\begin{aligned} & (sv_{good} \wedge n_{sv_{good}} \wedge (\\ & (type(sv_{obj}) = User \wedge n_{sv_{user}} = sv_{user} \wedge \\ & \wedge parent(n_{sv_{obj}}) = None \wedge n_{sv_{action}} \in \{w, x_0\} \wedge \\ & \wedge \Delta'(n_{sv_{user}}, n_{sv_{obj}}, n_{sv_{action}})) \vee \\ & \vee (type(sv_{obj}) = Dir \wedge n_{sv_{user}} = sv_{user} \wedge sv_{obj} = \\ & = parent(n_{sv_{obj}}) \wedge type(n_{sv_{obj}}) = Dir \wedge action = x_0 \wedge \\ & \wedge n_{sv_{action}} \in \{w, x_0\} \wedge \Delta'(n_{sv_{user}}, n_{sv_{obj}}, n_{sv_{action}})) \vee \end{aligned}$$

$$\begin{aligned} & \vee (type(sv_{obj}) = Dir \wedge n_{sv_{user}} = sv_{user} \wedge sv_{obj} = \\ & = parent(n_{sv_{obj}}) \wedge type(n_{sv_{obj}}) = File \wedge sv_{action} = \\ & = x_0 \wedge n_{sv_{action}} = w \wedge \Delta'(n_{sv_{user}}, n_{sv_{obj}}, n_{sv_{action}})) \vee \\ & \vee (type(sv_{obj}) = File \wedge n_{sv_{user}} = sv_{user} \wedge n_{sv_{obj}} = parent(sv_{obj}) \wedge \\ & \wedge n_{sv_{action}} = x \wedge sv_{action} \in \{r, x\} \wedge \Delta'(sv_{user}, sv_{obj}, sv_{action}) \wedge \\ & \wedge \Delta'(n_{sv_{user}}, n_{sv_{obj}}, n_{sv_{action}})) \vee \\ & \vee (type(sv_{obj}) \in \{File, Dir\} \wedge n_{sv_{user}} \neq sv_{user} \wedge n_{sv_{obj}} = sv_{obj} \wedge \\ & \wedge n_{sv_{action}} = r \wedge sv_{action} = w \wedge \Delta'(sv_{user}, sv_{obj}, sv_{action})) \vee \\ & \vee (type(sv_{obj}) = Dir \wedge n_{sv_{user}} = sv_{user} \wedge n_{sv_{obj}} = \\ & = parent(sv_{obj}) \neq None \wedge n_{sv_{action}} = x \wedge sv_{action} \in \{r, x\} \wedge \\ & \wedge \Delta'(sv_{user}, sv_{obj}, sv_{action})) \vee \\ & \vee (type(n_{sv_{obj}}) = User \wedge n_{sv_{user}} = sv_{user} \wedge parent(sv_{obj}) = \\ & = None \wedge n_{sv_{action}} = a_{none} \wedge sv_{action} \in \{r, x\} \wedge \\ & \wedge \Delta'(sv_{user}, sv_{obj}, sv_{action})) \vee \\ & \vee (\neg n_{sv_{good}} \wedge n_{sv_{user}} = n_{sv_{obj}} = n_{sv_{action}} = BadState). \end{aligned}$$

Предикат $\Delta'(u, o, a)$ строится на основе определенного ранее предиката $\Delta(s, o, a)$ таким образом, что их значения совпадают при условии $user(s) = u$.

Факторизация полученной модели в целях уменьшения числа состояний может быть проведена с использованием следующего подхода. Рассмотрим множества файлов — элементов одного и того же подкаталога в древовидной иерархии, определяемой отношением $parent$. Будем использовать тот факт, что нередко в Unix-подобных ОС в одном каталоге содержатся большие группы файлов с одним и тем же владельцем и одинаковыми заданными правилами доступа к ним. Каждое такое множество факторизуем с помощью группировки по эквивалентным правилам доступа, заданным для элементов-файлов. Отношение эквивалентности формулируется следующим образом: $eq(o_1, o_2) \equiv parent(o_1) = parent(o_2) \wedge \forall u \in U \forall a \in A: \Delta'(u, o_1, a) = \Delta'(u, o_2, a)$.

Дальнейшая группировка объектов доступа может использовать, например, следующее свойство. Рассмотрим некоторый каталог в иерархии объектов файловой системы и поддереву с вершиной в нем в такой иерархии. Пусть выполнено условие, что все каталоги и файлы в этом поддереве принадлежат одному и тому же владельцу, а правила доступа других пользователей к ним совпадают. Тогда рассматриваемое поддерево может быть заменено на поддерево с той же вершиной, однако состоящее только из объектов доступа, соответствующих файлам в оригинальном поддереве, без структуры каталогов. Применяя к такому набору файлов факторизацию по отношению eq , получим поддерево с сокращенным числом элементов. Указанный подход к сокращению числа объектов доступа позволяет отождествить в модели объекты доступа, соответствующие

однотипному содержимому в пользовательских и в системных каталогах файловых систем. Следует отметить, что факторизация множества объектов доступа приводит к тому, что в проверяемых свойствах не представляется возможным указать некоторые объекты доступа. Возможно указать только классы эквивалентности, к которым они принадлежат. Это обстоятельство необходимо учитывать при формулировке свойств для последующей проверки.

В применении к типовым дистрибутивам ОС Linux и к другим Unix-подобным ОС представленные подходы позволяют провести сокращение числа объектов доступа с нескольких сотен тысяч до тысячи и менее. Получаемый объем системы переходов позволяет осуществлять проверку свойств в модели с использованием универсальных средств верификации на модели.

В качестве языка спецификации свойств выбрана линейная временная логика (LTL), расширенная операторами прошедшего времени [5, 6]. Типовое свойство (см. раздел 2) прохождения информационного потока от сущностей класса E_1 к сущностям класса E_2 через сущности выделенного класса E_3 в модели записывается на входном языке LTL-спецификаций NuSMV [6] следующим образом:

$$I_{E_1}(sv) \rightarrow \neg (sv_{good} U(F(I_{E_2}(sv) \wedge \wedge sv_{good} \wedge H \neg I_{E_3}(sv))))),$$

где $I_{E_1}(sv)$, $I_{E_2}(sv)$, $I_{E_3}(sv)$ — предикаты соответствия состояний системы переходов $sv = (sv_{user}, sv_{obj}, sv_{action}, sv_{good})$ классам E_1 , E_2 и E_3 .

Для апробации предлагаемых методов проверки свойств использовали средство символьной верификации на модели NuSMV [6]. Входные данные для средства проверки формировали набором программ, осуществляющих сбор данных о правилах разграничения доступа и преобразование их к модели, в которой происходит проверка свойств. Верифицируемые свойства задавали на языке входных данных для средства проверки с использованием генератора типовых свойств по заданным шаблонам, включая шаблон для формул, соответствующих свойству о прохождении информационного потока через заданный класс сущностей модели ЛРД.

В случае истинности проверяемого свойства в модели дополнительных действий, направленных на интерпретацию результата проверки, как правило, не требуется. Интерес представляет случай отрицательного результата проверки свойства. В такой ситуации средство верификации на модели предоставляет контрпример — трассу в системе переходов, на которой заданное свойство не выполняется. Для интерпретации подобной трассы в терминах исследуемой модели ЛРД необходимо выполнить преобразование из системы переходов

в набор сущностей оригинальной модели разграничения доступа, а именно — отобразить состояния в трассе как субъекты и объекты доступа в модели ЛРД, а переходы — как действия, проводимые над объектами доступа. Следует отметить, что ввиду группировки объектов доступа при формировании системы переходов одному состоянию может соответствовать большое число сущностей из оригинальной модели ЛРД.

Заключение

Как продемонстрировано в настоящей работе и в ряде других работ в области формальной верификации, для решения задачи проверки свойств в моделях ЛРД могут быть использованы методы верификации на модели. В целях применения подобных методов на практике необходима адаптация способов проверки к типовым моделям логического разграничения доступа в современных автоматизированных системах. Важным примером таких моделей являются модели ЛРД, которые реализуются в механизмах разграничения доступа в операционных системах. В настоящей работе показан подход к подобной адаптации методов проверки на основе учета иерархии объектов доступа и способов их группировки в целях сокращения размера проверяемой модели. Полученные результаты позволяют проводить проверку типовых моделей ЛРД с использованием универсальных средств верификации. К дальнейшей работе в этом направлении следует отнести создание оптимизированных алгоритмов проверки, а также разработку методов и средств для формирования типовых свойств и для интерпретации результатов проверки.

Список литературы

1. Васенин В. А., Шапченко К. А., Андреев О. О. Математические модели и механизмы логического разграничения доступа в операционной системе Linux: текущее состояние и перспективы развития // Матер. Второй междунар. науч. конф. по проблемам безопасности и противодействия терроризму. Пятая общеросс. науч. конф. "Математика и безопасность информационных технологий" (МаБИТ-06). МГУ им. М. В. Ломоносова, 25–26 октября 2006 г. М.: МЦНМО, 2007. С. 159–171.
2. Шапченко К. А. К вопросу о средствах ОС Linux для управления доступом при использовании ролевых политик безопасности // Математика и безопасность информационных технологий. Матер. конф. в МГУ 2–3 ноября 2005 г. М.: МЦНМО, 2006. С. 257–281.
3. Guelev D. P., Ryan M. D., Schobbens P.-Y. Model-checking access control policies // In Information Security Conference, number 3225 in Lecture Notes in Computer Science. Springer-Verlag, 2004.
4. Guttman J. D., Herzog A. L., Ramsdell J. D., Skorupka C. W. Verifying information flow goals in security-enhanced Linux // Journal of Computer Security. 2004. Vol. 13.
5. Кларк Э. М., Грамберг О., Пелед Д. Верификация моделей программ: Model Checking: Пер. с англ. / Под. ред. Р. Смелянского. М.: МЦНМО, 2002.
6. Cimatti A., Clarke E. M. et al. NuSMV 2: An Open Source Tool for Symbolic Model Checking // In Proc. of International Conference on Computer-Aided Verification (CAV 2002). Copenhagen, Denmark, July 27–31, 2002. <http://nusmv.irst.itc.it/>.

УДК 004.42

В. А. Мочалов, аспирант,
Московский технический университет связи
и информатики (МТУСИ),
e-mail: mvaproduct@mail.ru

Алгоритмы размещения транзитных узлов в сенсорной сети

Предлагаются алгоритмы размещения транзитных узлов (Т-узлов) в сенсорной сети (СС). Описываются алгоритмы базового размещения Т-узлов и даются оценки эффективности работы таких алгоритмов. В структуру СС согласно одному из алгоритмов отказоустойчивого (ОУ) размещения Т-узлов добавляются избыточные Т-узлы, которые увеличивают вероятность связности функциональных узлов СС с сервером. В работе рассматриваются два алгоритма ОУ-размещения Т-узлов: первый основан на использовании генетических алгоритмов, второй размещает Т-узлы согласно заданной степени связности функциональных узлов с сервером.

Ключевые слова: сенсорные сети, отказоустойчивость, связность, алгоритмы.

Сенсорная сеть и задача построения отказоустойчивой СС

Сенсорная сеть (СС) — это распределенная сеть *необслуживаемых* миниатюрных электронных устройств (узлов сети), обменивающихся информацией *по беспроводной связи* и способных регистрировать данные о параметрах внешней среды и передавать их на базовую станцию с помощью ретрансляции от узла к узлу [1] (рис. 1).

Узел сети, называемый *сенсором*, содержит датчик, воспринимающий данные от внешней среды (собственно сенсор), микроконтроллер, память, радиопередатчик, автономный источник питания [2]. Полагаем, что СС должна быть размещена на некотором объекте, распределенном в пространстве, причем заданы точки этого объекта, в которых обязательно должны быть размещены узлы сети, осуществляющие сбор информации в некоторой окрестности каждой точки и называемые *функциональными узлами* (Ф-узлы). Информация, полученная этими узлами, должна быть передана с требуемой степенью надежности в центральный узел сети (сервер) с помощью определенным образом размещенных в пределах данного объекта *транзитных узлов* (Т-узлов), осуществляющих только передачу информации [3].

Считаем, что известны надежность Ф-узлов и Т-узлов, радиоканалов связи, создаваемых ими, радиусы передачи радиосигнала и стоимость узлов. Искомая структура СС представляется графом, вершинам которого соответствуют функциональные и транзитные узлы СС, состоящие из некоторого числа (от 1 до m) элементарных узлов (ЭУ) и размещенные в определенных точках объекта, причем точки размещения функциональных узлов фиксированы, а точки размещения транзитных узлов должны быть найдены. Ребрам графа соответствуют радиоканалы между узлами сети [4].

Требуется сформировать отказоустойчивую структуру СС, которая обеспечивает "наилучшее значение" функции приспособленности СС, характеризующей как степень надежности сети, так и затраты, требуемые для достижения необходимо-

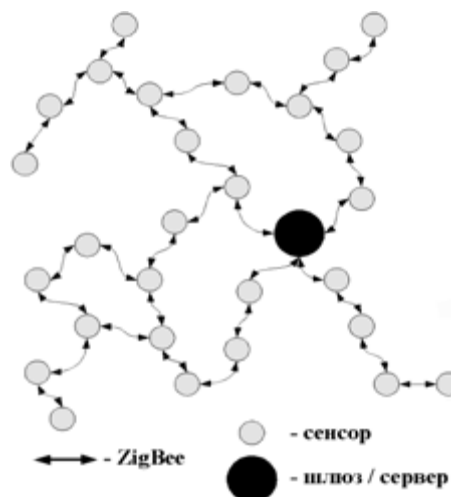


Рис. 1. Структура сенсорной сети

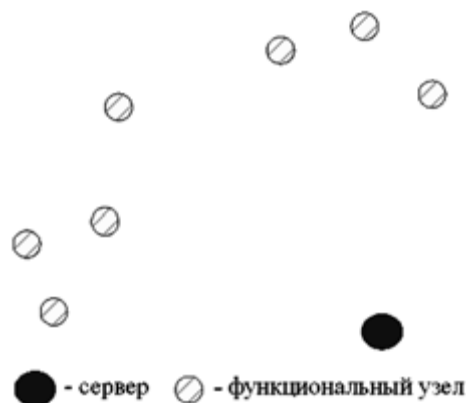


Рис. 2. Пример исходных данных в задаче размещения Т-узлов

го уровня ее отказоустойчивости. Для выполнения этой задачи рассмотрим алгоритмы размещения транзитных узлов и продемонстрируем результаты работы каждого алгоритма на примере, показанном на рис. 2. В примере заданы координаты Ф-узлов и сервера, требуется найти координаты Т-узлов.

Алгоритмы базового размещения транзитных узлов

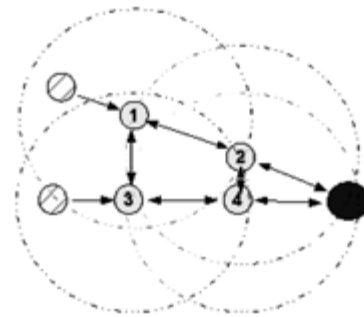
Простое размещение транзитных узлов (ПРТУ). Простым способом размещения транзитных узлов (Т-узлов) является их установка по прямой линии от Ф-узла к серверу на расстоянии радиуса действия друг от друга [5]. Такое простое размещение предоставляет каждому Ф-узлу свой собственный путь к серверу (рис. 3). Вычислительная сложность данного алгоритма $O(n_s d_m)$, где n_s — число Ф-узлов, а d_m — максимальное число ретрансляции от одного Ф-узла к серверу. Число размещенных Т-узлов оценивается величиной $O(n_s d_m)$.

На рис. 4 показано размещение Т-узлов в рассматриваемом примере алгоритмом ПРТУ. Пунктирной линией показана прямая линия от Ф-узла к серверу.

Простое размещение многократно используемых транзитных узлов (ПРМИТУ). Легко убедиться, что алгоритм ПРТУ размещает большое число избыточных Т-узлов. Например, в размещении, показанном на рис. 3, один из Т-узлов 2 или 4 может быть удален, при этом итоговая структура сети все еще будет связной. Алгоритм может быть улучшен путем удаления избыточных Т-узлов и многократного использования уже развернутых Т-узлов. При этом каждый функциональный узел соединяется с ближайшим уже размещенным Т-узлом либо с сервером. На рис. 5 показано размещение Т-узлов в рассматриваемом примере алгоритмом ПРМИТУ.

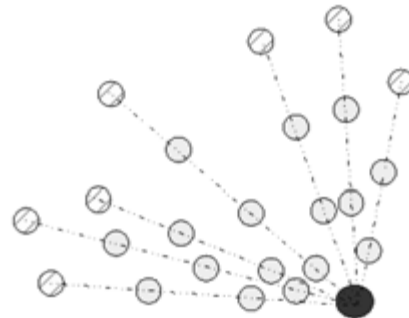
Описание алгоритма ПРМИТУ:

- 1) создать отсортированный по критерию близости к серверу массив Ф-узлов M_F ;
- 2) создать пустое множество размещенных узлов $\Omega_{ТУ}$ и добавить в него сервер;
- 3) создать переменную i , которая хранит индекс текущего Ф-узла из массива M_F и инициализировать ее значением 0 ($i = 0$). Ближайший к серверу Ф-узел принять текущим Ф-узлом F_c ($F_c = M_F[0]$);
- 4) взять ближайший к F_c узел ($T_{БЛ}$) из множества $\Omega_{ТУ}$. Если расстояние между F_c и $T_{БЛ}$ больше радиуса передачи радиосигнала Ф-узла, то установить по прямой линии от F_c к $T_{БЛ}$ Т-узлы на расстоянии радиуса действия друг от друга. Добавить размещенные Т-узлы во множество $\Omega_{ТУ}$;



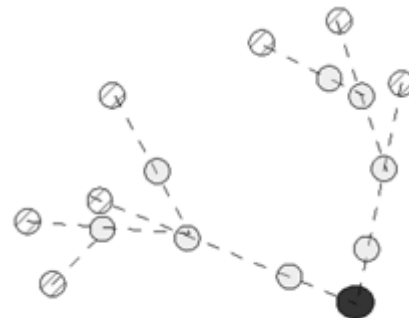
● - сервер ○ - Ф-узел ○ - Т-узел ↔ - ZigBee

Рис. 3. Пример простого размещения транзитных узлов сенсорной сети



● - сервер ○ - функциональный узел ○ - транзитный узел

Рис. 4. Размещение Т-узлов по алгоритму ПРТУ в рассматриваемом примере



● - сервер ○ - функциональный узел ○ - транзитный узел

Рис. 5. Размещение Т-узлов по алгоритму ПРМИТУ в рассматриваемом примере

5) если i меньше числа Ф-узлов, то увеличить i ($i = i + 1$) и принять в качестве текущего узла следующий ближайший к серверу Ф-узел $F_c = M_F[i]$. Повторять шаги 4 и 5 пока i не станет равным числу Ф-узлов.

Алгоритм ПРМИТУ работает медленнее по сравнению с ПРТУ, но в сети с большей плотностью размещения Ф-узлов ПРМИТУ разместит меньшее число Т-узлов. Для сети с малой плотностью Ф-узлов число размещенных Т-узлов оценивается величиной $O(n_s \sqrt{d_m})$, где n_s — число Ф-узлов, а d_m — максимальное число ретрансляций от одного Ф-узла к серверу. С увеличением плот-

ности Φ -узлов число T -узлов асимптотически стремится к значению $O(d_m^2)$. Таким образом, алгоритм ПРМИТУ, который в лучшем случае размещает d_m^2 T -узлов, лучше масштабируется по сравнению с алгоритмом ПРТУ, число размещенных узлов которого оценивается значением $O(n_s d_m)$. Вычислительная сложность алгоритма ПРМИТУ будет $O(n_s d_m^2)$.

Групповое размещение транзитных узлов (ГРТУ). Число размещенных T -узлов может быть уменьшено путем объединения Φ -узлов в группы. Задача заключается в поиске групп Φ -узлов, покрытии минимальным связным графом T -узлов каждой такой группы и осуществлении связи каждой группы с сервером с помощью алгоритма ПРМИТУ (рис. 6, см. третью сторону обложки).

Описание алгоритма ГРТУ:

1) сформировать такие группы Φ -узлов, чтобы расстояние между любыми двумя Φ -узлами одной и той же группы не превышало величины D , задаваемой проектировщиком. Φ -узел, являющийся единственным Φ -узлом группы, называется одиночным Φ -узлом;

2) создать пустое множество узлов (Ω_y), которые будем связывать с сервером; добавить одиночные Φ -узлы во множество Ω_y ;

3) для каждой группы, состоящей более чем из одного Φ -узла, найти минимальную покрывающую Φ -узлы окружность, в центре которой следует разместить центральный T -узел группы. Все Φ -узлы группы соединить с центральным T -узлом с помощью алгоритма ПРМИТУ. Добавить центральные T -узлы групп во множество Ω_y ;

4) связать все узлы множества Ω_y с сервером по алгоритму ПРМИТУ.

Вычислительная сложность алгоритма объединения Φ -узлов в группы оценивается величиной $O(n_s^2)$, где n_s — это число Φ -узлов. Вычислительная сложность же осуществления связи узлов множества Ω_y с сервером с помощью алгоритма ПРМИТУ оценивается величиной $O(n_c d_m^2)$, где n_c — это число элементов множества Ω_y , d_m — максимальное число ретрансляций от одного Φ -узла к серверу.

Оптимизация алгоритмов базового размещения T -узлов. Решения, полученные алгоритмами базового размещения T -узлов, могут быть не оптимальными. Так, некоторые T -узлы могут быть удалены, при этом Φ -узлы останутся связными с сервером. Для определения T -узлов, которые могут быть удалены, необходимо выполнить следующую процедуру:

1) удалить все ребра графа, описывающего CC ;

2) каждый Φ -узел соединить с таким T -узлом в пределах радиуса действия, расстояние от которого к серверу минимально;

3) удалить T -узлы имеющие меньше двух связей с другими узлами сети (Φ -, T -узлами или сервером).

Алгоритмы базового размещения транзитных узлов не учитывают возможность отказов T -узлов и каналов связи и обеспечивают для каждого Φ -узла наличие хотя бы одного пути между данным Φ -узлом и сервером.

Алгоритмы отказоустойчивого размещения транзитных узлов

Структуры CC , полученные с помощью алгоритмов базового размещения транзитных узлов, связаны до тех пор, пока работоспособны все T -узлы. Если же один или несколько T -узлов или каналов связи выйдут из строя, то часть CC становится несвязной с сервером (рис. 7, см. третью сторону обложки).

Основным подходом к повышению надежности CC является добавление избыточных путей таким образом, чтобы каждый Φ -узел имел более одного независимого пути к серверу.

Степень связности данного Φ -узла с сервером будем характеризовать числом независимых путей (HP) от данного Φ -узла к серверу. Степень связности, при которой каждый Φ -узел имеет k независимых путей к серверу, назовем k -уровнем связности. На рис. 8 (см. третью сторону обложки) показана структура CC со степенью связности, равной 2.

Простейшим способом увеличения степени связности Φ -узлов в k раз является размещение k T -узлов на позициях, выбранных одним из алгоритмов базового размещения T -узлов. Этот подход является простым, но в большинстве случаев можно достичь той же степени связности Φ -узлов, задействовав при этом меньшее число T -узлов. Более того, вследствие нахождения CC во враждебной внешней среде, достаточно велика вероятность одновременного отказа всех k -узлов, находящихся в одной пространственной точке CC . Рассмотрим алгоритмы отказоустойчивого размещения T -узлов.

Генетический алгоритм размещения транзитных узлов (ГАРТУ). Метод, позволяющий осуществить построение отказоустойчивой структуры CC при учете требований к вероятности связности ее Φ -узлов с сервером и стоимости внедрения описан в работе [4]. Этот метод основан на применении адаптированных генетических алгоритмов (ГА). Основными особенностями алгоритма ГАРТУ является прикладная реализация генетических операторов скрещивания и направленной мутации, а также функции приспособленности структуры CC на основе оценки вероятности связности (ОВС) ее узлов с помощью описываемого в статье [4] эвристического алгоритма.

В алгоритме ГАРТУ CC представляется в виде двумерной сетки $[N \times M]$, "покрывающей" объект

размещения СС, где N — число рядов (строк), а M — число столбцов. Размер $h \times w$ ячейки этой сетки, от которого зависят значения N и M (при заданной площади объекта), определяется заданным радиусом R передачи радиосигнала сенсора. Длина ребра ячейки не должна превышать R . На пространстве, соответствующем каждой ячейке сетки, может быть размещено не более одного узла СС (Ф или Т), состоящего из некоторого числа (от 1 до m) элементарных узлов. Поэтому при выборе структуры СС необходимо каждой ячейке сетки поставить в соответствие определенное число k узлов, причем для каждой из ячеек, содержащих заданные точки размещения Ф-узлов, $k = 1, \dots, m$, а для остальных ячеек, в которых могут быть размещены только Т-узлы, $k = 0, \dots, m$ [3]. Размещение Т-узлов в алгоритме ГАРТУ осуществляется с помощью генетического алгоритма по критерию вероятности связности Ф-узлов с сервером и стоимости всей СС. Рассмотрим основные определения и особенности алгоритма.

Популяция — множество возможных структур СС. **Хромосома** — структура СС (рис. 9, см. третью сторону обложки). Каждая хромосома состоит из N генов. Ген представляет собой конкретный горизонтальный разрез сенсорной сети (точнее, покрывающей ее сетки), состоящей из M ячеек. Каждая ячейка содержит тип хранимого узла (Ф или Т) и число элементарных узлов, из которых строится комплексный узел. Ф-узел может быть построен на основе как принципа голосования, так и принципа резервирования замещением. Т-узел строится только на основе резервирования замещением [4].

Инициализация — формирование исходного множества возможных структур СС (начальной популяции). Для создания начальной популяции необходимо первоначально сформировать такую базовую структуру (БС) СС, которая имеет ненулевую вероятность связности всех функциональных узлов с сервером и небольшую стоимость. БС формируется на основе одного из алгоритмов базового размещения Т-узлов. Структуры, входящие в начальную популяцию, формируются путем случайного добавления к базовой структуре Т-узлов, образующих с БС связный граф.

Оценка структуры СС осуществляется на каждой итерации ГА на основе *функции приспособленности* (ФП), которая учитывает оценки надежности СС и стоимости ее внедрения. Один из возможных подходов — определение ФП путем умножения значений показателя, характеризующего надежность СС, на значения функции F_{sf} , которая характеризует степень (в процентах) удовлетворения потребности проектировщика в стоимости проекта СС. Надежность структуры СС при заданном множестве ее Ф-узлов $\{D_f\}, f = 1, \dots, k$, определяется набором значений *вероятностей связ-*

ности P_{fs} каждого из узлов заданного множества $\{D_f\}$ с сервером. В качестве показателя надежности структуры СС принимается *минимальное* из этих значений $P_{FS} = \min\{P_{fs}\}, f = 1, \dots, k$. С помощью алгоритма *оценки вероятности связности* вычисляется нижняя оценка вероятности связности двух узлов сенсорной сети (P_{fs}). Алгоритм имеет линейную вычислительную сложность. Он основан на уменьшении размерности графа, описывающего СС, до такого размера, при котором два рассматриваемых узла СС становятся соседями, после чего методом прямого перебора вычисляется вероятность связности этих узлов в небольшой окрестности. Уменьшение размерности осуществляется путем последовательного сжатия графа в 4 раза на каждой итерации процедуры сжатия графа, после чего осуществляется подсчет вероятностей связи друг с другом сжатых соседних узлов.

Применение генетических операторов к хромосомам (структурам СС), отобраным с помощью селекции, приводит к формированию новой популяции структур СС. В процедуре применяются два основных генетических оператора: *оператор скрещивания* [8] и *направленный оператор мутации* [4]. Вероятности скрещивания и мутации (СМ) выбираются проектировщиком. Для увеличения скорости работы алгоритма может быть применена функция изменения вероятностей СМ [4].

На рис. 10 (см. четвертую сторону обложки) показан пример размещения с помощью алгоритма ГАРТУ Т-узлов на базовую структуру СС, полученную с помощью алгоритма ГРТУ (см. рис. 6).

Основным достоинством алгоритма ГАРТУ является высокое качество размещения Т-узлов. К недостаткам алгоритма ГАРТУ можно отнести высокую вычислительную и пространственную сложность, а также сложность реализации алгоритма.

Избыточное размещение транзитных узлов (ИРТУ). Имеется граф, содержащий вершины, соответствующие Ф-узлам, серверу и Т-узлам, размещенным в соответствии с одним из алгоритмов базового размещения Т-узлов, и называемый далее *начальным графом*.

Алгоритм ИРТУ добавляет избыточные Т-узлы в начальный граф для достижения требуемой степени связности Ф-узлов. Алгоритм ИРТУ состоит из двух процедур.

1. **Процедура вычисления количества независимых путей (ВКНП)** определяет для начального графа *степень связности* каждого Ф-узла с сервером. Процедура ВКНП выполняется в два шага. На первом шаге двунаправленный связный граф (ДСГ), представляющий СС, преобразуется в ориентированный связный граф (ОСГ) для вычисления независимых путей (НП) с помощью упрощенного алгоритма Форда — Фалкерсона (Ф-Ф). На втором шаге процедуры вычисляется макси-

мальный поток каждого Φ -узла к серверу, который и определяет количество независимых путей Φ -узлов с сервером. С помощью алгоритма Φ - Φ , описанного в работах [6, 7], осуществляется вычисление максимального потока от истока к стоку во взвешенном ориентированном графе. Для использования алгоритма Φ - Φ необходимо преобразовать ДСГ в ОСГ путем замены всех пар Т-узлов, связанных двунаправленным ребром, на четыре Т-узла, циклически связанных однонаправленными ребрами. Таким образом, один Т-узел преобразуется в два Т-узла, причем все входящие связи с других узлов поступают на первый Т-узел, исходящие же связи идут от второго Т-узла (рис. 11, см. четвертую сторону обложки). Необходимо подчеркнуть, что имеется только одно преобразование из ДСГ в ОСГ, которое не зависит от способа обхода ДСГ.

Для того чтобы использовать одно ребро не более чем в одном пути, установим поток через каждое ребро в единицу. Следовательно, упрощенный алгоритм Φ - Φ для поиска максимального потока между двумя узлами (Φ -узлом и сервером) вычисляется следующим образом:

1) находится возможный путь в графе с использованием как недействованных, так и уже действовавших ребер, но направленных в противоположном направлении;

2) ребра найденного пути помечаются как задействованные, ребра же, направленные в противоположном направлении найденного пути, помечаются недействованными;

3) повторяются шаги 1 и 2 до тех пор, пока не найдутся новые пути;

4) максимальный поток от Φ -узла к серверу равен числу независимых путей от Φ -узла к серверу.

На рис. 12 (см. четвертую сторону обложки) показан пример нахождения НП между Φ -узлом и сервером. Вначале находится путь с использованием недействованных ребер с потоком 1, эти ребра помечаются как задействованные [5] (рис. 12, б). Затем находится новый путь с использованием одного задействованного ребра, но направленного в противоположном направлении (рис. 12, в). После этого ребра найденного пути отмечаются как задействованные, вес ребра, направленного в противоположном направлении, устанавливается в 0 (рис. 12, г). Таким образом, суммарное число НП равняется двум.

2. Процедура размещения избыточных Т-узлов (РИ Т-узлов) осуществляет добавление дополнительных Т-узлов в структуру СС для достижения требуемой степени связности Φ -узлов с сервером.

Описание алгоритма процедуры РИ Т-узлов:

1) создать отсортированный по критерию близости к серверу массив Φ -узлов M_F ;

2) создать переменную i , которая хранит индекс текущего Φ -узла из массива M_F и инициализировать ее значением 0 ($i = 0$). Принять в качестве текущего Φ -узла F_c ближайший к серверу Φ -узел ($F_c = M_F[0]$);

3) для Φ -узла F_c вычислить количество НП с помощью процедуры ВКНП. Если число НП меньше требуемой степени связности, добавить новые НП с помощью алгоритма ПРМИТУ: ввести связи F_c с ближайшим узлом множества Ω_{CY} , в которое входит сервер и Т-узлы, не принадлежащие множеству Т-узлов независимых путей F_c ;

4) если i меньше числа Φ -узлов, то положить $i = i + 1$ и принять в качестве текущего узла следующий ближайший к серверу узел $F_c = M_F[i]$. Повторять шаги 3 и 4 пока i не станет равным числу Φ -узлов.

Вычислительная сложность алгоритма ИРТУ согласно [5] оценивается величиной $O(rn_s(fn_e n_r)')$, n_s — число Φ -узлов; n_r — число Т-узлов; r — требуемая степень связности Φ -узлов; n_e — число ребер в графе; f — суммарный максимальный поток.

На рис. 13 (см. четвертую сторону обложки) показано размещение на начальный граф, показанный на рис. 6, Т-узлов с помощью алгоритма ИРТУ. Здесь степень связности Φ -узлов равна двум ($k = 2$).

Решение, полученное алгоритмом ИРТУ, может быть неоптимальным. Так, некоторые Т-узлы могут быть удалены, при этом требуемая степень связности Φ -узлов не изменится. Для определения Т-узлов, которые могут быть удалены, выполняется следующий алгоритм:

1) создать отсортированный по критерию дальности к серверу массив Т-узлов M_T ;

2) создать переменную i , которая хранит индекс текущего Т-узла из массива M_T и инициализировать ее значением 0 ($i = 0$). Принять в качестве текущего Т-узла T_c — самый дальний от сервера Т-узел ($T_c = M_T[0]$);

3) временно удалить T_c и проверить после этого степени связности Φ -узлов с помощью функции ВКНП. Если без временно удаленного T_c степень связности Φ -узлов сети не уменьшается, то T_c удаляется из сети;

4) если i меньше числа Т-узлов, то положить $i = i + 1$ и принять в качестве текущего Т-узла T_c следующий узел $T_c = M_T[i]$. Повторять шаги 3 и 4, пока i не станет равным числу Т-узлов.

Результаты работы алгоритмов

Результаты работы описываемых в статье алгоритмов продемонстрированы на следующей тестовой задаче: Φ -узлы случайным образом размещаются на объекте квадратной формы, с длиной

Таблица 1
Сравнение неоптимизированных решений алгоритмов базового размещения

Алгоритм	200 Ф-узлов		500 Ф-узлов	
	Число Т-узлов	Время, с	Число Т-узлов	Время, с
ПРТУ	1091	0,069	2713	0,12
ПРМИТУ	248	4,3	458	73
ГРТУ	199	0,30	237	0,61

Таблица 2
Сравнение оптимизированных решений алгоритмов базового размещения

Алгоритм	200 Ф-узлов		500 Ф-узлов	
	Число Т-узлов	Время, с	Число Т-узлов	Время, с
ПРТУ	215	1,8	275	11
ПРМИТУ	145	4,7	223	79
ГРТУ	140	0,69	187	1,9

Таблица 3
Сравнение неоптимизированных решений алгоритма ИРТУ

Степень связности	200 Ф-узлов		500 Ф-узлов	
	Число Т-узлов	Время, с	Число Т-узлов	Время, с
2	279	2,1	387	6,9
3	405	6,3	562	24
4	529	14	742	53

Таблица 4
Сравнение оптимизированных решений алгоритма ИРТУ

Степень связности	200 Ф-узлов		500 Ф-узлов	
	Число Т-узлов	Время, с	Число Т-узлов	Время, с
2	217	14,9	323	59
3	330	69	484	170
4	436	160	665	410

Таблица 5
Результаты работы алгоритма ГАРТУ

Алгоритм	200 Ф-узлов		500 Ф-узлов	
	Число Т-узлов	Время, с	Число Т-узлов	Время, с
ГАРТУ (*)	170	150	450	970

стороны 400 м и площадью, соответственно, 160 000 м². Радиус передачи радиосигнала Ф-узла равен 15 м, а Т-узла — 30 м. В тестовой задаче оценивается число размещенных Т-узлов и время работы алгоритмов размещения.

В табл. 1 и 2 приведены результаты работы алгоритмов базового размещения Т-узлов по количеству размещенных Т-узлов и времени работы алгоритма. В табл. 1 показаны результаты без оптимизации, в табл. 2 результаты с оптимизацией.

В табл. 3, 4 показаны результаты работы алгоритма ИРТУ с различными заданными степенями связности Ф-узлов с сервером.

Результаты работы алгоритма ГАРТУ показаны в табл. 5. Исходные данные для поиска оптимальной структуры СС с помощью алгоритма ГАРТУ:

- уровень связности каждого Ф-узла с сервером должен быть не менее 0,99;
- ограничение по стоимости (число Т-узлов не должно быть меньше числа Ф-узлов).

Заключение

В статье рассмотрены алгоритмы базового и отказоустойчивого размещения Т-узлов, показаны примеры размещения Т-узлов этими алгоритмами, а также даны сравнительные таблицы эффективности работы алгоритмов. Таким образом, рассмотренные в статье алгоритмы размещения Т-узлов можно использовать в системах поддержки проектирования отказоустойчивых сенсорных сетей. Решение о применении того или иного алгоритма зависит от требований проектировщика к СС. Согласно результатам экспериментов можно сделать вывод, что среди рассмотренных алгоритмов базового размещения Т-узлов лучшим является алгоритм ГРТУ. Алгоритм ИРТУ формирует отказоустойчивую структуру СС, имеющую требуемую степень связности Ф-узлов с сервером, в то время как алгоритм ГАРТУ осуществляет построение отказоустойчивой структуры СС, учитывая требования по вероятности связности ее Ф-узлов с сервером и стоимости внедрения.

Список литературы

1. Nitaigour P. Mahalik (Ed.). Sensor networks and configuration fundamentals, standards, platforms, and applications. Springer, 2007. 516 p.
2. Турута Е. Н., Мочалов В. А. Проблемы проектирования отказоустойчивых сенсорных сетей // Труды Московского технического университета связи и информатики. М.: ИД Медиа Паблшер, 2008. Т. 1. С. 128—132.
3. Мочалов В. А., Турута Е. Н. Интеллектуальная процедура выбора отказоустойчивой топологии и компонентов сенсорной сети // Труды Международных научно-технических конференций "Интеллектуальные системы" (AIS'08) и "Интеллектуальные САПР" (CAD-2008). М.: Физматлит, 2008. Т. 1. С. 385—392.
4. Мочалов В. А., Турута Е. Н. Метод построения отказоустойчивой структуры сенсорной сети, основанный на применении генетического алгоритма // Труды Московского технического университета связи и информатики. М.: МТУСИ, 2009.
5. Ahlberg M., Vlassov V., Yasui T. Router Placement in Wireless Sensor Network. // Mobile Adhoc and Sensor Systems (MASS), 2006. IEEE International Conference. Oct. 2006. P. 538—541.
6. Goodrich M.-T., Tamassia R. Algorithm Design: Foundations, Analysis, and Internet Examples. John Wiley & Sons, 2002.
7. Форд Л., Фалкерсон Д. Поток в сетях: Пер. с англ. М.: Мир, 1966. 276 с.
8. Гладков Л. А., Курейчик В. В., Курейчик В. М. Генетические алгоритмы: учебное пособие. 2-е изд. М.: Физматлит, 2006.

В. А. Огнев, аспирант,
С. Р. Иванов, канд. техн. наук, доц.,
 МГТУ им. Н. Э. Баумана,
 e-mail: smarserg@mtu.-nt.ru

Математическая модель приемовычислителя систем ГЛОНАСС/GPS

Рассмотрена структура и приведена упрощенная математическая модель приемовычислителя, работающего по сигналам спутниковых навигационных систем ГЛОНАСС/GPS, позволяющая оценить влияние на его работу гауссовых помех. Показано, что помехоустойчивость приемовычислителя определяется схемами слежения, осуществляющими оценку информационных параметров сигналов. Проведено сравнение различных вариантов построения следящих схем. Представлены результаты имитационного моделирования.

Ключевые слова: спутниковая навигационная система, приемовычислитель, помехоустойчивость, гауссова помеха, следящая схема, соотношение сигнал/шум.

В настоящее время широкое распространение получила технология определения координат подвижных объектов по сигналам спутниковых навигационных систем (СНС). Для определения координат объектов используются устройства, называемые приемовычислителями (ПВ) СНС, которые в настоящее время являются частью множества информационно-навигационных систем. В связи с этим актуальными становятся вопросы анализа работы ПВ в условиях действия помех различного происхождения (индустриальных или преднамеренных), затрудняющих выделение полезной информации из сигналов.

Данная статья отражает результаты работы по построению математической модели сигналов СНС и приемовычислителя, работающего в режиме сопровождения сигналов навигационных космических аппаратов (НКА). Данная модель синтезируется для изучения влияния широкополосной помехи (белого гауссова шума) на работу ПВ. При этом преследуются следующие цели: выявить факторы, определяющие помехоустойчивость ПВ, и определить граничное значение соотношения сигнал/шум, при котором еще происходит выделение информации.

На вход ПВ поступает сумма сигналов $\sum_{i=1}^n S_i$ от n источников навигационных сигналов. Для описания сигналов стандартной точности, излучаемых

НКА СНС ГЛОНАСС [1] и GPS [2] используется следующая модель [3]:

$$S_i(t) = \sqrt{2P_i} G_i(t - \tau_i) D_i(t - \tau_i) \times \\ \times \cos((\omega_{0i} + 2\pi f_{\text{доп}i})t + \varphi_{0i})$$

где P_i — мощность сигнала; $G_i(t - \tau_i)$ — функция модуляции дальномерным кодом — псевдослучайной последовательностью (ПСП), принимающей значения -1 или 1 ; τ_i — начальная фаза ПСП (задержка распространения сигнала от НКА до ПВ); $D_i(t - \tau_i)$ — функция модуляции навигационным сообщением (принимает значения -1 или 1 , которые, в общем случае, равновероятны [3]), длительность бита $T = 20$ мс); $\omega_{0i} = 2\pi f_{0i}$ — несущая частота; $f_{\text{доп}i}$ — доплеровский сдвиг частоты ($f_{\text{доп}i} = -5...+5$ кГц); φ_{0i} — начальная фаза сигнала.

Из теории распространения радиоволн известно, что на основе радионавигационных параметров τ_i и $f_{\text{доп}i}$ можно вычислить расстояние от НКА до ПВ $PR = ct$ (c — скорость света) и взаимную скорость НКА и ПВ $PRR = f_{\text{доп}i} \lambda$ (λ — длина волны). Рассчитав для каждой точки траектории объекта PR_i и PRR_i и пересчитав их в соответствующие τ_i и $f_{\text{доп}i}$, можно учитывать влияние динамики движения ПВ на параметры принимаемого сигнала:

$$S_i(t) = \sqrt{2P_i} G_i\left(t - \frac{PR_i}{c}\right) D_i\left(t - \frac{PR_i}{c}\right) \times \\ \times \cos\left(\left(\omega_{0i} + 2\pi \frac{PRR_i}{\lambda_i}\right)t + \varphi_{0i}\right).$$

И наоборот, измеряя в ПВ τ_i и $f_{\text{доп}i}$ для $i = 1, \dots, n$, можно вычислить набор PR_i (что позволяет вычислить пространственные координаты ПВ) и PRR_i (что позволяет определить скорость движения объекта). Таким образом, в основе функционирования ПВ лежит процесс измерения параметров сигнала τ_i и $f_{\text{доп}i}$.

С точки зрения функционирования ПВ различают три основных режима:

- обнаружение (т. е. грубая оценка параметров τ_i и $f_{\text{доп}i}$) сигналов НКА (в рамках данной статьи этот режим не исследуется);
- слежение за сигналом (некогерентный режим работы);
- прием навигационной информации (когерентный режим работы).

В некогерентном режиме работы ПВ оценивается только скорость изменения фазы (частота) принимаемого сигнала (т. е. проводится усреднение по начальной фазе сигнала). В когерентном

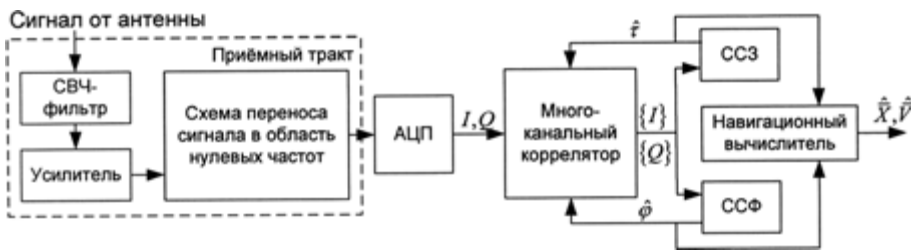


Рис. 1. Структурная схема ПВ с двухэтапной обработкой сигналов НКА

режиме проводится точная оценка мгновенной фазы сигнала.

Таким образом, вектор информационных параметров сигнала в некогерентном режиме $\chi_i = \{\tau_i, f_{\text{доп}i}\}$, а в когерентном $\chi_i = \{\tau_i, \phi_i\}$, где $\phi_i(t) = 2\pi f_{\text{доп}i}t + \phi_{0i}$.

В ПВ с классической двухэтапной обработкой сигналов (рис. 1) на первом этапе осуществляется выделение информационных параметров $\tau_i, f_{\text{доп}i}$ для всего набора $\{S_i\}$, а на втором этапе, на базе множества рассчитанных $\{PR_i, PRR_i\}$ проводится оценка векторов текущих координат X и скоростей V . Таким образом, способность ПВ вычислять пространственное положение потребителя в присутствии помех определяется помехоустойчивостью схем оценки информационных параметров χ_i .

Элементы приемного тракта ПВ одинаково преобразуют полезный сигнал и помехи, попадающие в полосу приема. Влияние приемного тракта на помехоустойчивость ПВ (которая характеризуется отношением сигнал/шум) заключается в генерировании теплового (гауссового) шума со спектральной плотностью:

$$N_0 = kT_0K_{\text{ш}},$$

где $k = 1,38 \cdot 10^{-23}$ Вт · с/К — постоянная Больцмана; T_0 — температура окружающей среды; $K_{\text{ш}}$ — коэффициент шума приемного тракта приемника.

Для современных ПВ легко достижим $K_{\text{ш}} = 2$ (т. е. $10\log(2) = 3$ дБ), что при $T_0 = 293$ К дает $N_0 = 1,21 \cdot 10^{-20}$ Вт/Гц. В то же время мощность полезного сигнала [1, 2] не превышает $P_c = -160$ дБ · Вт

(т. е. $10^{-160/10} = 10^{-16}$ Вт). Отметим, что при этом в полосе полезного сигнала $B_c = 1$ МГц (ширина спектра сигнала СТ ГЛОНАСС) мощность теплового шума будет составлять $P_T = N_0B = 1,21 \cdot 10^{-20}$ Вт/Гц · 10^6 Гц = $1,21 \cdot 10^{-14}$ Вт (-130 дБ/Вт), т. е. на три порядка (30 дБ) превышает мощность полезного сигнала. Как видно, на выходе приемного тракта соотношение сигнал/шум очень мало и для выделения информационных параметров необходимо проводить дополнительную обработку сигнала.

Согласно работе [4] при использовании АЦП с числом разрядов $k \geq 8$ и генератора опорной частоты с относительной нестабильностью менее 10^{-7} (что имеет место при построении современных ПВ) влиянием ошибок квантования и апертурной погрешности АЦП, формирующего синфазную I и квадратурную Q компоненты, на качество работы ПВ можно пренебречь.

Тогда, отношение сигнал/шум на выходе АЦП $SNR_{\text{АЦП}} = \frac{P_c}{N_0B + P_{\text{п}}}$, где $P_{\text{п}}$ — мощность помехи.

В случае гауссовой помехи (белый шум со спектральной плотностью $N_{\text{п}}$) с учетом того, что помеха не коррелирована с тепловым шумом приемного тракта, $SNR_{\text{АЦП}} = \frac{P_c}{(N_0 + N_{\text{п}})B}$.

В каждом канале коррелятора проводится вычисление набора корреляционных интегралов для одного из НКА:

$$I_P(t) = \int_{t-T}^t I(t)G_i(t - \bar{\tau})\cos(2\pi\bar{f}t + \varphi),$$

$$Q_P(t) = \int_{t-T}^t Q(t)G_i(t - \bar{\tau})\sin(2\pi\bar{f}t + \varphi);$$

$$I_E(t) = \int_{t-T}^t I(t)G_i(t - \bar{\tau} - d)\cos(2\pi\bar{f}t + \varphi),$$

$$Q_E(t) = \int_{t-T}^t Q(t)G_i(t - \bar{\tau} - d)\sin(2\pi\bar{f}t + \varphi);$$

$$I_L(t) = \int_{t-T}^t I(t)G_i(t - \bar{\tau} + d)\cos(2\pi\bar{f}t + \varphi),$$

$$Q_L(t) = \int_{t-T}^t Q(t)G_i(t - \bar{\tau} + d)\sin(2\pi\bar{f}t + \varphi),$$

где $\bar{\tau}, \bar{f}$ — оценки задержки и остаточной (доплеровской) частоты входного сигнала; d — временной сдвиг между локально генерируемыми копиями опорного сигнала, $d < \frac{\tau_{\text{и}}}{2}$, где $\tau_{\text{и}}$ — длительность элементарного элемента ПСП.

Согласно работе [5] корреляционная обработка повышает устойчивость к широкополосной гауссовой помехе, обеспечивая прирост отношения сигнал/шум на $\Delta SNR = 10\log_{10}(B_cT_c)$ [дБ], где B_c — ширина спектра сигнала ($B_c \approx 1$ МГц для сигнала ГЛОНАСС СТ и $B_c \approx 2$ МГц — для GPS

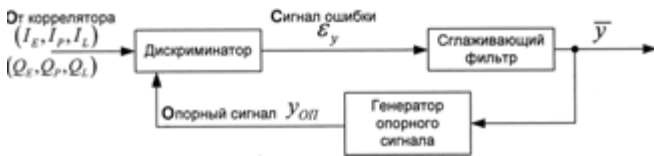


Рис. 2. Обобщенная структура следящих схем, используемых в ПВ

Таблица 1

Сравнительные данные по дисперсии шума на выходе дискриминатора ССЗ при различных реализациях этого дискриминатора

Вид следящей схемы	Формула для вычисления значения сигнала рассогласования на выходе дискриминатора	Дисперсия шума на выходе дискриминатора
ССЗ (когерентный режим)	$D1 = I_E - I_L$	$\sigma_{\tau}^2 \approx \frac{B_L d}{2SNR}$
ССЗ (некогерентный режим)	$D2 = (I_E^2 + Q_E^2) - (I_L^2 + Q_L^2)$	$\sigma_{\tau}^2 \approx \frac{B_L d}{2SNR} \left(1 + \frac{2}{(2-d)SNR \cdot T_1} \right)$
ССЗ (некогерентный режим)	$D3 = I_P(I_E - I_L) + Q_P(Q_E - Q_L)$	$\sigma_{\tau}^2 \approx \frac{B_L d}{2SNR} \left(1 + \frac{1}{SNR \cdot T_1} \right)$

Примечание: SNR — отношение сигнал/шум на входе следящей схемы; B_L — полоса пропускания системы по отношению к шуму (определяется ожидаемым диапазоном изменения оцениваемого информационного параметра); T_1 — время, в течение которого вычисляются корреляционные интегралы.

С/А); T_c — длительность (период повторения) сигнала (для сигналов ГЛОНАСС СТ и GPS С/А $T_c = 1$ мс). Соответственно, на выходе коррелятора обеспечивается прирост отношения сигнал/шум на 33 дБ для сигнала GPS С/А, и на 30 дБ для сигнала ГЛОНАСС СТ.

Для дальнейшего повышения отношения сигнал/шум значения корреляционных интегралов, вычисленные на миллисекундных интервалах, также можно суммировать, что эквивалентно увеличению параметра T_c и, следовательно, ведет к увеличению отношения сигнал/шум в M раз, где M — число суммирования. Однако поскольку изменение значения бита навигационного сообщения $D(t)$ приводит к смене знака сигнала, очевидно, что максимальное время суммирования ограничено длительностью бита навигационного сообщения, которая для сигналов ГЛОНАСС СТ и GPS С/А равна 20 мс (т. е. $M \leq 20$).

В работе [3] показано, что оптимальные оценки параметров τ , f должны формироваться параллельно работающими следящими схемами, обобщенная структура которых показана на рис. 2. В этой схеме под параметром y может пониматься как τ (в схеме слежения за задержкой (ССЗ)), так и f (в схемах слежения за частотой (ССЧ)) или φ

(в схеме слежения за фазой (ССФ)). Именно качество работы этих схем (а именно, минимально допустимое соотношение сигнал/шум на входе, при котором не происходит срыва слежения за информационным параметром y) определяет помехоустойчивость ПВ.

В табл. 1 приведены выражения для вычисления значения сигнала рассогласования на выходе дискриминатора петли ССЗ. Как видно, в дискриминаторе корреляционные интегралы подвергаются нелинейной обработке. Вследствие этого прямой анализ работы данной схемы аналитическими методами затруднен, такая имитационная модель ССЗ также имела бы высокую вычислительную сложность. В работе [6] показано, что при воздействии на вход ПВ аддитивного белого гауссова шума, шум на выходе дискриминатора — также гауссов с математическим ожиданием, равным нулю, и дисперсией, указанной в табл. 1. Поэтому, используя линейную аппроксимацию, представим сигнал рассогласования на выходе дискриминатора ССЗ в виде $\varepsilon_{\tau} = \tau + v_{\tau}$ (рис. 3), где τ — идеальное текущее значение задержки сигнала, вычисленное из модели динамики объекта; v_{τ} — шум (с нулевым математическим ожиданием и дисперсией σ_{τ}^2) на выходе дискриминатора.

Структуры ССФ и ССЧ идентичны, поскольку фаза $\varphi(t)$ и частота $f_{\text{доп}}(t)$ математически однозначно связаны между собой (частота — производная от фазы, $f(t) = \varphi'(t)$). Дискриминатор в ССФ осуществляет оценивание текущего рассогласования по фазе между приходящим и опорным сигналами, а дискриминатор ССЧ оценивает соответствующее рассогласование по частоте. Поэтому далее будем рассматривать построение моделей ССФ и ССЧ на примере ССФ.

Если строить математическую модель ССФ в строгом соответствии с рис. 2, то модель ССФ будет иметь вид, представленный на рис. 4. Здесь квадратурные отсчеты входного сигнала $I_{\text{вх}}(\varphi, \sigma_{\text{ш}})$ и $Q_{\text{вх}}(\varphi, \sigma_{\text{ш}})$ зависят как от текущей фазы сигнала φ (которая, в свою очередь, зависит от параметров дви-



Рис. 3. Модель ССЗ при линейной аппроксимации сигнала на выходе дискриминатора



Рис. 4. Структура неоптимизированной модели схемы слежения за фазой сигнала

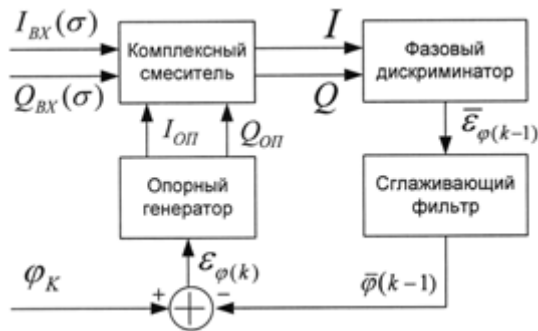


Рис. 5. Структура оптимизированной модели ССФ

Таблица 2

Сравнительные данные по дисперсии шума на выходе дискриминаторов ССЧ и ССФ

Вид следящей схемы	Формула для вычисления значения сигнала рассогласования на выходе дискриминатора
ССЧ	$D4_k = I_{k-1}Q_k + I_kQ_{k-1}$
ССФ	$D5 = Q\text{sign}(I)$
ССФ	$D6 = \arctg \frac{Q}{I}$

жения объекта: $\varphi_i(t_0 + \Delta t) = \varphi(t_0) + \int_{t_0}^{t_0 + \Delta t} \frac{PRR_i(t)}{\lambda} dt$,

так и от среднего квадратического отклонения аддитивного белого шума $\sigma_{ш}$, отражающего воздействие на ПВ широкополосной помехи. Для упрощения модели целесообразно отдельно учитывать влияние этих факторов, т. е. отдельно вводить их в модель, что обеспечивает, например, модель ССФ, представленная на рис. 5, где одним из входных параметров является идеальная (т. е. сформированная только на основе модели движения объекта) фаза входного сигнала $\varphi(k)$.

В оптимизированной модели опорный генератор формирует не полные копии отсчетов входного сигнала, а только сигнал рассогласования. На комплексный смеситель подаются значения $I_{ВХ}(\sigma_{ш})$, $Q_{ВХ}(\sigma_{ш})$, сформированные с учетом аддитивного наложения шумовой помехи с задаваемой

дисперсией $\sigma_{ш}^2$. Влияние перемещения объекта на параметры принимаемого сигнала учитываются путем введения в модель идеальной (незашумленной) текущей фазы сигнала φ_K , которая зависит только от динамики объекта: $\varphi_i(t_0 + \Delta t) =$

$$= \varphi(t_0) + \int_{t_0}^{t_0 + \Delta t} \frac{PRR_i(t)}{\lambda} dt.$$

Сглаживающие фильтры ССЗ, ССЧ, ССФ представляют собой, как правило, фильтры нижних частот 3-го порядка и эквивалентны друг другу (см. работу [6]). Структура сглаживающего фильтра для ССФ представлена на рис. 6.

Математическая модель такого фильтра:

$$d2\text{phase}(n + 1) = d2\text{phase}(n) + k_3\varepsilon(\varphi_{n + 1})$$

$$d\text{phase}(n + 1) = d\text{phase}(n) + d2\text{phase}(n + 1) + k_2\varepsilon(\varphi_{n + 1}) + k_1(\varepsilon(\varphi_{n + 1}) - \varepsilon(\varphi_n))$$

$$\text{phase}(n + 1) = \text{phase}(n) + d\text{phase}(n + 1).$$

Член $k_1(\varepsilon(\tau_n) - \varepsilon(\tau_{n-1}))$ необходим для устранения расходимости фильтра, которая возникает при последовательном включении двух интегрирующих звеньев. Коэффициенты фильтра k_1 , k_2 , k_3 выбираются в зависимости от шага дискретизации по времени и требуемой полосы пропускания следящей схемы (см. [7]).

Приведенные соотношения использованы для построения имитационной модели ПВ в среде MATLAB 2007.

Моделирование работы ПВ проводилось при следующей динамике объекта: начальная скорость объекта — 20 м/с; ускорение — 2g (B_L установлена в 5 Гц); $d = 100$ мкс.

По результатам моделирования получена зависимость средней квадратической ошибки (СКО) определения скорости PRR и дальности PR от эффективного отношения сигнал-шум SNR при работе ССФ (дискриминатор D6), ССЧ (дискриминатор D4) и ССЗ (дискриминатор D3). Варьируя СКО внешней шумовой помехи, можно определить

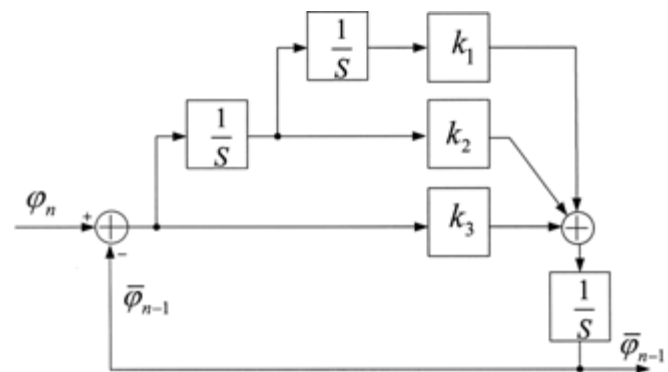


Рис. 6. Структура сглаживающего фильтра ССФ

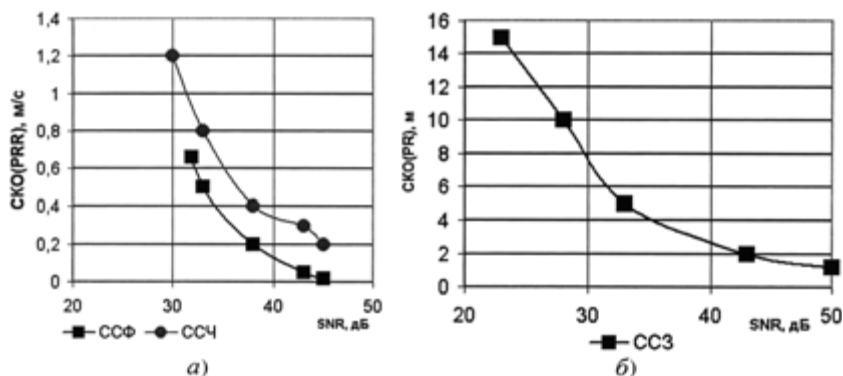


Рис. 7. Результаты моделирования работы ССФ, ССЧ (а) и ССЗ (б)

граничное значение отношения сигнал/шум, при котором происходит срыв слежения за соответствующим информационным параметром.

Эффективное отношение сигнал/шум вычислено по формуле $SNR = \frac{Am}{\sigma_{ш} \sqrt{2}}$, где Am — амплитуда сигнала (принята равной 1).

Как видно из рис. 7, граничное значение соотношения сигнал/шум для функционирования ССФ равно 33 дБ, для ССЧ = 30 дБ, для ССЗ = 23 дБ.

Выводы

1. Как было показано, значение выигрыша в отношении сигнал/шум в результате накопления сигнала в корреляторе ограничивается длительностью бита навигационного сообщения. Поэтому можно заключить, что помехоустойчивость ПВ определяется структурой используемых при работе сигналов (длиной ПСП и длительностью бита навигационного сообщения) и качеством работы схем оценивания информационных параметров сигналов — ССЗ и ССЧ (ССФ).

Использование при работе сигналов с более длинной ПСП и немодулированных навигационным сообщением $D(t)$ должно вести к повышению помехоустойчивости ПВ.

2. Анализ результатов моделирования (см. рис. 7) показывает, что в некогерентном режиме работы помехоустойчивость ПВ выше, чем в когерентном, т. е. ССЧ имеет больший запас помехоустойчивости, чем ССФ, однако точность оценивания скорости PRR при работе ССЧ ниже.

Поскольку декодирование навигационного сообщения возможно только в когерентном режиме (при работе ССФ), а ССФ является наименее помехоустойчивой следящей системой ПВ, при наличии помех в первую очередь нарушается прием навигационных сообщений от НКА.

3. Построенная имитационная модель адекватно (с качественных позиций) описывает процес-

сы, происходящие в ПВ при первичной обработке сигналов СНС (выделении информационных параметров сигналов). Полученные результаты в целом совпадают с результатами других исследователей [3, 8].

Итак, в результате проведенных исследований получена упрощенная модель приемовычислителя СНС, которая позволяет изучить влияние одного из видов помех — гауссовых — на приемовычислитель, работающий по существующим в настоящее время сигналам СНС ГЛОНАСС/GPS.

Разработанная модель может стать ядром перспективного программно-моделирующего комплекса для анализа помехоустойчивости ПВ. Для этого ее в дальнейшем целесообразно модифицировать в целях обеспечения возможности оценки влияния на ПВ не только широкополосной шумовой, но и других (узкополосных, импульсных) типов помех.

Поскольку ожидается, что в результате проводящейся модернизации СНС ГЛОНАСС/GPS и развертывания новой европейской СНС GALILEO в 2009...2012 гг. в эксплуатацию будут введены сигналы СНС с новой структурой (с более длинными ПСП и компонентами, свободными от модуляции навигационным сообщением), является актуальным внесение необходимых изменений в разработанную модель, чтобы с ее помощью провести оценку потенциальной помехоустойчивости перспективных ПВ, использующих при работе новые сигналы СНС.

Список литературы

1. Глобальная навигационная спутниковая система ГЛОНАСС. Интерфейсный контрольный документ. Редакция 5.0. Москва, 2002 г.
2. ICD-GPS-200D-001. NAVSTAR GLOBAL POSITIONING SYSTEM. INTERFACE SPECIFICATION. Navstar GPS Space Segment / Navigation User Interfaces. March 2006.
3. Перов А. И., Харисов В. Н. ГЛОНАСС. Принципы построения и функционирования. Изд. 3-е, перераб. М.: Радиотехника, 2005. 688 с.
4. Cutright C., Burns J. R., Braasch M. S. Characterization of Interference Mitigation Performance versus Quantization Error in Software Radios // Proc. of conference. "ION Annual Meeting 2003", USA, 2003. P. 323—332.
5. Варакин Л. Е. Системы связи с шумоподобными сигналами. М.: Радио и связь, 1985. 384 с.
6. Bradford W. Parkinson, James J. Spilker Jr. Global Positioning System: Theory and Applications. Vol. I. American Institute of Aeronautics and Astronautics, Inc. 1996. 620 p.
7. Stephens S. A., Thomas J. B. Controlled-Root Formulation for Digital Phase-Locked Loops // IEEE Transactions on Aerospace and Electronic Systems. 1995. Vol. 31. N 1.
8. Перов А. И., Болденков Е. Н., Григоренко Д. А. Упрощенная аналитическая методика оценки потенциальной помехоустойчивости оптимальных следящих систем приемников спутниковой навигации // Радиотехника. 2003. № 7. С. 79—87.

УДК 658.512.011.56

А. А. Петунин, канд. техн. наук, проф.,
Государственное образовательное учреждение
высшего профессионального образования
"Уральский государственный
технический университет —
УПИ им. первого Президента России Б. Н. Ельцина",
г. Екатеринбург,
e-mail: sansan_2007@mail.ru

Автоматизация нерегулярного фигурного раскроя: основные подходы к разработке систем проектирования раскроя и управляющих программ для машин резки листовых материалов

Описывается структура программного обеспечения современной CAD/CAM-(Computer-Aided-Design/Computer-Aided Manufacturing) системы, предназначенной для автоматизированного проектирования раскроя листовых материалов на заготовки сложной геометрической формы и для подготовки управляющих программ технологического оборудования с числовым программным управлением, которое использует термическую или гидроабразивную технологию резки материала. Излагается новый подход к созданию подсистемы автоматического нерегулярного фигурного раскроя. Формулируются и обосновываются требования, предъявляемые к основным модулям системы.

Ключевые слова: CAD/CAM-система, нерегулярный фигурный раскрой, резка материала, подсистема САПР, машины с ЧПУ, управляющая программа.

Введение

В машиностроении и других отраслях промышленности значительная часть продукции изготавливается из заготовок, получаемых из листовых материалов на технологическом оборудовании с числовым программным управлением (ЧПУ) для лазерной, плазменной, газовой, электроэрозионной и гидроабразивной резки материала. В условиях снижения серийности производства, которое в настоящее время характерно для большинства промышленных предприятий и в России, и за рубежом, именно ма-

шины с ЧПУ становятся основным видом станков, удовлетворяющих требованию максимальной простоты адаптации к постоянным изменениям номенклатуры получаемых заготовок. Следует отметить, что процесс "индивидуализации" производства никак не связан с сегодняшним глобальным финансовым кризисом. Как известно, использование систем автоматизированного проектирования (САПР) управляющих программ (УП) для станков с ЧПУ (САМ-систем — *Computer-Aided Manufacturing*) обеспечивает значительное сокращение времени подготовки программ в сравнении с "ручным" проектированием. Проектирование УП для машин термической и гидроабразивной резки предполагает предварительное получение раскроя листового материала. Уже не одно десятилетие во многих странах Европы и в США разрабатывалось математическое и программное обеспечение САПР для решения задачи оптимизации раскроя материала на заготовки сложной (непрямоугольной) геометрической формы в единичном производстве. Это одна из задач раскроя-упаковки (Cutting & Packing, C&P) [1], известная как проблема нерегулярного фигурного раскроя или проблема "нестинга". В общем случае под "нестингом" понимается оптимальное размещение геометрических объектов сложной формы в заданных областях, а программное обеспечение, автоматизирующее процесс "нестинга", принято относить к CAD (*Computer-Aided Design*)-модулям.

Параллельно с разработками в области оптимизации раскроя материала развивались средства автоматизации проектирования УП для машин с ЧПУ. Созданные на основе этих разработок CAD/CAM-системы различаются между собой не только применяемыми математическими моделями, методами и алгоритмами, но также и различной степенью универсальности, которая касается как широты охватываемых задач, так и спектра технологического оборудования, поддерживаемого системой. Если говорить о зарубежных САПР, то на российском рынке сейчас в основном преобладают узкоспециализированные CAD/CAM-системы, приобретаемые предприятиями вместе с конкретной машиной с ЧПУ (примером может служить программное обеспечение *ToPs*, поставляемое с оборудованием германской фирмы *Trumpf*, или система *Columbus*, предлагаемая покупателям машин для термической резки материала шведской компании *ESAB*). Такие системы, как правило, имеют хороший

CAM-модуль, разработанный специально для данного типа машин, и позволяют также эффективно решать определенный (хотя и весьма ограниченный) круг задач раскроя. Вместе с тем, применение этих САПР вне рамок их специализации чаще всего бывает нецелесообразно, либо невозможно. Прежде всего, набор постпроцессоров, поставляемых с оборудованием, чаще всего ограничен списком этого оборудования. Кроме того, сами функции САМ-модуля для автоматизации проектирования УП не позволяют обычно расширить применение системы для другого оборудования. Например, программное обеспечение для подготовки программ лазерной резки может не учитывать особенности плазменной резки или наличие у покупателя машины с ЧПУ с трехрезовым блоком для одновременной резки заготовки и снятия фасок под сварку и т. д.

Ясно, что требования универсальности и эффективности являются взаимно противоречивыми: высокоэффективной может быть только специализированная система. Поиск "золотой середины" между универсальностью и эффективностью остается в настоящее время одной из ключевых проблем для разработчиков САПР раскроя.

Российские разработчики, работающие в этой предметной области, могут похвастаться серьезными успехами в теоретических исследованиях вопросов оптимизации для многообразных задач раскроя-упаковки [2—6]. Гораздо более скромны их достижения в теории и практике создания эффективных САПР раскроя промышленных материалов, включающих в себя средства подготовки УП для технологического оборудования с ЧПУ, использующего термическую и гидроабразивную технологию резки листовых материалов. В данной статье рассматриваются вопросы, связанные со структурой современной CAD/CAM-системы оптимального раскроя материала в единичном производстве, ориентированной на использование в различных отраслях промышленности и на максимально широкий спектр технологического оборудования с ЧПУ для производства заготовок из листа.

Состав и особенности программного обеспечения САПР оптимального раскроя материала на заготовки сложной формы

CAD/CAM система, предназначенная для автоматизированного проектирования нерегулярного раскроя, содержит два основных модуля: собственно модуль раскроя и САМ-модуль для последующего проектирования УП для машины с ЧПУ. Кроме того, САПР нерегулярного раскроя должна включать программные средства, которые реализуют функции описания геометрической информации для двумерных объектов и импорта

такого рода информации из других CAD-систем, а также функции формирования задания на раскрой и верификации УП программы, полученной в САМ-модуле.

Каждый из перечисленных модулей и каждую из этих функций естественно реализовать в виде отдельной самостоятельной подсистемы САПР. Рассмотрим некоторые особенности указанных подсистем.

Подсистема описания и визуализации геометрической информации. Один из главных вопросов при разработке подсистемы описания геометрической информации заключается в том, насколько целесообразно для этих целей использовать существующие широкораспространенные универсальные CAD-системы типа AutoCAD (Autodesk) или T-Flex CAD (Топ Системы), либо ориентироваться на разработку собственного специализированного модуля, который не будет иметь очевидную избыточность функций, характерную для универсальных систем. Отметим, что в настоящее время при разработке специализированных САПР применяют как первый [7, 8], так и второй подход [9, 10]. Первый подход характерен для разработок, в которых универсальная CAD-система применяется не только как подсистема описания геометрических объектов, но и как базовая графическая среда, интегрированная со многими специализированными внешними модулями, в которых используются функции геометрического моделирования и визуализации. Наличие в CAD-системах соответствующего инструмента API (*Application Program Interface*), который ориентирован на создание пользовательских программных модулей, интегрированных с базовым программным продуктом, может существенно сократить время разработки специализированной САПР раскроя материала. Еще одно преимущество использования такого подхода для разработчика приложений заключается в уменьшении эффекта "старения" созданного прикладного программного обеспечения, так как часть решения проблемы "старения" берет на себя разработчик универсальной САПР, который в условиях жесткой конкуренции вынужден постоянно обновлять версии своего программного продукта. Пожалуй, наиболее популярной в настоящее время технологией создания API становится NET-технология, разработанная компанией *Microsoft*. Стоит отметить, что данная технология не является единственной, позволяющей интегрировать внешние модули в систему. Большинство универсальных CAD-систем также поддерживают СОМ (*Component Object Model*)-технология, позволяющую внедрять внешние объекты в приложение, но ввиду сложности разработки СОМ-компонентов данная технология не получила широкого распространения.

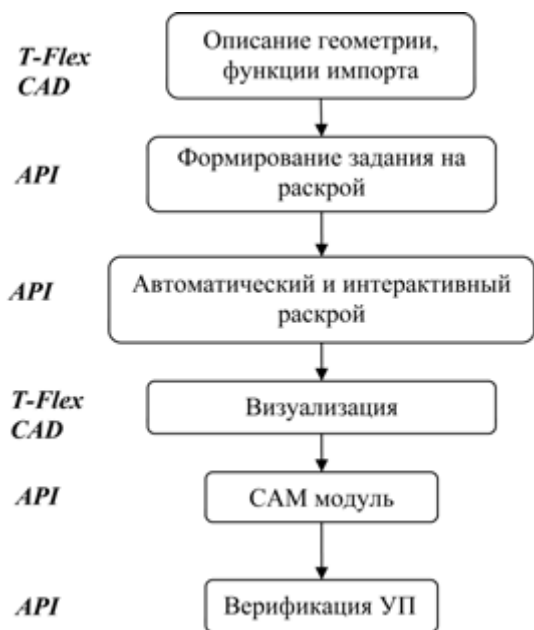


Рис. 1. Состав подсистем САПР нерегулярного фигурного раскрой

Если при разработке подсистемы описания геометрии и других модулей используется, например, API российской системы T-Flex CAD, то общую структурную схему подсистем САПР раскрой и подготовки УП для машин резки материала можно представить следующим образом (рис. 1)

Что касается CAD/CAM-систем раскрой, в которых все подсистемы реализованы без использования универсальных CAD-систем, то их доля, как нам представляется, будет постепенно уменьшаться.

В любом случае одним из основных требований, предъявляемых к подсистеме описания геометрической информации, помимо удобства геометрического моделирования двумерных геометрических объектов, является наличие функций импорта-экспорта геометрической информации, которые должны обеспечивать взаимодействие с другими CAD-системами посредством стандартных геометрических форматов DXF, DWG, STL и др.

Подсистема формирования задания на раскрой. Подсистема формирования задания помимо основной функции формирования списка и количества заготовок, подлежащих размещению на раскраиваемом материале, должна содержать и необходимые функции работы с базой данных заготовок. К ним следует отнести следующие:

- возможность выборки из базы данных заготовок по их приоритету, а также по марке и толщине материала;
- визуализация геометрии заготовок и раскраиваемого материала;
- вычисление суммы площадей отобранных в задании заготовок и раскраиваемого материала;

- контекстный поиск по наименованию;
- предварительный автоматический раскрой.

Особо отметим необходимость в подсистеме формирования задания последней функции, поскольку именно она позволяет осуществлять адекватный выбор заготовок в задание и во многом может способствовать замене "ручного" предварительного планирования раскрой, которое существует сегодня на большинстве предприятий с единичным типом производства. Использование на этих предприятиях специальных подразделений, выполняющих функции предварительной "прикидки" "вручную", обусловлено, в первую очередь, отсутствием в имеющихся САПР раскрой эффективного инструмента формирования задания на раскрой, который бы учитывал множество разных факторов и позволял за короткое время получать и оценивать предварительные варианты раскрой еще до процесса проектирования окончательного варианта.

Подсистема проектирования раскрой. Современный опыт разработки и использования САПР раскрой материала показывает, что ключевая для САПР нерегулярного фигурного раскрой подсистема, реализующая функции "нестинга", структурно должна состоять из трех модулей:

- модуля автоматического раскрой;
- модуля автоматического выбора алгоритма раскрой из числа имеющихся и определения параметров алгоритма;
- модуля интерактивного раскрой.

Известно, что задача "нестинга" относится к классу *NP*-трудных задач, для которых не существует алгоритмов решения полиномиальной сложности. Опыт разработки приближенных алгоритмов решения, реализуемых в виде специализированного программного обеспечения автоматического раскрой, выявил еще одну особенность этой задачи: эффективность созданных алгоритмов существенно зависит от условий задачи: геометрической формы деталей, их ассортимента, количественных характеристик, размеров материала и пр., причем характер этой зависимости остается крайне мало исследованным. На возможность применения того или иного алгоритма влияют и особенности оборудования, используемого для раскрой (например, гильотинные ножницы или машина с ЧПУ для плазменной резки материала). Следует отметить, что практически все приближенные алгоритмы имеют целый набор числовых параметров, которые задают не только условия задачи (например, значение припуска на рез или свойство анизотропности материала), но и особенности применяемого метода оптимизации (например, ограничение на глубину рекурсии для рекурсивного алгоритма и временные характеристики). Это означает, что фактически пользователь

имеет дело не с одним алгоритмом, а с целым спектром алгоритмов. Несмотря на то, что результат минимального изменения параметров алгоритма часто приводит к значительному изменению результата оптимизации, метод научного "тыка" часто остается единственным методом, который может посоветовать пользователю при выборе параметров автор алгоритма. Если же пользователь имеет возможность применения нескольких принципиально различных алгоритмов раскроя, то проблема выбора наиболее эффективного из них превращается в неразрешимую задачу. В этом случае наличие в подсистеме раскроя модуля, который бы в автоматическом режиме определял наиболее "перспективный" алгоритм и его параметры, становится необходимостью. Разработка такого рода модулей должна, естественно, проводиться на основе статистического анализа применения алгоритмов при решении различных типов задач раскроя. Одна из таких работ описана, в частности, в [11]. При этом модуль автоматического выбора алгоритма и его параметров должен сам иметь ряд управляющих параметров, задаваемых пользователем.

Еще одна особенность задач нерегулярного фигурного раскроя заключается в том, что для определенного класса таких задач достаточно эффективно работает эвристический прием формирования групп заготовок (или *bits*), имеющих в задании на раскрой. Под *bit* в этом случае понимается набор заготовок, положения которых зафиксированы относительно друг друга. При размещении такой *bit* ведет себя как одна заготовка, т. е. все преобразования по перемещению/вращению проводятся одновременно со всеми заготовками, входящими в группу. Группы (*bits*), в основном, состояются из однотипных заготовок, но могут содержать и заготовки различной конфигурации. Например, все одинаковые прямоугольные треугольники целесообразно объединять парами в группу, имеющую форму прямоугольника, с размерами, равными катетам треугольника, а крупное внутреннее отверстие какой-либо заготовки удобно предварительно плотно "заполнить" мелкими заготовками.

Таким образом, одна из функций модуля автоматического выбора алгоритма должна заключаться в анализе геометрии заготовок, входящих в задание на раскрой и формировании соответствующих *bits*. При этом, применение (или неприменение) механизма формирования групп должно определяться пользовательским параметром, так же как и, естественно, использование самого модуля выбора автоматического раскроя.

Поскольку (в силу *NP*-сложности задачи) ни один из алгоритмов решения не может гарантировать получение глобального "оптимума", то ин-

терактивный метод проектирования раскроя материала в ряде случаев становится единственным способом получения удовлетворяющего пользователя варианта решения. Кроме того, интерактивная корректировка результата автоматического раскроя необходима, когда полученный компьютером результат неудовлетворителен вследствие технологических особенностей вырезки заготовок, например, из-за неизбежных термических деформаций заготовки при ее термической резке по полученной раскройной карте. Интерактивный режим позволяет опытному пользователю учесть все технологические требования изготовления заготовок на конкретном технологическом оборудовании и получить качественную раскройную карту (в том числе и не уступающую компьютерному варианту по коэффициенту использования материала). Этот факт часто становится основной причиной того, что на многих предприятиях, эксплуатирующих САПР раскроя, интерактивный модуль все еще остается основным инструментом при проектировании раскройных карт. Естественно, что его эффективность определяется и богатством интерактивных функций и степенью использования в нем автоматических и полуавтоматических функций, которые применяют для ускорения процесса интерактивного проектирования. Вместе с тем, повышение эффективности алгоритмов автоматического раскроя определяет современную тенденцию их преимущественного использования при практическом проектировании.

Исходя из изложенного, общая структурная схема подсистемы проектирования раскроя современной САПР нерегулярного фигурного раскроя имеет вид, показанный на рис. 2.

Подсистема проектирования УП для машины с ЧПУ (САМ-подсистема). При создании такой подсистемы необходимо предусмотреть в ней максимально полное использование всех возможностей современных машин термической резки с ЧПУ. В частности, САМ-модуль интегрированной САПР раскроя материала должен обеспечивать программирование следующих функций:

- резка материала несколькими суппортами;
- резка трехрезаковым блоком для получения фасок;
- резка нескольких заготовок (не включая резака) для уменьшения числа врезок;
- резка с перемычками;
- резка по эквидистантному контуру;
- совмещенный рез;
- использование подпрограмм;
- разработка макропрограмм с помощью встроенного макроязыка и т. д.

Так же как и у подсистемы раскроя, САМ-модуль должен иметь автоматический и интерактивный режимы проектирования управляющей про-

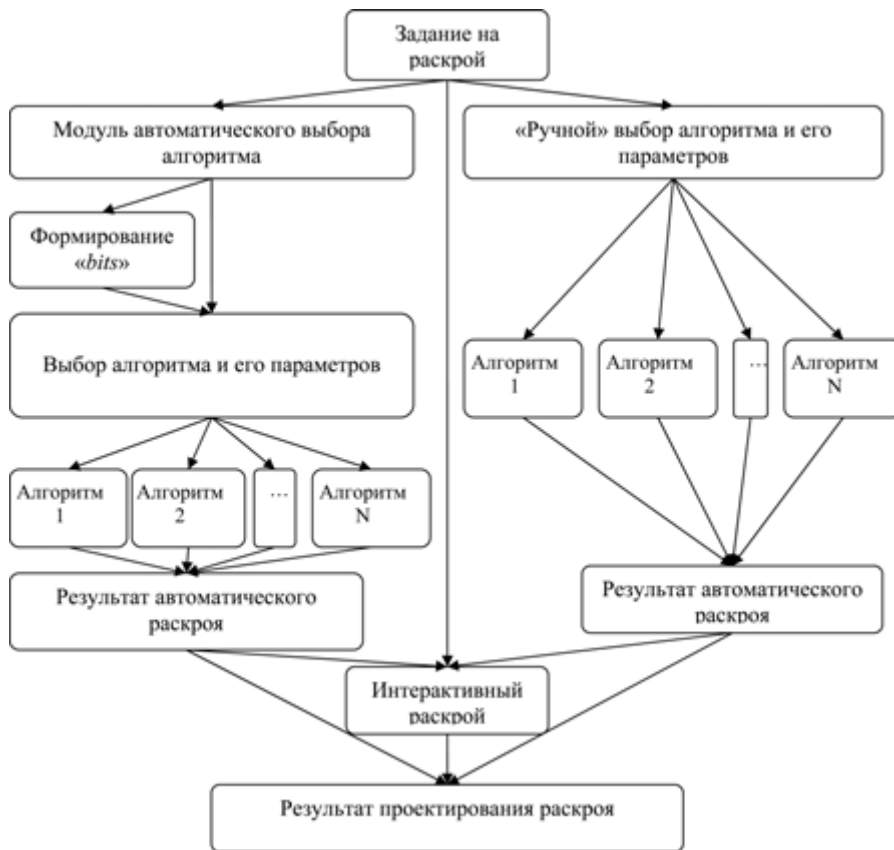


Рис. 2. Структурная схема подсистемы проектирования раскрой

граммы. Это связано с тем, что для многих вариантов раскройных карт существующие алгоритмы автоматического проектирования маршрута резки не обеспечивают выполнение необходимых технологических требований резки материала, что часто приводит к термическим деформациям заготовок. В этом случае только интерактивный режим позволяет опытному пользователю избежать брака.

Следует отметить, что минимизация термических деформаций заготовок при резке является приоритетной задачей проектирования управляющей программы. Другая задача, возникающая при подготовке УП, — минимизация холостого хода инструмента, по мнению автора, носит чисто теоретический характер и не может использоваться на практике, поскольку ее решение часто существенно увеличивает тепловые деформации материала, а во многих случаях приводит к порче заготовок. Кроме того, сокращение времени резки одного листа от оптимизации холостого хода составляет, как правило, несколько секунд, в то время как суммарное время резки "толстого" листа может измеряться часами. Поэтому функция автоматического назначения маршрута резки с учетом минимизации холостого хода инструмента не является обязательной.

САМ-модуль современной САПР должен предусматривать возможность проектирования УП не

только для машин термической резки, но и для машин гидроабразивной резки, которые становятся все более популярными в последнее время, поскольку позволяют выполнять резку специфического материала, в частности камня. Модуль должен обеспечивать автоматизацию подготовки управляющих программ и уже давно применяющихся в раскройно-заготовительном производстве машин с ЧПУ, использующих электроэрозионный способ резки материала, а также маркирующих машин с ЧПУ для разметки материала керном или порошком. В этом смысле, универсальность САМ-модуля может стать главным конкурентным преимуществом на рынке "неуниверсальных", узкоспециализированных САД/САМ-систем раскрой.

Еще одно из требований, предъявляемых к современным подсистемам проектирования УП, — наличие программного обеспечения, служащего для

визуального анализа и редактирования управляющих программ в кодах конкретного станка ЧПУ. По существу это программное обеспечение выполняет функцию контроля УП перед ее выполнением, которая предварительно была получена либо в САМ-модуле после постпроцессирования, либо "вручную". Особенно полезно иметь такой модуль в организациях, где часть управляющих программ получена ранее с использованием различных САМ-систем и хранится длительное время для будущего повторного использования. Модуль для проверки "правильности" и редактирования управляющей программы чаще всего представляет собой отдельное приложение. Примером может служить, в частности, программный продукт NCManager [12].

Заключение

Сформулируем основные результаты, полученные в данной работе:

- описана структура программного обеспечения современной САД/САМ-системы, предназначенной для автоматизированного проектирования раскрой листовых материалов на заготовки сложной геометрической формы и для подготовки управляющих программ технологического оборудования с ЧПУ;

- предложена и обоснована новая схема создания подсистемы автоматизированного проектирования нерегулярного раскроя;
- сформулированы основные требования, предъявляемые к программным модулям системы.

Список литературы

1. Dykhoff Н. A typology of cutting and packing problems // European Journal of Operational research. 1990. N 44. P. 145—159.
2. Мухачева Э. А., Верхотуров М. А., Мартынов В. В. Модели и методы расчета раскроя-упаковки геометрических объектов // Программные продукты и системы. 2000. № 3. С. 47—48.
3. Мухачева Э. А., Валеева А. Ф., Картак В. М., Мухачева А. С. Методы локального поиска оптимума в задачах ортогонального раскроя и упаковки: аналитический обзор и перспективы развития // Информационные технологии. Приложение. 2004. № 5. С. 2—18.
4. Фроловский В. Д. Оптимальное группирование геометрических объектов при проектировании карт раскроя материалов // Программные продукты и системы. 2000. № 3. С. 47—48.
5. Верхотуров М. А. Задача нерегулярного раскроя плоских геометрических объектов: моделирование и расчет рационального раскроя // Информационные технологии. 2000. № 5. С. 37—42.

6. Петунии А. А., Мухачева Э. А., Мухачева А. С. Метод прямоугольной аппроксимации для решения задач нерегулярного фигурного раскроя-упаковки // Информационные технологии. 2008. № 1. С. 28—31.
7. Мазурин А., Пирогов И., Пирогов В. Раскрой листового материала и ЧПУ-обработка на базе T-FLEX CAD // САПР и графика. 2000. № 10. С. 39—43.
8. Петунии А., Галкин И. О применении API T-FLEX CAD при разработке подсистемы автоматизированного проектирования раскроя материала // САПР и графика. 2008. № 12. С. 86—88.
9. Кириленко В. Техтран. Раскрой листового материала 2 // CADmaster. 2001. #09/4 (октябрь—декабрь). С. 28—30.
10. Петунии А. А. Интегрированная САПР "Сириус" для автоматизации раскройно-заготовительного производства. Концепция. Опыт разработки и внедрения // Ресурсосберегающие технологии: математическое обеспечение оптимизационных задач в системах автоматизированного проектирования. (Сборник докладов Первой всероссийской научно-практической конференции по вопросам решения оптимизационных задач в промышленности). Санкт-Петербург: Изд-во ЦНИИТС, 2001. С. 126—129.
11. Полевов А. В. Алгоритмы автоматического раскроя листового материала в условиях единичного производства: дис. ... канд. техн. наук. Уральский государственный технический университет. Екатеринбург, 2005, 127 с.
12. Кац Е., Вахатов Ю., Емелин А., Старостин Н. NCManager: визуализация, контроль и редактирование программ ЧПУ // САПР и графика. 2000. № 9. С. 54—57.

УДК 517.518.2 + 004.0891:673.6

В. А. Немтинов, д-р техн. наук, доц.,
С. Я. Егоров, канд. техн. наук, доц.,
П. И. Пахомов, инженер,
 Тамбовский государственный
 технический университет,
 e-mail: nemtinov@mail.gaps.tstu.ru

Применение теории нечетких множеств и экспертных систем при автоматизированном выборе элемента технической системы

Рассмотрены вопросы разработки процедурной модели автоматизированного выбора элемента сложного технического объекта.

Ключевые слова: теория нечетких множеств, автоматизированный выбор, процедурная модель, элемент технической системы.

Введение

В настоящее время развитие многих технических систем и внедрение энергосберегающих технологий дали толчок к расширению номенклатуры разнообразных устройств — элементов, используемых для управления. Насыщение рынка

зарубежными и отечественными образцами, большое разнообразие условий, в которых работают элементы сложных технических объектов (арматура, контрольно-измерительные приборы и т. п.), специфичность требований, предъявляемых к ним, вопросы надежности и долговечности, наличие большого числа конструктивных типов усложняют выбор элементов для конкретных условий эксплуатации.

В связи с этим в данной статье рассматривается подход, позволяющий в интерактивном режиме выбрать вариант элемента технической системы исходя из известных потребительско-эксплуатационных показателей и их оценок, выполненных различными группами специалистов-экспертов.

Оценка потребительско-эксплуатационных показателей элемента технической системы

Для различных элементов сложного технического объекта могут быть использованы разные показатели. Например, для трубопроводной арматуры, являющейся элементом водопроводных и тепловых систем, наиболее значимыми показателями являются: герметичность затвора (высокая); гидравлическое сопротивление (низкое); частое срабатывание (открывание — закрывание); дросселирование; равнопроходность арматуры и трубопровода; ремонтпригодность и др. Информация о значениях этих показателей часто носит нечеткий характер, так как

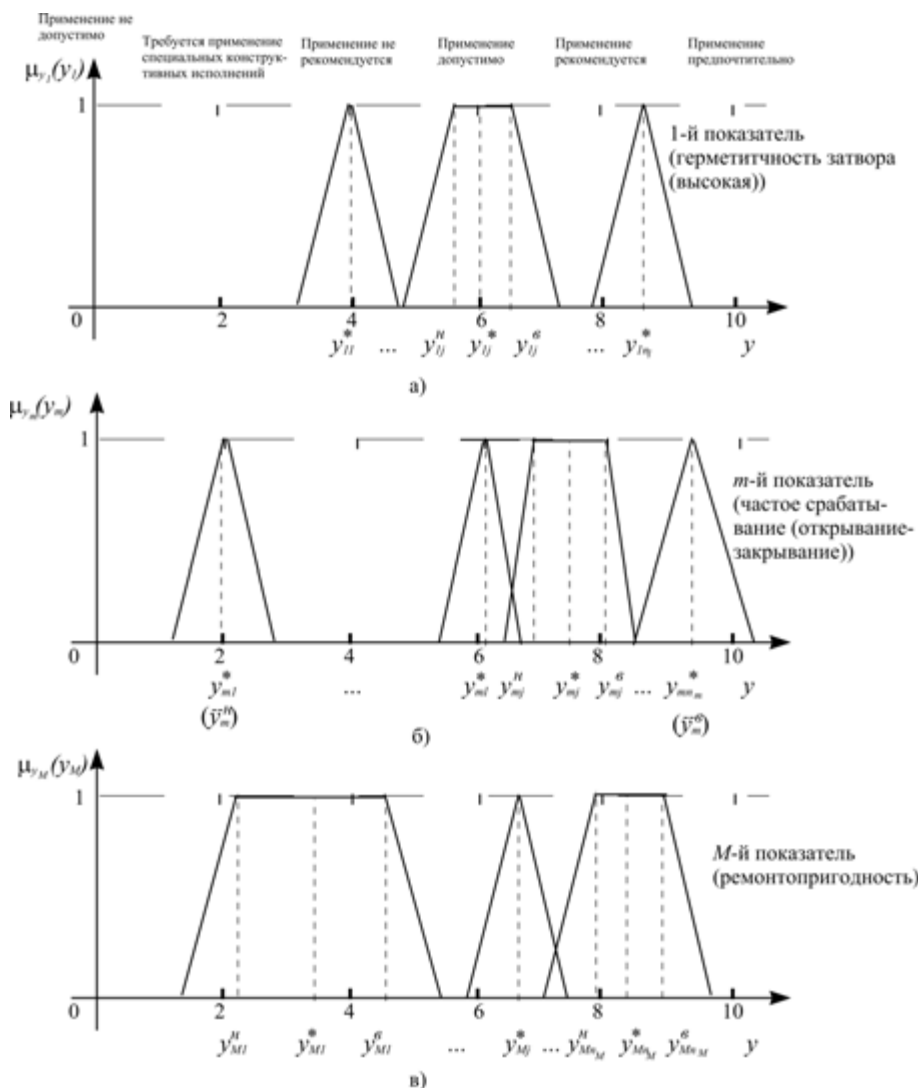


Рис. 1. Дефаззификация нечеткой информации при оценке потребительско-эксплуатационных показателей различных типов элемента (на примере арматуры — элемента трубопроводных систем)

является результатом экспертных оценок различных групп специалистов: конструкторов, производителей, слесарей-ремонтников и т. п.

Для информационного анализа и обработки такой информации целесообразным является использование математического аппарата экспертных систем в качестве систем поддержки принятия решений.

Пусть $R = (r_m | m = \overline{1, M})$ — множество потребительско-эксплуатационных показателей. Для каждого m -го показателя может быть получена качественная или приблизительная количественная информация о его значении. Оценка отдельных показателей может быть описана словами, которые в теории нечетких множеств называются термами [1]. В таком случае значением лингвистической переменной, например показателя "герметичность затвора (высокая)", для конкретного типа элемента могут быть такие термы, как: "применение не рекомендуется", "применение до-

пустимо", "применение рекомендуется", "применение предпочтительно", "требуется применение специальных конструктивных исполнений" и др. Для определения значений лингвистической переменной необходимо определить четкие значения ее термов, позволяющие перейти от лингвистических переменных к их числовым аналогам. В этом случае группа специалистов-экспертов задает вид функции принадлежности, определяющей степень "принадлежности" каждого точного значения к одному из термов лингвистической переменной.

В настоящее время сформировалось понятие о так называемых стандартных функциях принадлежности: Λ -функции, Π -функции, Z -функции, S -функции. Для отображения информации, приведенной выше, целесообразно использовать Λ - и Π -функции, вид которых представлен на рис. 1.

При переходе от нечетких значений величин к вполне определенным необходимы специальные математические методы [1]. Для устранения нечеткости окончательного результата существует несколько методов дефаззификации.

Наиболее часто используемым является метод центра максимума. Рассмотрим его применительно к нашей задаче.

Для тех значений показателя y , которые описаны Λ -функцией, в качестве центра максимума будем использовать величину y_{mj}^* , а при использовании Π -функции — $y_{mj}^* = (y_{mj}^B + y_{mj}^H)/2$ (рис. 1). Взвешенное значение y_m для m -го показателя определим по формуле

$$y_m = \frac{\sum_{j=1}^{n_m} \rho_{mj} y_{mj}^*}{\sum_{j=1}^{n_m} \rho_{mj}}, \quad (1)$$

где ρ_{mj} — коэффициент значимости информации j -й группы специалистов для m -го показателя. Примем:

$$\rho_{mj} = 1 \text{ — при использовании } \Lambda\text{-функции; } \quad (2)$$

$$\rho_{mj} = \frac{y_{mj}^B - y_{mj}^H}{\bar{y}_m^B - \bar{y}_m^H} \text{ — для П-функции,} \quad (3)$$

где \bar{y}_m^H, \bar{y}_m^B — соответственно нижняя и верхняя границы диапазона возможного изменения показателя y_m (см. рис. 1, б).

Процедурная модель выбора элемента технического объекта

В связи с большим разнообразием выпускаемых промышленностью различных типов одного и того же элемента технической системы, характеризующихся одинаковыми основными параметрами, возникает необходимость выбора такого элемента, который бы отвечал множеству требований R конкретного потребителя.

Нами предлагается процедурная модель автоматизированного выбора элемента технической системы, характеризующегося наилучшими заданными потребительскими показателями для каждого конкретного случая.

На первом этапе выбора для заданных конструктивных размеров элемента $K = (k_l | l = \overline{1, L})$ (например, для элемента трубопроводной системы — трубопроводной арматуры — условного прохода и условного давления) формируется подмножество допустимых типов, выпускаемых промышленностью $T^{dop} \subset T$, где T — множество всех типов выбираемого элемента. Формирование T^{dop} осуществляется с использованием базы данных типов элемента, созданной в соответствии с используемой для них классификацией.

На следующем этапе автоматизированного выбора элемента среди подмножества типов T^{dop} необходимо выбрать такой тип, который бы удовлетворял выбранному потребителем подмножеству потребительских требований (показателей) $R^{zad} \subset R$.

Вследствие значительного числа критериев оценки (для отдельных элементов более 20), которые могут быть использованы потребителем при выборе критерия, нами предложена следующая схема выбора. При этом решается задача, математическая постановка которой формулируется следующим образом:

для заданного функционального назначения, основных потребительских параметров K и выбранных потребительско-эксплуатационных показателей $R^{zad} \subset R$ найти такой тип элемента, для которого справедливо условие

$$t_{opt} = \arg \min_{t \in T^{dop}} F(t) \quad (4)$$

при условии, что с позиций выбранных эксплуатационных показателей применение t -го типа элемента возможно:

$$y_m(t) \neq 0, m \in R^{zad}, t \in T^{dop}, \quad (5)$$

где $y_m(t)$ — количественная оценка m -го показателя для t -го типа элемента.

В такой постановке рассматриваемая задача относится к классу многокритериальных задач. Критерий оптимальности $F(t)$ представим в виде суммы взвешенных относительных потерь критериев: стоимости элемента и некоторого подмножества потребительских показателей $R^{zad} \subset R$, задаваемого лицом, принимающим решение (ЛПР), и выберем соответствующие методы нормализации множества критериев и их ранжирования [2–4]. В этом случае критерий F можно записать как

$$F(t) = \sum_{i=1}^{R^{zad}} \rho_i \omega_i(t), \quad (6)$$

где $\rho_1, \rho_2, \dots, \rho_i, \dots, \rho_{R^{zad}}$ — весовые коэффициенты:

$$\rho = \{\rho_i\} = \{\rho_i : \rho_i > 0, i = 1, \dots, R^{zad}, \sum_{i=1}^{R^{zad}} \rho_i = 1\}; \quad (7)$$

$\rho_i \omega_i(t)$ — взвешенные потери по i -му критерию; $\omega_i(t) = \omega_i(F_i(t))$, $i = 1, \dots, R^{zad}$, $t \in T^{dop}$ — монотонные функции, преобразующие каждую функцию цели $F_i(t)$, $i = 1, \dots, R^{zad}$, $t \in T^{dop}$, к безразмерному виду.

$F_1(t)$ — экономический критерий, включающий в себя стоимость элемента технической системы; $F_i(t)$, $i = \overline{2, R^{zad}}$ — оценка потребительско-эксплуатационных показателей. Для функции цели $F_1(t)$ находится минимум, а для остальных — максимум:

$$\omega_1(t) = \frac{F_1(t) - F_{10}}{F_{1(\max)} - F_{10}}, \quad t \in T^{dop}; \quad (8)$$

$$\omega_i(t) = \frac{F_{i0} - F_i(t)}{F_{i0} - F_{i(\min)}}, \quad t \in T^{dop}, \quad i = \overline{2, R^{zad}}, \quad (9)$$

где $F_{1(\max)}$ — наибольшее значение минимизируемой функции $F_1(t)$, $t \in T^{dop}$, на множестве допустимых альтернатив T^{dop} ; $F_{i(\min)}$, $i = \overline{2, R^{zad}}$, — наименьшее значение максимизируемых функций $F_i(t)$, $t \in T^{dop}$, на множестве допустимых альтернатив T^{dop} ; F_{i0} , $i = \overline{1, R^{zad}}$, — оптимальные значения функций цели. Значения $\omega_i(t)$, $i = \overline{1, R^{zad}}$, $t \in T^{dop}$, лежат в пределах от 0 до 1.

Как и при решении других многокритериальных задач, в данном случае необходимо найти такую компромиссную альтернативу $t \in T^{dop}$, которая может не являться оптимальной ни для одной функции цели $F_i(t)$, но оказываться приемлемой для интегрального критерия $F(t)$.

Для выбора компромиссного решения в задаче принятия сложного решения зададим весовые коэффициенты $\rho_i, i = \overline{1, R^{zad}}$, удовлетворяющие соотношению (7) и отражающие относительную важность всех функций цели. Наиболее эффективными подходами к определению этого предпочтения являются методы ранжирования и приписывания баллов [3] (последний применен в данной работе).

Процедура решения задачи (4)—(9) сводится к формированию множества типов элемента системы заданного функционального назначения, отвечающих основным потребительским параметрам и выбранным ЛПР эксплуатационным показателям $R^{zad} \subset R$, для которых выполняется условие (5). Далее с использованием метода полного перебора выбирается такой тип t_{opt} для которого значение критерия $F(t)$ имеет минимальное значение.

Практическая реализация процедурной модели

Рассмотрим реализацию предложенной модели принятия решений при выборе элемента технической системы на примере выбора трубопроводной арматуры.

Насыщение рынка зарубежными и отечественными образцами, большое разнообразие условий, в которых работает арматура, специфичность требований, предъявляемых к ней, вопросы надежности и долговечности, наличие большого числа конструктивных типов усложняют выбор арматуры для конкретных условий эксплуатации в современных системах.

В настоящее время общепринятой является классификация трубопроводной арматуры по следующим параметрам:

- по способу перекрытия потока среды;
- по области применения;
- по материалу корпусных деталей;
- по методу управления;
- по функциональному назначению и т. п.

Запорная арматура различных размеров и типов является стратегически важными узлами в установках, в которых используются жидкости и газы. Протечка лишь в одном из них может привести к необходимости прекращения всего процесса. Замена поврежденной арматуры является сложным и трудоемким процессом, приводящим к производственным убыткам и существенным затратам и вынуждающим компании вкладывать

Фрагмент базы данных оценок потребительско-эксплуатационных показателей типов арматуры

Наименование показателя	Назначение арматуры	Тип арматуры	Оценка показателя y_m
Герметичность затвора (высокая)	Запорная	Задвижки с выдвигным шпинделем шиберные	8,4
		Задвижки с неподвижным шпинделем шланговые	9,2
		Клапаны запорные сальниковые с наружной резьбой проходные	8,2
		Клапаны обратные поверхностные	3,9
	
...
Масса (небольшая)	Запорная	Задвижки с выдвигным шпинделем клиновые двухдисковые	0
		Задвижки с выдвигным шпинделем шиберные	6,8
		Клапаны запорные бессальниковые диафрагмовые	7,5
		Клапаны обратные безударные	8,9
	
...
Высокое давление	Запорная	Клапаны запорные сальниковые с внутренней резьбой угловые	9,4
		Клапаны запорные сальниковые с внутренней резьбой прямоточные	3,5
		Краны пробковые конические сальниковые с подъемом пробки	0,8
		Краны шаровые проходные	7,5
		Клапаны обратные подъемные	8,2
...
...

значительные денежные средства в запчасти, для того чтобы сократить время простоя.

При проведении ремонта сетей специалистам очень часто приходится решать вопросы, связанные с заменой с трубопроводной арматуры.

Предложенная процедурная модель автоматизированного выбора трубопроводной запорной арматуры реализована в виде комплекса программ, написанных в среде системы программирования Visual Basic с использованием СУБД MS Access.

Количественная оценка потребительско-эксплуатационных показателей арматуры получена в ходе обработки разнообразной качественной и количественной информации в соответствии с подходом, описанным выше. В таблице приведен фрагмент базы данных количественных оценок потребительско-эксплуатационных показателей для отдельных типов запорной арматуры.

Проиллюстрируем реализацию модели автоматизированного выбора запорной арматуры на сле-

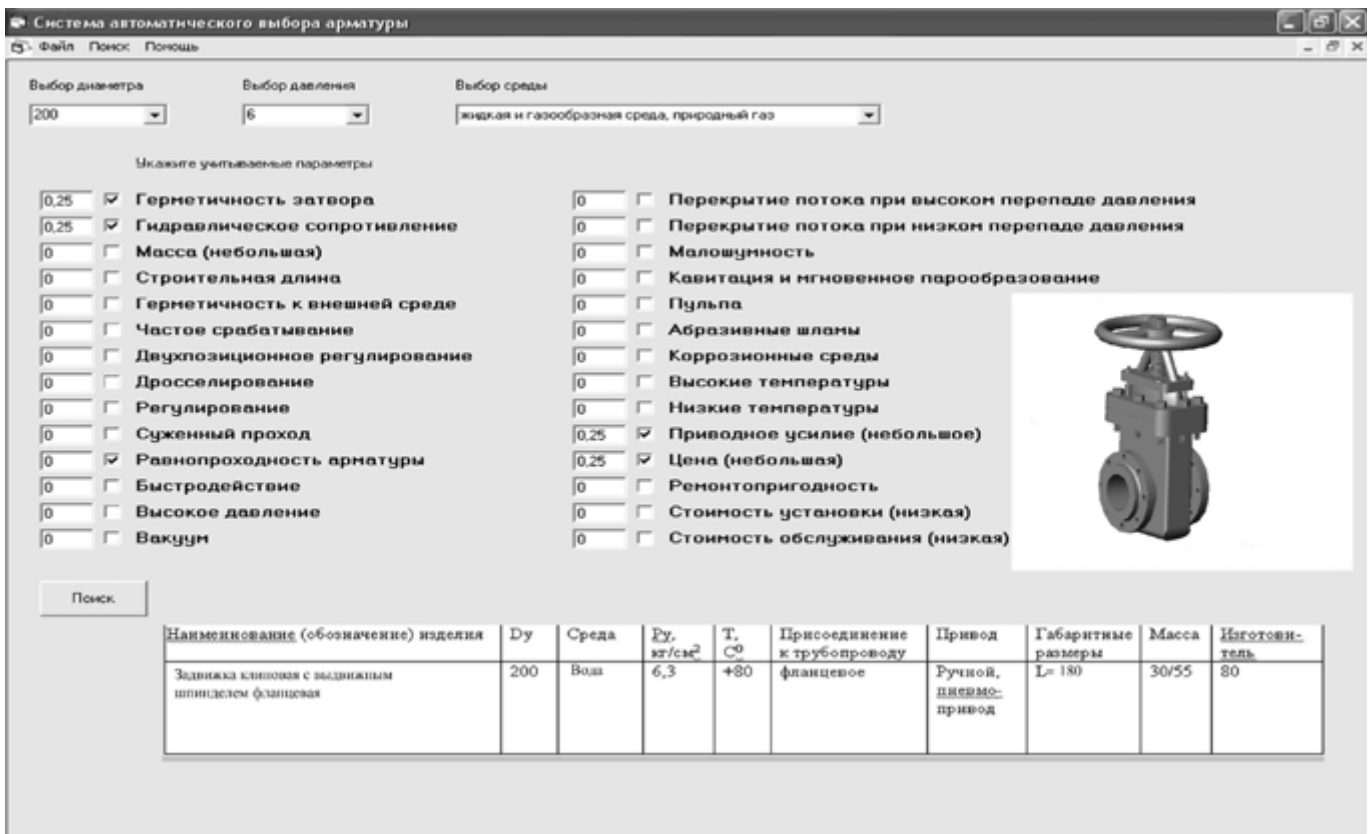


Рис. 2. Вид рабочего окна при выборе запорной трубопроводной арматуры

дующем примере. При ремонте водопроводной сети потребовалась замена трубопроводной запорной арматуры на трубопровод следующих параметров: $D_y^{зад} = 200$ мм и $P_y^{зад} = 58,8 \cdot 10^4$ Па. При выборе следует учитывать следующие потребительские показатели: высокую герметичность затвора, низкое гидравлическое сопротивление, небольшое приводное усилие и невысокую цену изделия.

Для каждого из используемых критериев (стоимость, высокая герметичность затвора, низкое гидравлическое сопротивление, небольшое приводное усилие) заданы равные веса $\rho_i = 0,25$, $i = \overline{1, 4}$. В конечном итоге оптимальным решением с позиций принятых критериев оказалась задвижка клиновая с выдвижным шпинделем фланцевая (рис. 2).

Заключение

Предложенная авторами процедурная модель автоматизированного выбора элемента техниче-

ской системы, реализованная на примере трубопроводной запорной арматуры для водопроводных и тепловых сетей в виде пакета программ, позволяет потребителю выбрать оптимальный вариант решения задачи с учетом его различных требований и экспертных оценок потребительско-эксплуатационных показателей элемента системы, выполненных разными группами специалистов.

Список литературы

1. **Нечеткие** множества и теория возможностей / Под ред. Р. Ягера. М.: Радио и связь. 1986. 342 с.
2. **Многовариантный** типологический подход в задачах обучения и обработки данных / Фетина Е. П., Кораблина Т. В., Криволапова Л. И. и др. // Известия вузов. Черная металлургия. 2000. № 4. С. 57—60.
3. **Михалевич В. С., Волкович В. Л.** Вычислительные методы исследования и проектирования сложных систем. М.: Наука, 1982. 286 с.
4. **Айзерман М. А., Алескеров Ф. Т.** Выбор вариантов. Основы теории. М.: Наука, 1990. 227 с.

УДК 025.171; 004.421.6

Алыгулиев Р. М., канд. физ.-мат. наук, зав. отд.,
Институт информационных технологий
Национальной Академии наук Азербайджана,
г. Баку, Азербайджан,
e-mail: a.ramiz@science.az

Реферирование документов через ранжирование предложений

Предлагается метод автоматического реферирования документов через ранжирование предложений. Для ранжирования предложений применяется модифицированный вариант алгоритма PageRank.

Ключевые слова: реферирование документа, ранжирование предложений, алгоритм PageRank, сопротивление.

Введение

Автоматическое реферирование текстовых документов является одним из перспективных направлений в области обработки текстов естественного языка и информационного поиска. Потребность в автоматическом реферировании обусловлена, в первую очередь, возрастающим потоком текстовой информации. Задача автоматического реферирования — это процесс создания сокращенной версии исходного документа (или документов), отражающей его (их) основной контент. Подходы к реферированию текстовых документов можно разбить на два класса [1]: абстрагирование и экстрагирование. Первый подход — абстрагирование генерирует реферат путем формирования новых предложений, необязательно теми же словами, которые есть в тексте. Другой подход — экстрагирование формирует реферат путем извлечения информативных предложений из исходного документа. Подавляющее большинство методов реферирования документов, за исключением нескольких, основано на извлечении из них информативных предложений. Оно связано со сложностью автоматического формирования грамматически правильных предложений. Поэтому все усилия исследователей направлены на создание рефератов путем извлечения информативных предложений. Реферат, полученный путем соединения предложений, извлеченных из документа-источника, лишен гладкости, его трудно читать,

однако он по сравнению с абстрактом, созданным человеком, лишен субъективизма.

Как отмечено выше, при автоматическом реферировании документов основной задачей является выявление информативных предложений. С этой целью разработаны разные методы. Методы, предложенные в работах [2—4], представляют документ в виде неориентированного взвешенного графа, где вершины соответствуют предложениям, а вес ребер — мере близости между ними. Степень информативности предложений определяется суммой весов ребер, соединяющих их с другими предложениями. С вычислительной точки зрения такой подход реферирования очень эффективен. Однако он имеет два существенных недостатка. Во-первых, реферат, созданный таким способом, не лишен избыточности, т. е. в нем могут повторяться близкие по смыслу предложения. Другим существенным недостатком этого подхода является то, что он не может выявить скрытые тематические разделы. Следовательно, он эффективен в случае, когда документ посвящен только одной теме. Если в документе затрагивается несколько тем, то такой подход не дает ожидаемых результатов. Для решения этой проблемы в работах [5—8] предложены методы, определяющие и тематические разделы, и информативные предложения в документе. Предлагаемые методы состоят из двух этапов. На первом этапе осуществляется кластеризация предложений по контенту. После кластеризации на каждом кластере (тематическом разделе) определяются информативные предложения, отражающие основной контент кластера.

В последние годы алгоритмы анализа Web-структуры успешно применяют не только для ранжирования результатов поиска, их также используют для других целей, таких как кластеризация Web-ресурсов, реферирование документов и т. д. Среди алгоритмов ранжирования, основанных на анализе Web-графа, наиболее популярным является алгоритм PageRank [9]. В первоначальном варианте алгоритма PageRank степень важности страниц определялась числом входящих гиперссылок, игнорируя при этом тематическую близость страниц. Для повышения эффективности в последующих работах [10—13] предложены некоторые модификации алгоритма PageRank, которые учитывают тематическую близость соседних страниц.

В настоящей статье предлагается такая модификация алгоритма PageRank, в которой при вычисле-

нии ранга страницы учитываются скрытые отношения между ними. Для выявления скрытых отношений между вершинами графа используется понятие "расстояние сопротивления" [14–16]. Предложенная модификация применена для реферирования документов и сравнена с другими методами.

Расстояние сопротивления и алгоритм PageRank

Суть алгоритма PageRank заключается в вычислении степени значимости вершин Web-графа, т. е. Web-страниц. Согласно этому алгоритму пользователь с вероятностью d следует по одной из исходящих гиперссылок или с вероятностью, равной $(1 - d)$, "прыгает" на случайную страницу в Web-графе. При этом предполагается, что пользователь никогда не возвращается к предыдущей странице, т. е. он двигается только в одном направлении — по исходящим гиперссылкам. Процесс моделируется итерационным процессом [9]:

$$PR(v_j) = \frac{(1-d)}{n} + d \sum_{v_i \in B(v_j)} \frac{PR(v_i)}{|F(v_i)|}, \quad (1)$$

где $PR(v_i)$ — PageRank-страницы v_i ; $B(v_i)$ — множество страниц, которые ссылаются на страницу v_j ; $F(v_i)$ — множество страниц, на которые ссылается страница v_i ; $d \in [0.8, 1]$ — коэффициент демпфирования. В формуле (1) $\frac{1}{n}$ означает, что вероятность выбора каждой n страницы одинаковая. Из этого алгоритма также видно, что вероятность выбора каждой исходящей ссылки одинаковая, равная $\frac{1}{|F(v_i)|}$, где $|F(v_i)|$ — число исходящих ссылок страницы v_i . Всем вершинам в начальной итерации присваивается одинаковый (ненулевой) ранг.

Очевидно, что от вершины v_i к вершине v_j можно прийти по разным маршрутам. При этом вклад ранга вершины v_i в вершину v_j в зависимости от маршрута будет разным. Для учета вклада вершины v_i в вершину v_j по всевозможным маршрутам предлагается следующая модификация алгоритма PageRank:

$$PR_{\text{resist}}(v_j) = \frac{(1-d)}{n} + d \sum_{\substack{v_i \in V \\ v_i \neq v_j}} \frac{PR_{\text{resist}}(v_i)}{r_{ij}}, \quad (2)$$

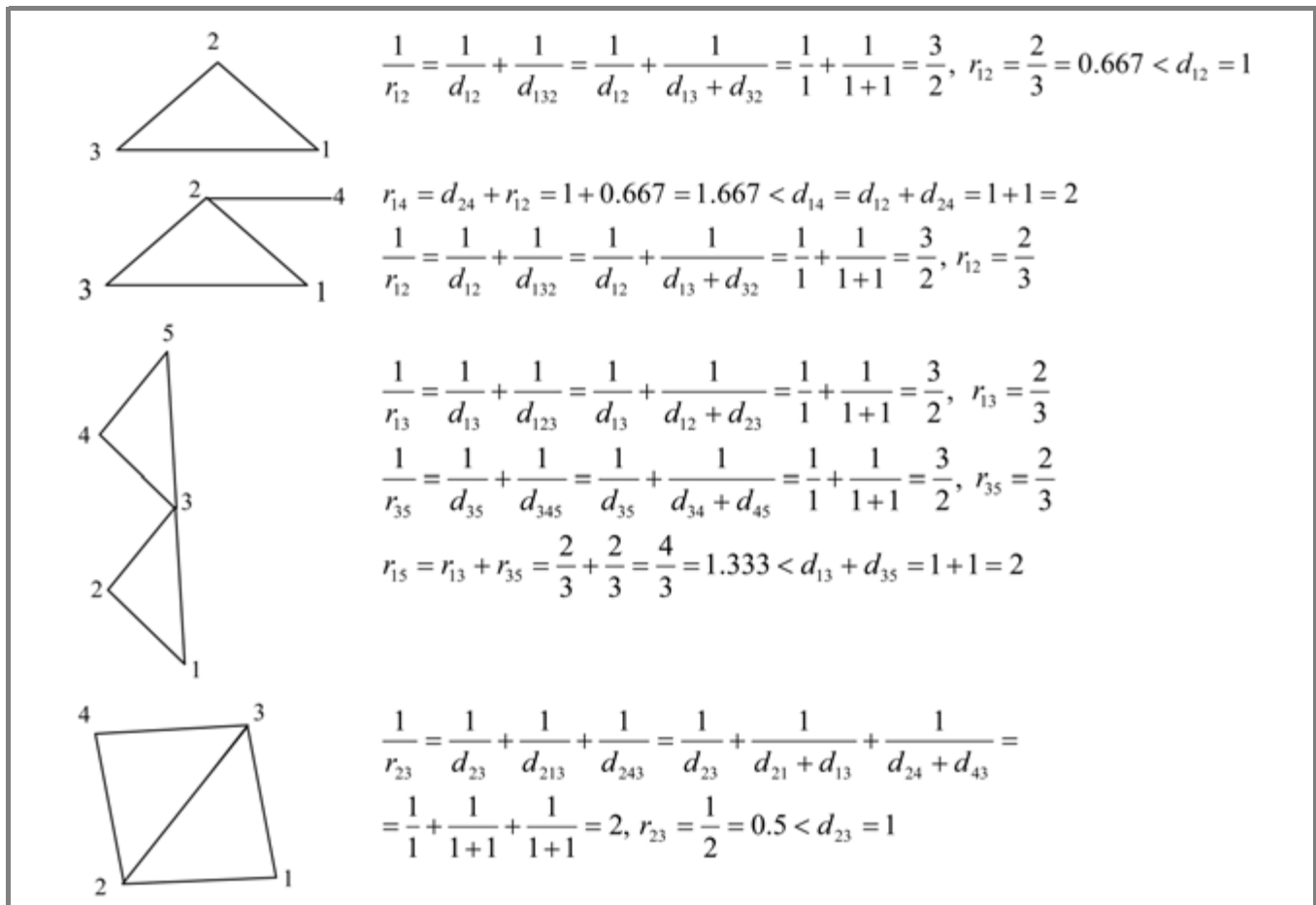


Рис. 1. Примеры вычисления расстояния сопротивления

где r_{ij} — расстояние сопротивления между вершинами v_i и v_j . Оно определяется аналогично вычислению сопротивления электрических цепей. Для более точного понимания расстояния сопротивления рассмотрим примеры [16], приведенные на рис. 1.

В примерах через d_{ij} обозначено топологическое расстояние, равное длине кратчайшего пути, соединяющего вершины v_i и v_j . Как видно, расстояние сопротивления всегда меньше, чем топологическое расстояние. Расстояние сопротивления r_{ij} можно считать результирующим, объединяющим в себе всевозможные пути между вершинами v_i и v_j .

Для вычисления расстояния сопротивления сначала формируется матрица Лапласа. Матрица Лапласа \mathbf{L} неориентированного графа определяется таким образом [14]:

$$\mathbf{L} = \mathbf{D} - \mathbf{A}, \quad (3)$$

где \mathbf{A} — матрица смежности, элементы которой определяются так:

$$a_{ij} = \begin{cases} 1, & \text{если вершины } v_i \text{ и } v_j \text{ смежны} \\ 0 & \text{в противном случае.} \end{cases} \quad (4)$$

В случае взвешенного графа элементы матрицы \mathbf{A} определяются так:

$$a_{ij} = \begin{cases} w_{ij}, & \text{если вершины } v_i \text{ и } v_j \text{ смежны} \\ 0 & \text{в противном случае,} \end{cases} \quad (5)$$

где $w_{ij} \geq 0$ является весом ребра, соединяющего вершины v_i и v_j .

Матрица \mathbf{D} с элементами $d_{ii} = [\mathbf{D}]_{ii} = a_{i*} = \sum_{j=1}^n a_{ij}$ — диагональная.

Для ориентированного графа матрица Лапласа определяется таким образом:

$$\mathbf{L} = \mathbf{A}\mathbf{A}^T, \quad (6)$$

где \mathbf{A}^T — транспонирование матрицы \mathbf{A} .

Матрица инцидентности \mathbf{A} ориентированного графа $G = (V, E)$ с $n \times m$ элементами (n — число вершин, m — число ребер) определяется так:

$$(\mathbf{A})_{ie} = \begin{cases} -1, & \text{если вершина } v_i \text{ является концом ребра } e, \\ 1, & \text{если вершина } v_i \text{ является началом ребра } e, \\ 0, & \text{если вершина } v_i \text{ и ребро } e \text{ несмежны.} \end{cases} \quad (7)$$

Из определения (3)—(7) следует, что для любого графа $\sum_{i=1}^n (\mathbf{L})_{ij} = \sum_{j=1}^n (\mathbf{L})_{ij} = 0$, т. е. матрица \mathbf{L} — сингулярная. Стандартный метод определения расстояния сопротивления сводится к вычислению псевдообращения \mathbf{L}^+ Мура—Пенроуза (Moore—Penrose) [14] сингулярной матрицы Лапласа \mathbf{L} :

$$r_{ij} = (\mathbf{L}^+)_{ii} + (\mathbf{L}^+)_{jj} - (\mathbf{L}^+)_{ij} - (\mathbf{L}^+)_{ji}. \quad (8)$$

В формуле (8) $(\mathbf{L}^+)_{ij}$ — ij -й элемент псевдообращения \mathbf{L}^+ . Как показано в работах [14, 15], вычисление псевдообращения \mathbf{L}^+ Мура—Пенроуза — трудоемкая работа, оно требует больших вычислительных затрат. Поэтому для упрощения вычисления расстояния r_{ij} в работе [15] предложена такая формула:

$$r_{ij} = \frac{\det \mathbf{L}(i, j)}{\det \mathbf{L}(i)}, \quad (9)$$

где подматрица $\mathbf{L}(i)$ получена из матрицы \mathbf{L} удалением i -й строки и i -го столбца, а подматрица $\mathbf{L}(i, j)$ получена удалением i -х и j -х строк и столбцов, $i \neq j$.

Расстояние сопротивления имеет одно очень существенное преимущество. Оно обнаруживает скрытое отношение между вершинами графа. Для демонстрации рассмотрим простой пример, в котором расстояние сопротивления сравнивается с алгоритмом *SimRank*. Алгоритм *SimRank* [17] основывается на идее, что если Web-страницы указываются подобными страницами, то они подобны:

$$\text{SimRank}(u, v) = \gamma \frac{\sum_{a \in I(u)} \sum_{b \in I(v)} \text{SimRank}(a, b)}{|I(u)| \cdot |I(v)|}, \quad (10)$$

где $\gamma \in (0, 1)$; $I(x)$ — множество страниц, ссылающих к странице x ; $|I(x)|$ — мощность множества $I(x)$, т. е. число входящих ссылок страницы x . Если множества $I(u)$ и $I(v)$ пустые, то согласно определению $\text{SimRank}(u, v) = 0$. Алгоритм *SimRank* начинает итерацию с $\text{SimRank}_0(u, v) = 1$ для $u = v$ и $\text{SimRank}_0(u, v) = 0$ для $u \neq v$. $\text{SimRank}(u, v)$ между u и v определяется как $\lim_{k \rightarrow \infty} \text{SimRank}_k(u, v)$.

Пусть задан Web-граф (числа на гиперссылках указывают их номера), изображенный на рис. 2.

Используя формулы (3)—(9) и (10), получим следующие матрицы:

$$\mathbf{RD} = \begin{pmatrix} 0.00 & 0.67 & 0.67 & 1.67 & 1.67 & 1.67 \\ 0.67 & 0.00 & 0.67 & 1.00 & 1.00 & 1.67 \\ 0.67 & 0.67 & 0.00 & 1.67 & 1.67 & 1.00 \\ 1.67 & 1.00 & 1.67 & 0.00 & 2.00 & 2.67 \\ 1.67 & 1.00 & 1.67 & 2.00 & 0.00 & 2.67 \\ 1.67 & 1.67 & 1.00 & 2.67 & 2.67 & 0.00 \end{pmatrix},$$

$$\text{SimRank} = \begin{pmatrix} 1.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 \\ 0.00 & 1.00 & 0.50 & 0.00 & 0.00 & 0.00 \\ 0.00 & 0.50 & 1.00 & 0.50 & 0.50 & 0.25 \\ 0.00 & 0.00 & 0.50 & 1.00 & 1.00 & 0.50 \\ 0.00 & 0.00 & 0.50 & 1.00 & 1.00 & 0.50 \\ 0.00 & 0.00 & 0.25 & 0.50 & 0.50 & 1.00 \end{pmatrix}.$$

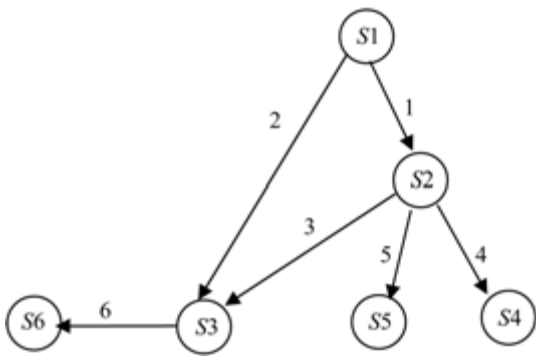


Рис. 2. Web-граф

Сравнение результатов показывает, что модификация алгоритма *PageRank*, использующая расстояние сопротивления, демонстрирует более разумный результат, чем алгоритм *SimRank*. Чтобы убедиться в этом, следует обратиться к следующему. Несмотря на то, что между S1 и S2, S1 и S3, S2 и S4, S2 и S5 существуют гиперссылки, алгоритм *SimRank* показывает, что мера подобия между ними равна нулю. Это, конечно, не соответствует действительности. Однако **RD** показывает, что не только они, все вершины графа в какой-то степени подобны между собой.

Реферирование документа ранжированием предложений

Пусть задан документ $D = \{S_1, \dots, S_n\}$, где n — число предложений в документе. Для применения изложенного выше алгоритма ранжирования документ должен представляться в виде взвешенного графа. Пусть $T = \{t_1, \dots, t_m\}$ — множество слов в документе. При использовании векторной модели (*Vector Space Model*) предложение S_i представляется как вектор m -мерного пространства $S_i = \{w_{i1}, \dots, w_{im}\}$, элементами w_{ik} которого являются веса слов t_k ($k = 1, \dots, m$) в этом предложении S_i . Вес слова определяется по схеме TF-ISF (*Term Frequency-Inverse Sentence Frequency*), $w_{ik} = f_{ik} \log\left(\frac{n}{n_k}\right)$. Здесь f_{ik} — частота появления слова t_k в предложении S_i , а n_k — число предложений, в которых встречается слово t_k . После такого представления, используя метрику косинуса, вычисляем меру близости между парами S_i и S_j :

$$w_{ij} = \sin(S_i, S_j) = \cos(S_i, S_j) = \frac{\sum_{k=1}^m w_{ik} w_{jk}}{\sqrt{\sum_{k=1}^m w_{ik}^2 \cdot \sum_{k=1}^m w_{jk}^2}},$$

$i, j = 1, \dots, n.$

Если мера близости $w_{ij} > 0$, то вершины S_i и S_j соединяются ребром. Таким образом, документ представляется в виде неориентированного взвешенного графа, где вес ребер соответствует мере близости вершин (предложений).

Для демонстрации эффективности предложенного алгоритма рассмотрим следующий пример. В этом примере наш алгоритм будет сравнен с алгоритмами *PageRank* и *Topic-Centric* [11]. Согласно алгоритму *Topic-Centric* ранг вершин определяется так:

$$PR_{TC}(v_j) = \frac{(1-d)}{n} + d \sum_{v_i \in B(v_j)} \frac{\text{sim}(v_i, v_j)}{\sum_{v_k \in F(v_i)} \text{sim}(v_i, v_k)} PR_{TC}(v_i).$$

Пример. Пусть документ состоит из восьми предложений. Представим его в виде неориентированного взвешенного графа (числа на дугах указывают степени близости вершин):

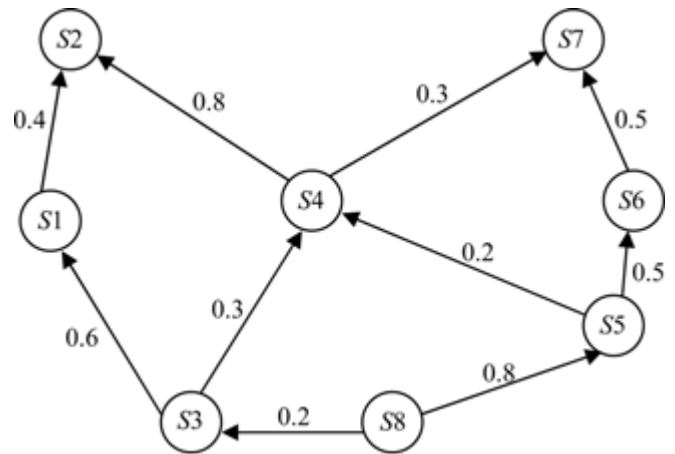


Рис. 3. Граф связей предложений

Матрица сопротивления представленного графа имеет вид:

$$RD = \begin{pmatrix} 0,000 & 1,702 & 1,316 & 1,965 & 3,465 & 4,254 & 4,053 & 3,579 \\ 1,702 & 0,000 & 1,965 & 1,053 & 3,079 & 3,746 & 3,342 & 3,404 \\ 1,316 & 1,965 & 0,000 & 1,693 & 2,842 & 3,711 & 3,640 & 2,825 \\ 1,965 & 1,053 & 1,693 & 0,000 & 2,289 & 2,895 & 2,386 & 5,088 \\ 3,465 & 3,079 & 2,842 & 2,289 & 0,000 & 1,658 & 2,921 & 1,114 \\ 4,254 & 3,746 & 3,711 & 2,895 & 1,658 & 0,000 & 2,386 & 2,614 \\ 4,053 & 3,342 & 3,640 & 2,386 & 2,921 & 2,386 & 0,000 & 2,456 \\ 3,579 & 3,404 & 2,825 & 5,088 & 1,114 & 2,614 & 2,456 & 0,000 \end{pmatrix}$$

Результаты ранжирования, полученные алгоритмами *PR*, *PR_{TC}* и *PR_{resist}*, показаны в табл. 1 (расчеты проведены для $d = 0,8$).

Анализ результатов (см. табл. 1) показывает, что наш алгоритм дает более разумный результат, чем алгоритмы *PageRank* и *Topic-Centric*. Для этого достаточно сравнить ранги топ 3 вершин, ранжи-

Таблица 1

Ранг предложений

Вер- шины	PR		PR _{TC}		PR _{resist}	
	Значение	Ранг	Значение	Ранг	Значение	Ранг
S1	0,039	4,5	0,040	6	2,350	5
S2	0,077	1	0,082	1	2,610	3
S3	0,035	6,7	0,029	7	2,512	4
S4	0,053	3	0,042	4	2,783	1
S5	0,035	6,7	0,048	3	2,621	2
S6	0,039	4,5	0,041	5	2,043	7
S7	0,056	2	0,063	2	1,951	8
S8	0,025	8	0,025	8	2,274	6

Таблица 2

Сравнение методов реферирования

Методы	ROUGE-2	ROUGE-SU4	Улучшение алгоритма PR _{resist} %	
			ROUGE-2	ROUGE-SU4
PR _{resist}	0,1207	0,1721	—	—
PR _{TC}	0,1037	0,1564	16,39	10,68
PR	0,0923	0,1436	30,77	20,54
LexRank	0,0987	0,1487	22,29	16,41
PPRSum	0,1195	0,1709	1,00	1,29

рованных по каждому алгоритму. Из рис. 3 видно, что вершина S4 с другими вершинами имеет более тесную связь, чем любая другая вершина. Аналогичным рассуждением можно прийти к заключению, что после S4 вершины S5 и S2 должны иметь более высокие ранги. Вместе с тем, сравнение результатов алгоритмов *PageRank* и *Topic Centric* показывает, что алгоритм *Topic-Centric* дает более разумный результат, чем алгоритм *PageRank*. Это, как показано и в предыдущих исследованиях [10, 11, 13], является следствием того, что в алгоритме *PageRank* не учитывается тематическая близость вершин, т. е. ранг страницы определяется только количеством входящих гиперссылок, игнорируя при этом степень близости страниц.

Эксперименты

На практике реферирование многократных документов более запрашиваемое действие, чем единичных документов, поэтому для оценивания нашего алгоритма применим его к реферированию многократных документов. Мы изучаем выполнение алгоритма на наборе документов DUC2007 [18]. DUC2007 состоит из 45 корпусов, где каждый корпус содержит 25 документов. Для оценки нашего алгоритма выбираем первые 10 документов из этого набора и реферат, содержащий 250 слов, созданный независимыми экспертами. В качестве критерия оценивания использу-

ется мера ROUGE-N, которая определяется следующей формулой:

$$\text{ROUGE-N} = \frac{\sum_{S \in \text{RefSumm}} \sum_{\text{gram}_N \in S} \text{Count}_{\text{match}}(\text{gram}_N)}{\sum_{S \in \text{RefSumm}} \sum_{\text{gram}_N \in S} \text{Count}(\text{gram}_N)},$$

где N означает длину N -грамм, $\text{Count}_{\text{match}}(\text{gram}_N)$ — максимальное число N -грамм, появляющихся и в автоматически созданном реферате, и в рефератах, выполненных независимыми экспертами, $\text{Count}(\text{gram}_N)$ — число N -грамм в автоматически созданном реферате.

Для оценки качества работы методов реферирования были вычислены меры ROUGE-2 и ROUGE-SU4. Алгоритм PR_{resist} будет сравнен с алгоритмами *PageRank* [9], *Topic-Centric* [11], *LexRank* [19] и *PPRSum* [20]. Результаты эксперимента показаны в табл. 2. Как можно было ожидать, наихудший результат показал алгоритм *PR* (*PageRank*). Среди других алгоритмов лучший результат получен алгоритмом *PPRSum* (*Personalized PageRank*). Из табл. 2 видно, что алгоритм PR_{resist} улучшает результат наилучшего метода *PPRSum* на 1,00 % и 1,29 % в мерах ROUGE-2 и ROUGE-SU4, соответственно.

Заключение

В алгоритме *PageRank* ранг страницы напрямую зависит от рангов ссылающихся страниц. Как известно, от страницы к странице можно прийти разными маршрутами. При этом по каждому маршруту вклад страниц будет разным. Для учета влияния маршрута на ранг страницы в статье предлагается модификация алгоритма *PageRank*. С этой целью используется расстояние сопротивления. Расстояние сопротивления позволяет выявить скрытое отношение между вершинами графа. Его можно считать результирующим, объединяющим в себе всевозможные пути между вершинами. При этом вклад вершины, который распространяется на другие страницы, обратно пропорционален расстоянию сопротивления.

Список литературы

1. Jones K. S. Automatic summarizing: the state of the art // Information Processing and Management. 2007. V. 43. N 6. P. 1449—1481.
2. Alguliyev R. M., Alguliyev R. M. Effective summarization method of text documents // Proc. of the 2005 IEEE/WIC/ACM International Conference on Web Intelligence (WI'05). France. 2005. P. 264—271.
3. Агулиев Р. М., Алыгулиев Р. М. Новый метод резюмирования текстовых документов и оценка результата классификации в трех аспектах // Телекоммуникации. 2006. № 3. С. 7—16.
4. Alguliyev R. M. Using the F-measure as similarity measure for automatic text summarization // Вычислительные технологии. 2008. Т. 13. № 3. С. 5—14.

5. **Alguliev R. M., Aliguliyev R. M., Bagirov A. M.** Global optimization in the summarization of text documents // Automatic Control and Computer Sciences. 2005. V. 39. N 6. P. 42–47.

6. **Aliguliyev R. M.** A Novel Partitioning-Based Clustering Method and Generic Document Summarization // Proc. of the 2006 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT 2006 Workshops) (WI-IATW'06). Hong Kong. 2006. P. 626–629.

7. **Aliguliyev R. M.** Automatic document summarization by sentence extraction // Вычислительные технологии. 2007. Т. 12. № 5. С. 5–15.

8. **Alguliev R. M., Alyguliev R. M.** Automatic text documents summarization through sentences clustering // Journal of Automation and Information Sciences. 2008. V. 40. № 9. P. 53–63.

9. **Brin S., Page L.** The anatomy of a large-scale hyper-textual web search engine // Computer Networks and ISDN systems. 1998. Vol. 30. NN 1–7. P. 107–117.

10. **Diligenti M., Gori M., Maggini M.** A unified probabilistic framework for web page scoring systems // IEEE Transactions on Knowledge and Data Engineering. 2004. Vol. 16. N 1. P. 4–16.

11. **Ingongngam P., Rungsawang A.** Topic-centric algorithm: a novel approach to Web link analysis // Proc. of the 18th International Conference on Advanced Information Networking and Applications (AINA'04). Japan. 2004. V. 2. P. 299–301.

12. **Алыгулиев Р. М.** Оптимизационная модель ранжирования Web-страниц // Системы управления и информационные технологии. 2006. № 3 (25). С. 4–7.

13. **Алыгулиев Р. М., Алыгулиев Р. М.** Ранжирование web-страниц с использованием взаимной информации между гиперссылками // Проблемы управления. 2007. № 4. С. 24–29.

14. **Klein D. J.** Resistance-distance sum rules // Croatica Chemica Acta. 2002. V. 75. № 2. P. 633–649.

15. **Bapat R. B., Gutman I., Xiao W.** A simple method for computing resistance distance // Z. Naturforschung — A Journal of Physical Sciences. 2003. V. 58a. NN 9–10. P. 494–498.

16. **Balaban A. T., Klein D. J.** Co-authorship, rational Erdos numbers, and resistance distances in graphs // Scientometrics. 2002. V. 55. N 1. P. 5–70.

17. **Jeh G., Widom J.** Simrank: a measure of structural-context similarity // Proc. of the 8th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. Canada. 2002. P. 538–543.

18. <http://duc.nist.gov>

19. **Erkan G., Radev D. R.** Lextrank: graph-based lexical centrality as salience in text summarization // Journal of Artificial Intelligence research. 2004. V. 22. P. 457–479.

20. **Liu Y., Wang X., Zhang J., Xu H.** Personalized PageRank based multi-document summarization // Proc. of the First IEEE International Workshop on Semantic Computing and Systems (WSCS2008). China. 2008. P. 169–173.

УДК 004.55

А. Н. Данчул, д-р техн. наук, проф.,
зав. кафедрой,
Российская академия государственной
службы при Президенте РФ,
e-mail: danch@ur.rags.ru,
А. П. Новиков, инженер,
Учебно-методический центр
по информационно-аналитической работе
департамента образования г. Москвы

Представление знаний о предметной области на основе гносеологической модели

Изложены концепции представления знаний о предметной области на основе гносеологической модели. Рассмотрено формирование сетевой структуры в гносеологической модели, которое приводит к представлению знаний о предметной области в виде квазиграфа. Отмечается, что созданная оболочка экспертной системы, реализующая предложенные концепции, обладает необходимой оперативностью актуализации знаний.

Ключевые слова: гносеологическая модель, представление знаний, сетевая структура.

Концепции представления знаний в моделировании предметной области

Предметная область — это часть реального мира, интересующая субъекта, который осуществляет некоторую деятельность (управленческую, про-

ектную, познавательную и т. п.). Как правило, при рассмотрении предметной области выделяется множество объектов, связанных между собой отношениями. В базах знаний используются структурные модели предметной области, в которых понятия-объекты и понятия-отношения ассоциируются с объектами и отношениями предметной области.

Конкретный смысловой оттенок "разграничение понятий на понятия-объекты и понятия-отношения" определяет весь процесс традиционного моделирования предметной области. В соответствии с целенаправленностью отбирается множество понятий-объектов, которые признаются существенными. Они отображаются узлами ориентированного псевдографа [1]. Каждое понятие-объект переопределяется в соответствии с целенаправленностью. Существенными признаются те проявления понятия-объекта, которые характеризуют его место в общей стройной системе понятий предметной области. Понятия-отношения каждого узла, соотносимого с понятием-объектом, отображаются именованными инцидентными дугами.

Смысловой оттенок "разграничение понятий предметной области на понятия-объекты и понятия-отношения" может по-разному учитываться при построении модели предметной области. Так, к примеру, в нотации Чена [2, с. 183] узлам могут соответствовать и "понятия-объекты", и "понятия-отношения". Неименованные дуги на ориентированном псевдографе при этом соединяют узлы-отношения с теми узлами-понятиями, кото-

рые связаны этим отношением, что придает сетевой модели предметной области двудольность.

Традиционно предметная область априори рассматривается как сложная целенаправленная система [3, с. 24]. Целенаправленные модели, как и модели, где однородная ориентация представлена однородностью отношений (в том числе классификационная модель "конкретизация \Rightarrow \Rightarrow обобщение"), являются вариантами реализации моделей, в которых при отборе понятий и определении каждого понятия предполагается наличие ориентации, характеризуемой двумя противоположными направлениями, на практике именуемыми двумя терминами (антонимами).

В традиционном моделировании предметной области используются понятия из гносеологической модели и методические приемы, сформированные в развитом механизме познания. *Гносеологической моделью* будем называть модель, которая формируется у субъекта, познающего реальный мир. Первично познание реального мира, моделирование предметной области вторично.

Представление знаний является завершением конкретных процессов познания, отражением ранее непознанных закономерностей реального мира. Представление знаний можно рассматривать и как процесс, и как результат моделирования реального мира на основе используемых методов и средств моделирования. Авторы предлагают выявить методы моделирования, являющиеся эффективными в познании реального мира на модели, в которой познающий субъект находится в пространстве единства реального мира, данного субъекту в ощущениях. Пространство реального мира объективно своим единством. Реальный мир остается единым, пока в нем: ничего не выявлено, не выделено и не разграничено. Всякое разграничение, выявление и выделение сущностей в пространстве единства реального мира субъективно, диктуется практической целесообразностью субъекта, выполняется субъектом и отражается в субъективной модели.

Выявление экземпляров сущностей и их отражение в гносеологической модели

Осуществляемые в процессе познания выявленные экземпляры сущностей и их отражение в гносеологической модели состоят из пяти операций.

1. Фиксация множества проявлений, которые характеризуют вновь выявляемую ситуацию, требующую выработки целенаправленных действий.

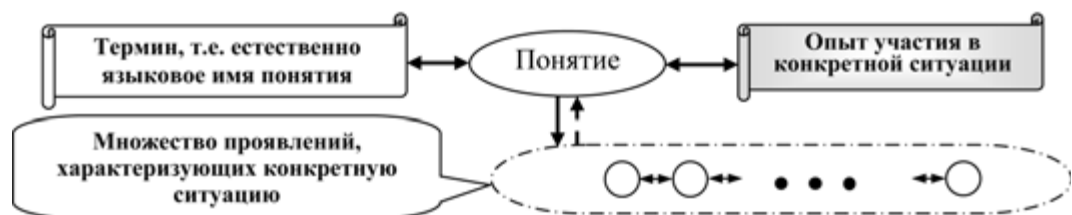
2. Выделение в единстве реального мира не выявленного ранее экземпляра сущности. В соответствии с законом единства формы и содержания множество проявлений, характеризующих ситуацию, определяет единственный экземпляр сущности. Экземпляр сущности, выделяемый субъектом, существует для субъекта в пространстве единства реального мира, если он воспринимается множеством проявлений, распознаваемых субъектом (может быть выявлен субъектом), и распознается в соответствии с каким-либо ранее сформированным определением в субъективной модели. Если экземпляр сущности проявляется, но не распознается, то субъект может его выявить, признать существенным и определить. Сущность, различаемая субъектом, — это абстракция, зафиксированная понятием в гносеологической модели, объединяющая экземпляры сущности и всякий раз ассоциируемая с конкретным экземпляром сущности. Формирование сущности базируется на практически целесообразном признании существования множества экземпляров конкретной сущности. Более подробно это пояснено далее.

3. Формирование понятия в знаниях о реальном мире (в гносеологической модели), ассоциируемого с конкретным экземпляром сущности.

4. Формирование определения соответствующего понятия (фиксация двунаправленной связи между новым понятием и множеством его проявлений). Связь между понятием и множеством его проявлений реализует возможность перехода (при общении) между понятием и множеством его проявлений в равной степени легко в одном и другом направлении. Поэтому фиксируемая связь должна быть двунаправленной. Рассмотрение в определении понятия множества проявлений подразумевает равноценность проявлений, принимаемую как адекватность первого типа. Типы адекватностей описаны далее.

5. Именованное сформированное понятие термином. В естественном языке только малая часть понятий имеет имя, выраженное простым уникальным термином (одним словом). Поскольку существенным является наличие понятия в гносеологической модели, а отсутствие сформиро-

Рис. 1. Определение (однообразная конструкция), соответствующее выявленному в познании экземпляру сущности



ванного для него имени не препятствует познанию и мышлению (в практике именуемому образным мышлением), при рассмотрении модели познания реального мира операция именованя понятия термином остается на втором плане.

Множественное последовательное выполнение механизмом познания этих пяти операций формирует множество никак не связанных между собой однообразных конструкций (рис. 1), в которых понятие связано с множеством его проявлений. На рис. 1—3 и далее также отражены связи, фиксирующие равноценность проявлений из множества, рассматриваемую ниже как адекватность первого типа.

Двунаправленность определений понятий в гносеологической модели

Во множество распознаваемых проявлений попадают понятия, ассоциируемые с ранее выявленными экземплярами сущностей. Рассмотрим ситуацию, когда понятие, ассоциируемое со вновь выявляемым экземпляром сущности, имеет множество проявлений (M_2), в котором можно выделить подмножество проявлений, равное множеству проявлений понятия (M_1), ассоциируемого с ранее выявленной сущностью, $M_1 \subseteq M_2$ (рис. 2). При фиксации определения понятия, ассоциируемого со вновь выявляемым экземпляром сущности, подмножество проявлений M_1 в множестве проявлений M_2 может быть заменено новым экземпляром ранее зафиксированного понятия, как это показано на рис. 3.

Определение понятия содержит определяемую часть, которую будем называть *меткой понятия*, и множество проявлений (определяющую часть), которое будем называть *содержанием понятия*. Таким образом, при "замене", показанной на рис. 3, меняется способ отображения понятия 1 в определении (однообразной конст-

рукции) понятия 2: вместо своего содержания оно отображается меткой. При множественных "заменах" каждое понятие в гносеологической модели представлено множеством экземпляров метки понятия, которое своим единством в гносеологической модели ассоциируется с конкретным экземпляром сущности.

На рис. 4 видно, что понятие (единство экземпляров метки этого понятия) при его размещении в гносеологической модели, с одной стороны, определено множеством своих проявлений (содержанием понятия в конкретном определении), с другой стороны — связано со множеством меток других понятий, для выявления и идентификации которых оно служит проявлением. Множество понятий, для выявления и идентификации которых конкретное понятие служит проявлением, составляет неотъемлемую часть его определения, так как активно используется в познании этого понятия, а также актуализации связанных с ним знаний. Таким образом, определение понятия в

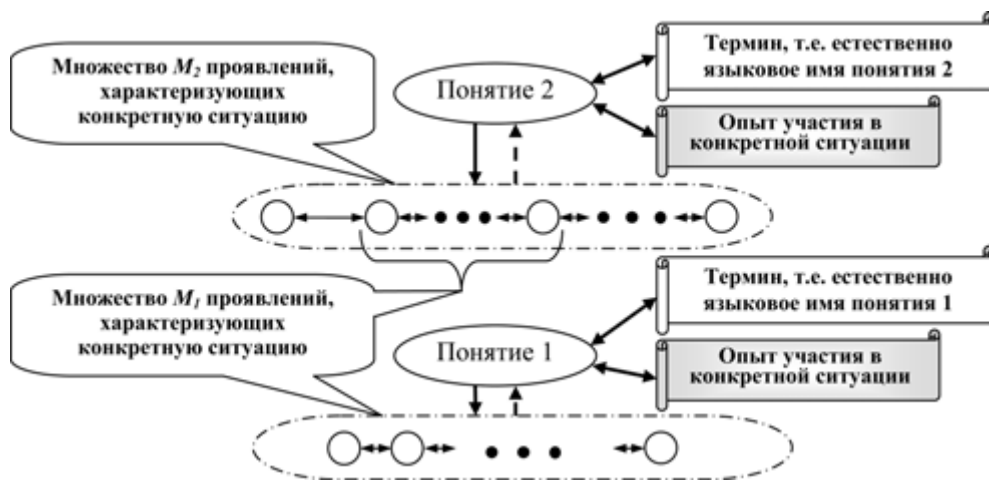


Рис. 2. Фиксируемое определение понятия 2 и ранее зафиксированное определение понятия 1 до "замены"

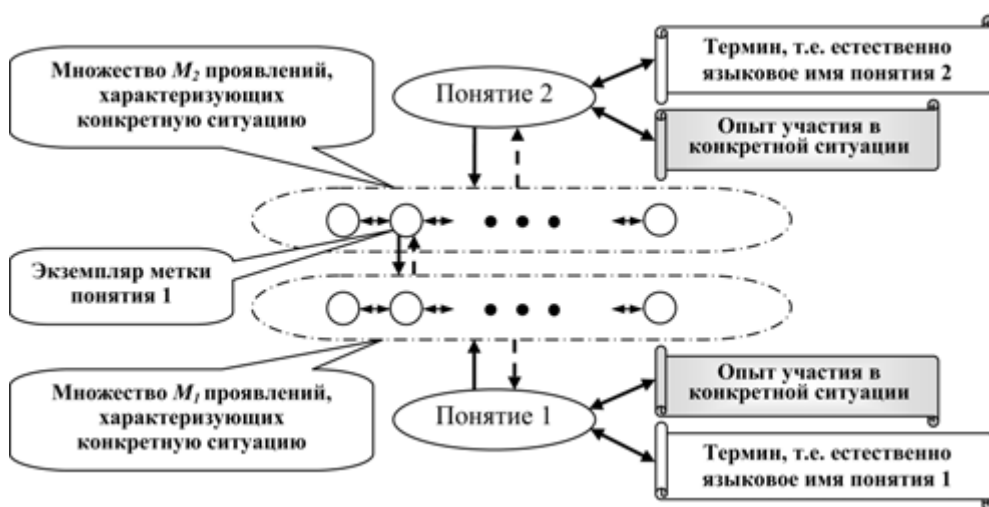


Рис. 3. Фиксируемое определение понятия 2 и ранее зафиксированное определение понятия 1 после "замены"

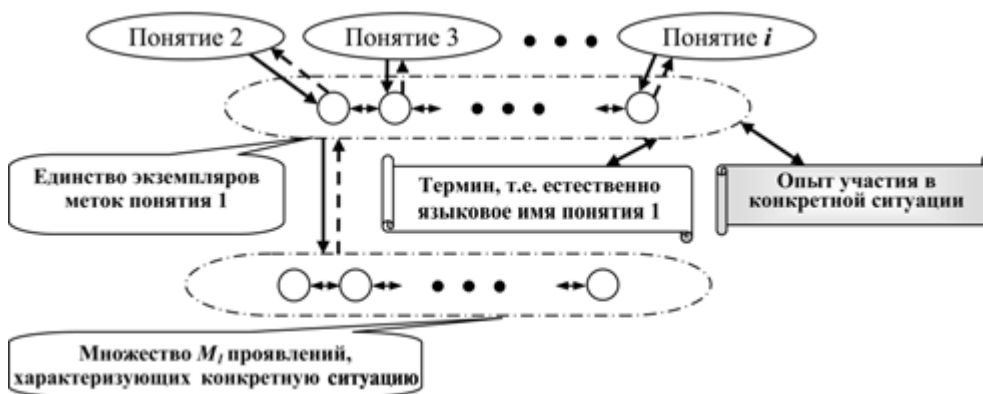


Рис. 4. Двухнаправленность определения понятия ("единства экземпляров меток понятия") в гносеологической модели

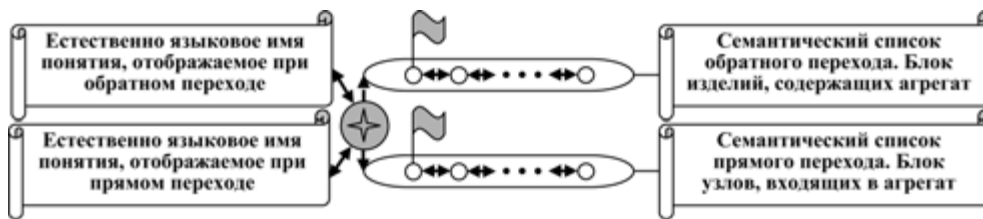


Рис. 5. Двухнаправленное определение понятия с указателями на естественно языковые имена понятия

гносеологической модели становится двухнаправленным, т. е. имеющим двухнаправленное содержание понятия. Естественно предположить, что в однообразной конструкции части двухнаправленного содержания понятия представлены одинаково, т. е. множество понятий, для выявления и идентификации которых конкретное понятие служит проявлением, представлено в однообразной конструкции множеством меток этих понятий, как это показано на рис. 5. В качестве примера двухнаправленного определения рассматривается агрегатная модель (рис. 5 и рис. 6, см. четвертую сторону обложки), в которой центральное в определении понятие соотносится с агрегатом. Метки понятий, соединенные исходящими дугами с агрегатом, соотносятся с изделиями, в которые они входят, а входящими дугами — с техническими узлами, из которых они состоят. Различение двух противоположных направлений однородной ориентации сказывается и на именах понятий. При выходе в общении на понятие по разным направлениям локации могут использоваться разные имена (такая возможность отражена на рис. 5).

Одно и то же конкретное понятие, соотносимое с сущностью, применяется в моделях многих предметных областей, где оно может быть неодинаково определено для отражения разных аспектов этого понятия (разные аспекты в определениях понятия — это разные подмножества проявлений). Из этого следует, что гносеологическая

модель составлена из множества определений каждого из множества понятий. "Содержание понятия" (два множества проявлений) является не тиражируемой частью конкретного двенаправленного определения конкретного понятия, выделяется в однообразную регулярную структуру, представляемую в виде двух семантических списков (см. рис. 5), и размещается в функциональной ячейке гносеологической модели. Под функциональной ячейкой гносеологической модели будем понимать контейнер, используемый для размещения однообразной регулярной структуры (с неограниченными диапазонами параметров этих структур [4]).

Формирование сетевой структуры гносеологической модели

Под сетевой структурой будем понимать всякую структуру с ориентированными связями, отличительной особенностью которой является наличие хотя бы одной структурной единицы (из множества и многообразия структурных единиц), имеющей более одной входящей и более одной исходящей связи.

Однородная ориентированная связь между понятиями (представленная на рис. 2—5 сплошной стрелкой) характеризует последовательность в определении одного понятия через другие: *понятия, ассоциируемые с позже выявленными экземплярами сущностей, определены через понятия, ассоциируемые с ранее выявленными экземплярами сущностей*. В сетевой структуре гносеологической модели также формируются в определении понятий, выраженных единством своих экземпляров, и обратные дуги, обозначаемые на рис. 2—4 и далее штриховыми стрелками. Развитие познания и гносеологической модели ведет к переопределению однородной ориентированной связи между понятиями: *менее устоявшиеся понятия определены через более устоявшиеся понятия*.

Рассматривая содержание понятия (не тиражируемую часть определения) в его двухнаправленном определении (рис. 5) как однообразную регулярную структуру, мы установили, что она состо-

ит из двух множеств: множества меток понятий, которые являются проявлениями конкретного понятия, и множества меток понятий, для выявления и идентификации которых конкретное понятие служит проявлением. Каждая метка каждого понятия из множества меток понятий, составляющих однообразную регулярную структуру, связана с конкретной ячейкой гносеологической модели, в которой помещено двунаправленное содержание понятия, соответствующее этой метке. Метки понятий, которые являются проявлениями рассматриваемого понятия, связаны исходящими связями со своими ячейками. Метки понятий, для выявления и идентификации которых рассматриваемое понятие служит проявлением, связаны входящими связями со своими ячейками. Структурной единицей сетевой структуры является ячейка гносеологической модели.

Самоорганизация познания

Практика субъекта инициирует познание не только отражения реального мира, но и своей самоорганизации. Самоорганизация познания нацелена на выявление и обобщение возможностей, а также создание и повышение эффективности способностей по переносу опыта предыдущих событий на события реального времени.

Важнейшим инструментом самоорганизации познания является типизация. *Типизация* — это познавательная процедура обоснованного сведения многообразия объектов к небольшому числу их типов, с последующим ассоциированием типа с каждым конкретным объектом. Типизация при самоорганизации познания реализуется системой различных субъективных методических приемов, влияющих как на механизм логического мышления, так и на механизм отражения реального мира.

В гносеологической модели типизация для самоорганизации познания распространяется на структурные единицы и связи:

- *типизация однообразных структурных единиц* модели предметной области — это определение и именование понятий, соотношенных с сущностями. Практика позволяет устанавливать субъективную эквивалентность некоторых экземпляров сущностей, признавая их представителями одной и той же сущности. Согласно с практикой, целесообразно отбирается подмножество (из множества выявленных субъектом и отраженных в гносеологической модели) экземпляров сущностей, которые признаются экземплярами (субъективными представителями) одной сущности. Для экземпляров этой сущности отбираются общие проявления, которые ставятся в соответствие сущности, объединяющей группу экземпляров. В гносеоло-

гической модели формируется определение нового понятия, соотносимого с сущностью и ассоциируемого с каким-либо одним экземпляром сущности;

- *типизация связей* в модели предметной области — это выявление практически целесообразных типов адекватностей, которое обеспечивает различие характеристик для типизации смысловых оттенков, отображаемых связями в гносеологической модели. Многообразии типов смысловых оттенков, которые могут быть использованы при моделировании предметной области, определяется множеством типов адекватностей, которые учитываются оболочкой в представлении знаний. Неименованные типизированные связи в гносеологической модели должны резервироваться двунаправленными, так как реализуют возможности перехода (при общении), которые в конкретном случае могут не назначаться, назначаться однонаправленными или двунаправленными. Структурная семантика по типу адекватности выражается структурным подобием (примеры см. на рис. 6, четвертая сторона обложки) в сетевой структуре гносеологической модели, не зависящим от конкретики предметной области. Структурная семантика по типам адекватностей достаточно наглядна на формально-логической модели. Наглядность смысловых оттенков, выражаемых структурной семантикой по типам адекватностей, достаточно высокая, если структурная семантика конкретизирована вместе с конкретизацией модели предметной области.

Моделирование предметной области на концептуальном уровне

По мнению проф. Е. С. Кузина: "В настоящее время достижения в направлении представления в компьютерных системах плохо формализуемых областей знаний весьма скромны. Это объясняется тем, что разработчики ... программного обеспечения пытаются описывать сложные проблемные среды, в которых решаются информационно сложные задачи, посредством однообразных регулярных структур (реляционных таблиц, объектов, фреймов, продукций и т. п.), слишком примитивных для отображения всего многообразия смысловых оттенков проблемной среды, хотя и удобных для последующей обработки информации" [5].

Сформулированное в [6] определение понятия *интенциональная семантическая сеть* (ИСС) позволяет предположить его эквивалентность понятию *гносеологическая модель*. ИСС — это непрерывно ассоциируемая с реальным миром модель предметной области, в которой структурированно представлены понятия и их проявления, так же как в

реальном мире сущности присутствуют своими проявлениями. ИСС представима множеством определений каждого из множеств понятий, в котором части каждого определения (метка понятия и два множества меток понятий) соединены типизированными связями, соответствующими (рассматриваемым ниже) типам адекватностей.

Рассмотрим основные типы адекватностей на ориентированном псевдографе, воспользовавшись отличительной особенностью сетевых структур. Структурной единицей сетевой структуры гносеологической модели является ее ячейка.

Неименованные типизированные связи сетевой структуры гносеологической модели продолжим внутрь ячеек гносеологической модели (до структурных единиц ячеек, которыми могут быть метки понятий или их множества). Структурная единица ячейки гносеологической модели (метка понятия или множество меток понятий) определяется в момент продолжения конкретной связи внутрь этой ячейки. Структурная единица "множество меток понятий" для соответствующих связей, продолжаемых внутрь ячейки гносеологической модели, становится множеством структурных единиц. Неоднозначность структурных единиц в ячейках гносеологической модели: *для одной и той же связи это может быть и метка понятия (в одной ячейке), и множество меток понятий (в другой ячейке)*, может быть разрешена структурированием каждого множества меток понятий. Каждое множество меток понятий, рассматриваемое как структурная единица, объединяется в цепной список с двунаправленной связью. Первой метке понятия в этом цепном списке присваивается особая роль: все связи со структурной единицей "множество меток понятий" перенаправлены на метку понятия, стоящую первой в организованном цепном списке.

Получаемую в итоге диаграмму будем называть квазиграфом, который в этом случае определяется как ориентированный псевдограф с неименованными узлами и неименованными типизированными дугами. Узлом квазиграфа является структурная единица "экземпляр метки понятия", которую будем называть микроструктурой.

Адекватность — это назначаемое типизированное сходство в гносеологической модели нескольких понятий или нескольких экземпляров одного понятия, определяющее их структурную семантику (примеры см. на рис. 6, четвертая сторона обложки) и собственно семантику сущностей реального мира. На формально-логическом уровне адекватность выражается типизированной (по одному из структурных свойств) двунаправленной связью внутри или между ячейками гносеологической модели, которая устанавливается на части одной или двух ячеек гносеологи-

ческой модели, соответствующих структурным единицам.

На формально-логическом уровне:

- адекватность 1-го типа — это принадлежность узла к семантическому списку прямого или обратного перехода (на рис. 6, *a* — стрелки коричневого цвета, см. четвертую сторону обложки);
- адекватность 2-го типа — отражается связью узла с его семантическими списками прямого и обратного переходов (на рис. 6, *a* — сплошные и штриховые стрелки синего цвета показывают порядок в списке адекватных);
- адекватность 3-го типа — это сходство узлов, имеющих переход на один и тот же семантический список прямого перехода (на рис. 6, *b* — стрелки сиреневого цвета);
- адекватность 4-го типа — это сходство узлов, имеющих переход на один и тот же семантический список обратного перехода (на рис. 6, *b* — стрелки красного цвета);
- адекватность 5-го типа — это сходство узлов, к которым подсоединена одна и та же висячая вершина. Им соответствует одно и то же имя в базе имен [7] (на рис. 6, *z* — стрелка фиолетового цвета).

Приведем примеры адекватностей 3-го и 4-го типов. Если из равных наборов радиодеталей изготовлены разные агрегаты, то эти агрегаты адекватны по адекватности 3-го типа. Если несколько агрегатов взаимозаменяемы (имеют равные эксплуатационные характеристики и посадочные размеры), но изготовлены из разных наборов радиодеталей, то эти агрегаты адекватны по адекватности 4-го типа. Приведенные примеры показывают, что ячейка гносеологической модели имеет двудольный характер. В каждой доле ячейки размещается множество меток понятий, каждая из которых не имеет прямых связей (связей внутри ячейки) с метками понятий другой доли. Две доли ячейки объединяются меткой понятия, которая задается этим определением (определяется одновременно двумя множествами проявлений, содержащимися в этих долях ячейки). Примеры показывают, что метки понятий могут определяться содержанием понятия из одной и той же доли одной ячейки, но при этом вторая часть содержания понятия у них разная. Таким образом, одна и та же доля ячейки является составной частью разных ячеек.

При представлении знаний *квазиграфом* (новый класс моделей представления знаний) каждый экземпляр метки каждого понятия (микроструктура) соединен с другими микроструктурами типизированными связями, соответствующими типам адекватностей.

Рассмотрим пример отображения квазиграфом фрагмента модели знаний, представленного ориентированным псевдографом на рис. 7, а (см. четвертую сторону обложки):

1. Формируются семантические списки всех узлов (см. рис. 7, б). Узлы в каждом списке объединяются в цепной список двунаправленными связями (стрелки коричневого цвета — адекватность 1-го типа). Для каждого узла указываются семантические списки прямого и обратного переходов.

2. Каждый узел каждого семантического списка (см. рис. 7, в) соединяется со своими списками прямого и обратного переходов в соответствии с тем, как это показано на диаграмме рис. 7, б (сплошные и штриховые стрелки синего цвета — адекватность 2-го типа).

3. При просмотре каждого из семантических списков выявляются те, на которые направлено более одного указателя. Узлы, от которых выявлен указатель на один и тот же семантический список прямого перехода, объединяются в цепной список с двунаправленными связями (стрелки сиреневого цвета — адекватность 3-го типа), как это выполнено на диаграмме рис. 7, г.

4. Узлы, от которых выявлен указатель на один и тот же семантический список обратного перехода, объединяются в цепной список с двунаправленными связями (стрелки красного цвета — адекватность 4-го типа), как это выполнено на диаграмме рис. 7, д.

5. Узлы квазиграфа, имеющие указатели на одно и то же имя в базе имен, объединяются в цепной список с двунаправленными связями (адекватность 5-го типа). На диаграмме рис. 7, з эта операция не отражена.

Во множестве концепций представления знаний в гносеологической модели учитываются принятые в настоящее время концепции моделирования предметных областей. Субъект, познающий единство реального мира, может выполнить расширение концепций моделирования предметных областей только привлечением концепций представления знаний в гносеологической модели. Например, описанные в статье концепции представления знаний в гносеологической модели позволили получить неограниченные диапазоны параметров сетевой структуры [4]; независимость времени реакции компьютера в интерактивном взаимодействии от размеров базы знаний

[7]; возможность создания редактора семантических сетей, доступного массовому пользователю, и другие преимущества оболочки интеллектуальных систем.

Состояние практических работ

Изложенные в статье концепции представления знаний были использованы при создании ОЭС FICONOCS [8], которая работает с детерминированными знаниями, содержит редактор классификационных сетей, доступный массовому пользователю, прошла государственную регистрацию и к ней выпущена техническая документация. Области применения этой ОЭС определяются подклассом нового класса моделей интеллектуальных систем, свободным от проникновения других оболочек [9]. Возможность достижения с помощью этой оболочки необходимой оперативности актуализации знаний в базе знаний делает такие интеллектуальные системы на ее базе экономически эффективными. Опыт внедрения ОЭС FICONOCS в различных предметных областях реального производства продемонстрировал универсальность редактора классификационных сетей.

Список литературы

1. Харари Ф. Теория графов / Под. ред. Г. П. Гаврилова. М.: МИР, 1973.
2. Коннолли Т., Бегг К., Страчан А. Базы данных: проектирование, реализация и сопровождение. Теория и практика. 2-е изд. / Пер. с англ. М.: Издат. дом "Вильямс", 2001. 1120 с.
3. Дьячко А. Г. Математическое и имитационное моделирование производственных систем. М.: МИСИС, 2007. 540 с.
4. Данчул А. Н., Новиков А. П. Конструктивный взгляд на использование семантических сетей в информационных системах поддержки решений // Матер. научно-практич. конф. "Ситуационные центры и перспективные информационно-аналитические средства поддержки принятия решений". Москва, 25—27 апреля 2007 г. М.: Российская академия государственной службы при Президенте РФ, 2007.
5. Кузин Е. С. Информационно-сложные задачи и технологии их решения // Новости искусственного интеллекта. 2003. № 1.
6. Абдикеев Н. М. Проектирование интеллектуальных систем в экономике: Учеб. М.: Экзамен, 2004.
7. Дьячко А. Г., Новиков А. П. Организация "базы имен" объектов и отношений семантической сети для неконтролируемого (бесконтрольного) ее использования // Матер. 2-й научн. конф. "Автоматизация в промышленности". Москва, 14 апреля 2008 г. М.: Институт проблем управления РАН, 2008.
8. Комплекс программ "FICONOCS". М.: ЦНИИ ЭИСУ, 1988.
9. Данчул А. Н., Новиков А. П. Самопрограммируемость интеллектуальных систем // Матер. 2-й научн. конф. "Автоматизация в промышленности". Москва, 14 апреля 2008 г. М.: Институт проблем управления РАН, 2008.

УДК 004.031.43

Э. А. Мухачёва, д-р техн. наук, проф.,
А. Р. Заминова, аспирант,
Уфимский государственный
авиационный технический университет,
e-mail: alina-z@mail.ru

Исследование эффективности эволюционных алгоритмов для решения задач дискретной оптимизации на примере задачи коммивояжера

Представлены методики использования эволюционных алгоритмов для решения задачи коммивояжера. В качестве алгоритмов рассматривались генетический алгоритм и эволюционные стратегии на примере $(\mu + \lambda)$ -ES, (μ, λ) -ES и частного случая — стратегии $(1 + 1)$ -ES. Приведены результаты численного эксперимента, подтверждающие эффективность простой эволюционной стратегии.

Ключевые слова: генетический алгоритм, эволюционные стратегии $(\mu + \lambda)$ -ES, (μ, λ) -ES и $(1 + 1)$ -ES, задача коммивояжера.

Введение

Рассмотрим базовую формулировку задачи коммивояжера. Дан полный неориентированный граф $G = (V, E)$, каждому из ребер которого $(u, v) \in E$ сопоставляется неотрицательная целочисленная стоимость $c(u, v)$. Необходимо найти путь минимальной длины (стоимости), проходящий в точности по одному разу через каждую вершину графа. Будем рассматривать частный вид задачи коммивояжера, где точка, являющаяся начальной и конечной для маршрута, заранее задана.

Как известно, задача коммивояжера представляет собой задачу оптимизации из класса NP-трудных проблем. Несмотря на результативность переборного метода "ветвей и границ" для достаточно большого числа пунктов, часто возникает задача коммивояжера, которую требуется решить в реальном времени. В таких случаях преимущественно применяются алгоритмы двух типов:

- алгоритмы, основанные на полном переборе с возможными накладываемыми ограничениями, сужающими область перебора;

- эвристические алгоритмы случайного поиска в заданном пространстве решений.

Существование большого числа практических оптимизационных задач, относящихся к классу NP-трудных, размерность которых не допускает возможности отыскания решений точными методами, способствовало развитию эвристических методов.

Одним из наиболее гибких и широко используемых алгоритмов эвристического поиска является *генетический алгоритм* (ГА), представляющий собой общую схему, основанную на схеме простого ГА, предложенной Д. Холландом [1] в 1975 г. и расширенную впоследствии до многофункционального математического аппарата [2, 3]. Благодаря гибкой настраиваемой структуре данный алгоритм может быть использован для поиска в пространствах параметров различных типов — как в пространствах бинарной и вещественной размерности [4], так и в целочисленных пространствах (например, при решении задач дискретной оптимизации). Возможность его применения для решения задач такого широкого круга достигается благодаря непривязанности общей схемы алгоритма к типу искомого решения и набору операторов кроссовера и мутации, позволяющих реализовывать различные стратегии конструирования и изменения решений.

Общая схема ГА может быть представлена последовательностью следующих этапов [5].

1. Предварительный этап заключается в выборе способа кодирования искомого решения и формирования начальной популяции решений. Далее, на каждой итерации алгоритма, пока не достигнуто условие останова, выполняются этапы 2—4.

2. Применение с заданной вероятностью оператора кроссовера, осуществляющего конструирование нового решения из уже существующих (чаще всего двух).

3. Применение с заданной вероятностью операторов случайного изменения решения — операторов мутации, инверсии, транслокации и сегрегации, осуществляющих случайные изменения по заданной схеме над некоторым решением.

4. Применение оператора селекции, выполняющего отбор из всех имеющихся на данной итерации решений в поколение для генерации решений на следующей популяции.

Существуют различные варианты критериев останова, в каждом конкретном случае выбор критерия определяется прежде всего особенностями поискового пространства и целевой функции (ЦФ). К наиболее популярным критериям останова относятся: выполнение заданного числа итераций, достижение заданной окрестности оптимума и отсутствие улучшения значения ЦФ по популяции, превышающего рекордное значение, за фиксированное число итераций.

Еще одним из направлений развития эволюционных алгоритмов (ЭА) стали эволюционные стратегии, предложенные впервые Реченбергом в 1973 г. [6]. Основными отличиями эволюционных стратегий от ГА являются наличие строго заданной простой схемы отбора решений и использование исключительно унарных эволюционных операторов генерации решений, таких как оператор мутации [7].

Общая схема эволюционных стратегий ($\mu + \lambda$, λ)-ES, (μ , λ)-ES может быть представлена в виде последовательности этапов, выполняемых на каждой итерации алгоритма над множеством решений текущей популяции.

1. Предварительный этап, заключающийся в выборе способа кодирования искомого решения и формирования начальной популяции из μ решений. Далее, на каждой итерации алгоритма, пока не достигнуто условие останова, выполняются этапы 2–4.

2. Применение с заданной вероятностью операторов случайного изменения решения для генерации λ новых решений.

3. Применение оператора селекции, выполняющего отбор μ лучших решений в новую популяцию из решений, сгенерированных на данной итерации, и решений, используемых для генерации, — для ($\mu + \lambda$, λ)-ES стратегии и отбор μ лучших из λ сгенерированных решений — для (μ , λ)-ES стратегии.

Далее будет рассмотрен подробнее частный случай стратегии ($\mu + \lambda$, λ)-ES при $\mu = \lambda = 1$, такую стратегию будем обозначать как (1 + 1)-ES. П. Борисовский и А. Еремеев показали, что при условии монотонности по вероятности оператора мутации алгоритм (1 + 1)-ES предпочтительнее других эволюционных алгоритмов [8]. Однако наличие этого условия в операторах мутации трудно подтвердить.

Исторически сложилось, что задача коммивояжера с применением эволюционных алгоритмов решалась с точки зрения двух основных подходов:

- когда в качестве хромосом рассматривали непосредственно наборы узловых точек (или ребер-путей между узловыми точками) [9];
- когда в качестве хромосомы использовали набор эвристик, последовательно используемых для построения маршрута [10].

В данной статье будет рассмотрен первый подход как предпочтительный с точки зрения значения выбора операторов кроссовера и мутации для скорости поиска оптимального решения и получения заведомо корректного решения в соответствии со спецификой решаемой задачи.

Эволюционные операторы, применимые для решения задач дискретной оптимизации

Рассмотрим задачу коммивояжера как задачу дискретной оптимизации и будем искать ее решение с использованием алгоритма, построенного на базе соответствующих эволюционных операторов.

Пусть задан набор из n городов $K = (k_1, \dots, k_n)$. задается матрица весов переходов между исходными городами $S(k_i, k_j)$, где k_i, k_j — некоторые города из списка городов K . Все элементы матрицы весов переходов между городами считаются заданными, а граф полным. В случае, если это не так, проводится пересчет незаданных расстояний между городами согласно алгоритму Дейкстры [11], и далее задача решается на полученном полном графе.

Необходимо найти маршрут индексов посещаемых городов вида $C = (c_0, \dots, c_{n-1})$, где начальная точка маршрута $c_0 = 0$ задана и является также для маршрута конечной, а остальные точки маршрута представляют собой последовательность неповторяющихся целочисленных элементов в заданном диапазоне. В качестве минимизируемой ЦФ данной задачи будем рассматривать суммарную длину маршрута, построенного указанным способом.

Таким образом, принимаем, что $c_0 = 0$, $c_j \in [1, n-1]$, $c_j \in Z, j = 1, n-1$. Отвечая за генерацию новых решений, оператор кроссовера во многом определяет эффективность работы алгоритма в целом. Дискретность пространства возможных решений накладывает ограничение на выбор используемого оператора кроссовера, поскольку в генерируемых хромосомах каждый ген должен принимать значения только из определенного набора. Область таких дискретных операторов в общем множестве существующих операторов кроссовера отражена в [12]. Помимо ограничений дискретности для задач генерации маршрута накладываем также дополнительное ограничение корректности маршрута: каждый пункт должен присутствовать в хромосоме и встречаться там не более одного раза.

В соответствии с указанными требованиями для решения рассматриваемой задачи построения маршрута коммивояжера было отобрано три возможных оператора кроссовера (ОК):

- "жадный" оператор кроссовера;
- частично соответствующий оператор кроссовера;
- циклический оператор кроссовера.

Ниже данные операторы кроссовера будут рассмотрены подробнее. Пусть $C_1 = (c_0^1, \dots, c_{n-1}^1)$ и $C_2 = (c_0^2, \dots, c_{n-1}^2)$ — две родительские хромосомы, выбранные для применения оператора кроссовера, где c_j^i — j -й бит закодированной i -й хромосомы. $C = (c_0, \dots, c_{n-1})$ — хромосома, отобранная для проведения мутации, где c_j — j -й бит выбранной хромосомы.

В результате применения рассматриваемых ниже операторов к выбранным хромосомам может быть получен либо один потомок $C' = (c_0', \dots, c_{n-1}')$, либо пара потомков: $C_1' = (c_0^{1'}, \dots, c_{n-1}^{1'})$ и $C_2' = (c_0^{2'}, \dots, c_{n-1}^{2'})$.

Определим функцию $Is_get(C, c_j)$, возвращающую значение *true*, если элемент c_j уже присутствует в хромосоме-маршруте C , и возвращающую значение *false* в противном случае:

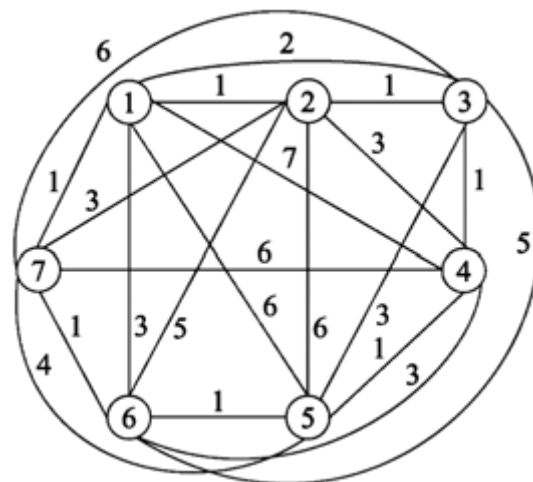
$$Is_get(C, c_j) = \begin{cases} true, & c_j \in C; \\ false, & c_j \notin C. \end{cases}$$

"Жадный" оператор кроссовера для задач дискретной оптимизации. Выполнение "жадного" оператора кроссовера (ЖОК) может быть представлено следующей схемой [5].

1. В конструируемую хромосому C' заносится начальная точка родительских маршрутов $c_0' = c_0$.

2. Из элементов, следующих в первой и второй родительских хромосомах за текущим элементом c_i' из конструируемого потомка, $c_{p_1}^1$ и $c_{p_2}^2$ в случае, если оба эти элемента ранее не были занесены в конструируемую хромосому C' , выбирается тот, расстояние до которого из c_i' минимально. Если ранее в хромосому C' был занесен только один из рассматриваемой пары элементов, то выбирается тот, который не был ранее занесен. Если же оба рассматриваемых элемента уже были предварительно занесены в C' , то выбирается тот элемент из непосещенных, расстояние до которого из текущего элемента c_i' минимально. Выбранный элемент записывается в конструируемый потомок:

$$c_{i+1}' = \begin{cases} c_{p_1}^1, & (Is_get(C', c_{p_1}^1) = false) \& \\ & \& (Is_get(C', c_{p_2}^2) = true); \\ c_{p_2}^2, & (Is_get(C', c_{p_1}^1) = true) \& \\ & \& (Is_get(C', c_{p_2}^2) = false); \\ c_{p_1}^1, & S(c_i', c_{p_1}^1) \geq S(c_i', c_{p_2}^2); \\ c_{p_2}^2, & S(c_i', c_{p_1}^1) < S(c_i', c_{p_2}^2); \\ c_{p^*}, & (S(c_i', c_{p^*}) = \min(\{S(c_i', c_p)\}_{p \in P})) \& \\ & \& (P: \forall p \in P, Is_get(C', c_p) = false). \end{cases}$$



Граф расстояний между городами

3. Шаг 2 выполняется до тех пор, пока все элементы не будут внесены в конструируемую хромосому и индекс текущего элемента не станет конечным: $i = n - 1$.

Рассмотрим применение оператора "жадного" кроссовера на примере.

Пример. Пусть дан граф G_{ex} исходных данных (см. рисунок).

Имеем следующую матрицу расстояний:

$$S_{ex} = \begin{bmatrix} 0 & 1 & 2 & 7 & 6 & 3 & 1 \\ 1 & 0 & 1 & 3 & 6 & 5 & 3 \\ 2 & 1 & 0 & 1 & 3 & 5 & 6 \\ 7 & 3 & 1 & 0 & 1 & 3 & 6 \\ 6 & 6 & 3 & 1 & 0 & 1 & 4 \\ 3 & 5 & 5 & 3 & 1 & 0 & 1 \\ 1 & 3 & 6 & 6 & 4 & 1 & 0 \end{bmatrix}.$$

При решении задачи поиска минимального маршрута по данному графу на некотором шаге с применением "жадного" оператора кроссовера, описанного выше, путем рекомбинации родительских особей были получены решения, представленные в табл. 1 и 2.

В табл. 1 представлен "плохой" пример выполнения оператора, полученный потомок обладает худшим значением целевой функции, чем родители, и в результате проведения селекции не будет включен в популяцию следующего поколения.

В табл. 2 представлен "хороший" пример выполнения оператора, полученный потомок обладает лучшим значением целевой функции, чем родители.

Частично соответствующий оператор кроссовера. Выполнение частично соответствующего оператора кроссовера (ЧСОК) может быть представлено следующей схемой [5].

Таблица 1

Пример выполнения "жадного" оператора кроссовера с ухудшением целевой функции полученного решения

Хромосома	Элементы хромосомы							Значение ЦФ
C_1	0	5	3	2	6	4	1	24
C_2	0	3	6	5	4	1	2	24
C'	0	5	4	1	2	6	3	30

Таблица 2

Пример выполнения "жадного" оператора кроссовера с улучшением целевой функции полученного решения

Хромосома	Элементы хромосомы							Значение ЦФ
C_1	0	5	3	2	4	6	1	18
C_2	0	1	2	5	6	4	3	20
C'	0	1	2	4	3	5	6	11

1. Случайно выбирается разрезающая точка p — реализация дискретной случайной величины с набором равновероятных возможных значений $\{1, \dots, n-1\}$, родительские хромосомы разбиваются на два сегмента:

$$C_1 = (c_0, \dots, c_{p-1}^1 | c_p^1, \dots, c_{n-1}^1)$$

и $C_2 = (c_0, \dots, c_{p-1}^2 | c_p^2, \dots, c_{n-1}^2)$.

2. Левый сегмент хромосомы C_1 переносится без изменений в конструируемую хромосому C'_1 , $C'_1 = (c_0, \dots, c_{p-1}^1)$.

3. Правая часть конструируемой хромосомы C'_1 заполняется следующим образом: если текущий вносимый элемент c_i^2 правой части хромосомы C_2 еще не содержится в конструируемой хромосоме C' , то он вносится в C'_1 на ту же позицию, на которой находился в C_2 ; если же c_i^2 уже присутствует в ранее составленной части C' , то на данную позицию в C'_1 заносится первый из элементов левой части хромосомы C_2 , еще не внесенный в C'_1 :

$$c'_i = \begin{cases} c_i^1, & i < p; \\ c_i^2, & (i \geq p) \ \& \ (Is_get(C'_1, c_i^2) = false); \\ c_{j^*}^2, & (j^* = \min(J)) \ \& \ (J : \forall j \in J, \\ & (Is_get(C'_1, c_j^2) = false) \ \& \ (j < p)); \end{cases}$$

$$i = \overline{0, n-1}.$$

4. Аналогичным образом, используя родительские хромосомы в обратном порядке, для той же разрезающей точки p строится хромосома C'_2 , в левую часть которой переносится левая часть хромосомы C_2 , а правая часть строится из правой части хромосомы C_1 с заменой повторяющихся элементов на не включенные ранее элементы из левой части хромосомы C_1 .

При решении задачи поиска минимального маршрута по рассмотренному в приведенном выше примере графу исходных данных G_{ex} на некотором шаге с применением описанного частично соответствующего оператора кроссовера путем рекомбинации родительских особей были получены решения, представленные в табл. 3.

Циклический оператор кроссовера. Выполнение циклического оператора кроссовера (ЦОК) может быть представлено в виде определенной последовательности шагов [9].

1. Формируется вектор меток циклов $S = (s_0, \dots, s_{n-1})$, все компоненты которого равны нулю, $s_i = 0, i = \overline{0, n-1}$; устанавливаем текущее число циклов $k = 0$.

2. Увеличиваем число циклов $k = k + 1$.

3. Случайным образом из хромосомы C_1 выбираем элемент c_j^1 для начала нового цикла, метка цикла для которого равна нулю, $c_j^1 : s_j = 0$.

4. В векторе меток цикла для индекса выбранного элемента устанавливаем $s_j = k$.

5. Рассчитываем индекс j^* как позицию в C_1 элемента с тем же значением, что и у c_j^2 из C_2 , т. е. находим такое значение j^* , при котором $c_{j^*}^1 = c_j^2$.

6. Если метка цикла для найденного индекса j^* нулевая ($s_{j^*} = 0$), то присваиваем $j = j^*$ и переходим к п. 4, иначе переходим к п. 7.

7. Цикл k считаем сформированным. Случайным образом выбираем одну из используемых хромосом C_1 или C_2 и заносим из нее все элементы, относящиеся к циклу k , в генотип конструируемой хромосомы C' : $c'_i = c_i^1$ либо $c'_i = c_i^2$ для всех i таких, что $s_i = k, i = \overline{0, n-1}$.

8. Если остался еще хотя бы один невыбранный индекс в хромосоме (не посещенный город), т. е. если $\exists i \in \{0, n-1\} : s_i = 0$, то повторяем действия, начиная с п. 2, иначе выход.

При решении задачи поиска минимального маршрута по рассмотренному выше в примере графу исходных данных G_{ex} на некотором шаге с применением циклического оператора кроссовера путем рекомбинации родительских особей были получены решения, представленные в табл. 4.

Таблица 3

Пример выполнения частично соответствующего оператора кроссовера

Хромосома	Элементы хромосомы							Значение ЦФ
C_1	0	3	2	1	4	5	6	18
C_2	0	1	6	4	2	3	5	18
C'_1	0	3	6	4	2	1	5	29
C'_2	0	1	2	3	4	5	6	7

Таблица 4

Пример выполнения циклического оператора кроссовера

Хромосома	Элементы хромосомы							Значение ЦФ
	0	2	3	1	5	4	6	
C_1	0	2	3	1	5	4	6	26
C_2	0	4	1	3	5	2	6	29
C'	0	2	1	3	5	4	6	18
№ цикла	1	2	3	3	4	2	5	

Оператор двухточечной мутации для дискретной оптимизации. Оператор мутации (ОМ) был реализован путем выбора двух позиций в хромосоме и перестановкой элементов хромосомы, находящихся на выбранных позициях.

В результате применения данного оператора мутации к хромосоме C создается потомок вида

$$C' = (c_0, c_1, \dots, c_{p_1-1}, c_{p_2}, c_{p_1+1}, \dots, c_{p_2-1}, c_{p_1}, c_{p_2+1}, \dots, c_{n-1}),$$

где p_1 — реализация дискретной случайной величины с набором равновероятных возможных значений $\{0, \dots, n-2\}$, а p_2 — реализация дискретной случайной величины с набором равновероятных возможных значений $\{p_1+1, \dots, n-1\}$.

Оператор инверсии для дискретной оптимизации. Оператор инверсии (ОИ) был реализован путем выбора двух позиций в хромосоме и перестановкой в обратном порядке всех элементов хромосомы, находящихся между выбранными позициями.

В результате применения данного оператора мутации к хромосоме C создается потомок вида

$$C' = (c_0, c_1, \dots, c_{p_1-1}, c_{p_2}, c_{p_2-1}, \dots, c_{p_1+1}, c_{p_1}, c_{p_2+1}, \dots, c_{n-1}),$$

где p_1 — реализация дискретной случайной величины с набором равновероятных возможных значений $\{0, \dots, n-2\}$, а p_2 — реализация дискретной случайной величины с набором равновероятных возможных значений $\{p_1+1, \dots, n-1\}$.

Оператор транслокации для дискретной оптимизации. Оператор транслокации (ОТ) был реализован путем выбора случайной позиции в хромосоме и перестановкой в обратном порядке всех элементов хромосомы, находящихся правее выбранной позиции.

В результате применения данного оператора мутации к хромосоме C создается потомок вида

$$C' = (c_0, c_1, \dots, c_{p_1-1}, c_{n-1}, c_{n-2}, \dots, c_{p_1+1}, c_{p_1}),$$

где p_1 — реализация дискретной случайной величины с набором равновероятных возможных значений $\{0, \dots, n-2\}$.

Экспериментальная оценка эффективности сконструированных модификаций эволюционных алгоритмов для решения задач дискретной оптимизации

Апробацию построенных модификаций ГА и $(1+1)$ -ES проводили на тестовых наборах данных, представляющих собой матрицы расстояний между заданным числом городов. В качестве первого примера тестовых данных для построенных алгоритмов рассмотрим исходные данные из 10 наборов карт для семи городов с усреднением результатов для каждого набора по 10 запускам. Для нахождения цикла минимального веса использовали ГА с популяцией из 12 решений, 20 генерируемыми поколениями и двумя элитными решениями в каждом поколении. Число особей в поколении и число поколений были выбраны равными 12 и 20 соответственно, так как позволяли достичь посредством применения ГА экспериментально найденного для всего набора данных оптимума — значения ЦФ, равного 23,66. В качестве результата рассматривали среднее и лучшее значения ЦФ последнего полученного поколения, усредненного по 10 используемым наборам данных. Вероятность использования оператора кроссовера была экспериментально подобрана равной 0,9, вероятности использования операторов мутации, инверсии и транслокации были выбраны равными 0,3.

Для алгоритма $(1+1)$ -ES текущее число итераций было установлено равным 480 (для того, чтобы число вычислений значений ЦФ совпало с числом их вычислений в ГА: $12 \cdot 20 \cdot 2 = 480$), вероятности использования операторов мутации, инверсии и транслокации были установлены равными 1. Результаты экспериментов представлены в табл. 5. В первом столбце таблицы указывается выбранная схема эволюционного алгоритма, во втором — используемый оператор кроссовера, далее "+" — признак использования соответствующего оператора в алгоритме, F_{best} — среднее по наборам средних по запускам ЦФ лучшего элемента, F_{stat} — среднее по наборам средних по запускам ЦФ среднего по популяции, T — время расчета заданным алгоритмом в секундах.

Анализируя полученные результаты, видим, что среди операторов случайного изменения решения (операторов мутации, инверсии и транслокации) наилучший результат при различных проводимых экспериментах для ГА с отдельным применением указанных операторов был получен при использовании оператора транслокации. В целом для ГА наилучшие результаты позволяло получить

Таблица 5

Результаты тестирования модификаций генетического алгоритма и эволюционной стратегии на тестовом примере из семи городов

ЭА	ОК	ОМ	ОИ	ОТ	F_{best}	F_{stat}	T, c
ГА	ЖОК				23,89	26,35	1,82
ГА	ЖОК	+			23,69	28,69	1,95
ГА	ЖОК		+		23,66	28,28	2,09
ГА	ЖОК			+	23,68	27,28	1,96
ГА	ЖОК	+	+	+	23,71	30,03	1,98
ГА	ЧСОК				24,28	26,63	1,25
ГА	ЧСОК	+			23,94	28,70	1,31
ГА	ЧСОК		+		23,89	27,00	1,25
ГА	ЧСОК			+	23,84	27,56	1,26
ГА	ЧСОК	+	+	+	23,76	29,32	1,35
ГА	ЦОК				25,49	27,72	1,81
ГА	ЦОК	+			24,54	29,78	2,00
ГА	ЦОК		+		24,31	29,42	1,98
ГА	ЦОК			+	24,28	28,30	1,92
ГА	ЦОК	+	+	+	24,01	32,52	2,01
(1 + 1)-ES	—	+			24,65	—	0,09
(1 + 1)-ES	—		+		23,77	—	0,18
(1 + 1)-ES	—			+	26,93	—	0,20

Таблица 6

Результаты тестирования модификаций генетического алгоритма и эволюционной стратегии на тестовом примере из 48 городов

ЭА	ОК	ОМ	ОИ	ОТ	D_{best}	D_{stat}	T, c
ГА	ЖОК		+		3,43	4,02	19,65
ГА	ЧСОК	+	+	+	5,67	6,03	18,98
ГА	ЦОК		+		8,68	9,68	19,21
(1 + 1)-ES	—	+			5,17	—	1,16
(1 + 1)-ES	—		+		3,40	—	1,21
(1 + 1)-ES	—			+	11,03	—	1,48

совместное использование всех трех операторов случайного изменения решений.

Среди операторов кроссовера лучшие решения были получены для "жадного" оператора (ЖОК) и частично соответствующего оператора (ЧСОК) с незначительными отличиями в эффективности в сочетаниях с использованием различных операторов случайного изменения решений.

Второй тестовый пример был взят из библиотеки тестов OR-library, доступной по адресу [13]. В качестве примера был использован тестовый набор данных att48.tsp с координатами 48 городов. Результаты тестирования приводятся в виде отношений найденных лучших решений к известному оптимальному решению. Для ГА брали популяцию из 100 особей, вычисления проводились для 40 поколений; для (1 + 1)-ES, соответственно, число итераций составляло 4000, таким образом, сравнение эффективности проводили на основании одинакового числа вычислений ЦФ обоими алгоритмами при заданных входных параметрах.

Для проведения эксперимента были выбраны те модификации ГА, которые были лучшими на задаче с меньшей размерностью. Результаты экс-

периментов представлены в табл. 6. Дополнительные обозначения в таблице: D_{best} — отношение среднего по запускам значения ЦФ лучшего элемента к оптимальному значению ЦФ, D_{stat} — отношение среднего по запускам значения ЦФ среднего по популяции к оптимальному значению ЦФ.

Полученные экспериментальные данные позволяют сделать вывод о том, что эффективность эволюционной стратегии (1 + 1)-ES сравнима с эффективностью хороших модификаций ГА, но в то же время алгоритм (1 + 1)-ES обладает значительным преимуществом по времени выполнения.

Кроме того, по результатам тестирования на задаче большего размера видим, что увеличение размерности решаемой задачи позволило более четко выделить наиболее эффективные модификации ГА, в частности, показало преимущество использования "жадного" оператора кроссовера по сравнению с частично соответствующим.

Заключение

В данной работе были рассмотрены основные схемы использования эволюционных алгоритмов для решения NP-трудной задачи дискретной оптимизации на примере задачи коммивояжера. Проведенные эксперименты позволили провести сравнение эффективности двух представителей эволюционных алгоритмов — генетического алгоритма и эволюционной стратегии (1 + 1)-ES — и показали, что эффективность лучшей модификации ГА сравнима с эффективностью (1 + 1)-ES, но незначительно ей уступает.

Использование тестовых задач разной размерности показало, насколько получаемая эффективность некоторой модификации ГА зависит от размерности решаемой задачи и позволила выявить наиболее эффективный оператор кроссовера из рассматриваемых в данной работе, им оказался "жадный" оператор кроссовера.

С точки зрения анализа успешности работы алгоритма (1 + 1)-ES предпочтение как наиболее эффективному оператору мутации должно быть отдано оператору инверсии.

Эксперименты П. Борисовского [8] с эволюционным алгоритмом (1 + 1)-ES для транспортной задачи и эволюционным алгоритмом моделирования отжига подтверждают эффективность (1 + 1)-ES, кроме того, многочисленные эксперименты в школе Э. А. Мухачевой для задач раскрытия и упаковки также указывают на превосходство данного алгоритма [7].

На основании проведенных экспериментов мы с полным основанием можем подтвердить принадлежность задачи коммивояжера к классу задач, эффективно решаемых алгоритмом (1 + 1)-ES.

Список литературы

1. **Holland J. H.** Adaptation in Natural and Artificial Systems: An Introductory Analysis with Application to Biology, Control, and Artificial Intelligence. USA: University of Michigan, 1975.
2. **Goldberg D. E.** Genetic Algorithms in Search, Optimization, and Machine Learning. USA: Addison-Wesley Publishing Company, Inc., 1989.
3. **Herrera F.** Tackling real-coded genetic algorithms: operators and tools for the behaviour analysis / F. Herrera, M. Lozano, J. L. Verdegay // Artificial Intelligence Review. 1998. Vol. 12. N 4. P. 265–319.
4. **Орехов Ю. В.** Исследование эффективности операторов генетического алгоритма при поиске экстремальных решений в бинарных и вещественных пространствах / Ю. В. Орехов, Э. Ю. Орехов, А. Р. Заминова // Информационные технологии. 2008. № 9. С. 65–72.
5. **Емельянов В. В.** Теория и практика эволюционного моделирования / В. В. Емельянов, В. М. Курейчик, В. В. Курейчик. М.: Физматлит, 2003. 431 с.
6. **Rechenberg I.** Evolutionsstrategie: Optimierung Technischer Systeme nach Prinzipien der Biologischen Evolution. Stuttgart: Formann-Holzboog Verlag. 1973.

7. **Филиппова А. С.** Моделирование эволюционных алгоритмов решения задач прямоугольной упаковки на базе технологии блочных структур // М.: Информационные технологии. 2006. № 6. 32 с.
8. **Борисовский П. А.** О сравнении некоторых эволюционных алгоритмов / П. А. Борисовский, А. В. Еремеев // Автоматика и телемеханика. 2004. № 3. С. 3–9.
9. **Батищев Д. И.** Применение генетических алгоритмов к решению задач дискретной оптимизации / Д. И. Батищев, Е. А. Неймарк, Н. В. Старостин // Учебно-метод. матер. по программе повышения квалификации "Информационные технологии и компьютерное моделирование в прикладной математике". Нижний Новгород: ННГУ. 2007. 88 с.
10. **Норенков И. П.** Эвристики и их комбинации в генетических методах дискретной оптимизации // Информационные технологии. 1999. № 1. С. 2–7.
11. **Ху Т. Ч.** Комбинаторные алгоритмы: Пер. с англ. / Т. Ч. Ху, М. Т. Шинг. Нижний Новгород: Изд-во Нижегородского гос. университета им. Н. И. Лобачевского. 2004. 330 с.
12. **Herrera F.** Hybrid Crossover Operators for Real-Coded Genetic Algorithms: An Experimental Study / F. Herrera, M. Lozano, A. M. Sanchez // Soft Comput. 2005. 9 (4). P. 280–298.
13. <http://www.iwr.uni-heidelberg.de/groups/comopt/software/TSPLIB95/tsp/>

УДК 621.396.6.049.75

Е. В. Воловикова, аспирантка,
С. У. Увайсов, д-р техн. наук, проф.,
Московский государственный институт
электроники и математики,
e-mail: Uvaysov@yandex.ru

Метод диагностирования радиоэлектронных устройств на основе комплексного электротеплового моделирования

Предложен модифицированный вариант метода справочников для диагностирования аналоговых устройств, основанный на комплексном электротепловом моделировании. Усовершенствованный метод позволяет повысить достоверность диагностирования за счет учета влияния температур на параметры элементов.

Ключевые слова: аналоговые радиоэлектронные устройства, диагностирование радиоэлектронных устройств, метод справочников, электротепловое моделирование.

Для обеспечения надежности радиоэлектронных устройств (РЭУ) на стадии эксплуатации большое значение имеет эффективность решения задач диагностирования. В настоящее время разработаны и применяются такие методы и технологии проверки работоспособности РЭУ, как:

- технология проектирования контролепригодных схем DFT (*Design for Test*);
- периферийное сканирование JTAG (*Design for Test*);
- технология автоматической генерации образцов АТПГ (*Automatic test pattern generation*);
- встроенное самотестирование BIST (*Built-in self-test*).

Эти методы ориентированы на применение для диагностики цифровых электронных средств, печатных узлов и интегральных микросхем [1].

Для аналоговых схем широкое распространение на практике получил метод справочников, который имеет следующие достоинства:

- вычисления сводятся к выполнению простых математических операций, что облегчает его реализацию на ЭВМ;
- большая часть работ проводится на стадии проектирования, а не эксплуатации, как в других методах;
- возможность использования для составления справочника программ математического моделирования РЭУ, что позволяет значительно сократить временные расходы на составление таблицы неисправностей;
- возможность диагностирования катастрофических и постепенных неисправностей с большим уходом параметров;
- интуитивная понятность, что облегчает практическое применение метода.

Стандартный алгоритм метода справочников состоит в выполнении следующих этапов.

Первый этап заключается в формировании списка неисправностей, который охватывает наиболее вероятные из них. Список необходимо составить как можно более полный, так как впоследствии можно будет идентифицировать только эти неисправности. От числа неисправностей в списке также будет зависеть размер таблицы, что служит ограничением для этого метода.

На втором этапе проводится моделирование электрической схемы проверяемого устройства для каждого типа неисправности из списка. В результате формируются векторы входных воздействий и значений выходных характеристик в контрольных точках, на основе которых будут выявляться и локализовываться неисправности. Полученные наборы воздействий и характеристик сохраняются в базе данных. Чтобы обеспечить накопление минимального объема данных, позволяющего достичь желаемой степени диагностируемости, необходимо провести выбор входных воздействий и выходных характеристик.

В процессе проверки диагностируемое устройство подвергается воздействию того же входного сигнала, что был использован при составлении справочника. Полученные характеристики сравниваются с характеристиками, хранящимися в базе данных. Для определения имеющейся неисправности или для ее идентификации в множестве неоднозначно определенных возможных неисправностей вводится критерий локализации.

Формирование справочника начинается с выбора входных сигналов и контрольных точек. В зависимости от входного сигнала выбирается определенная область моделирования электрической схемы:

1. *Статическая область*, в которой рассчитываются напряжения постоянного тока в узлах схемы при произвольных входных воздействиях постоянного тока. Это наиболее простой с точки зрения моделирования способ, требующий минимальных временных и вычислительных ресурсов, а также объема памяти для хранения базы данных неисправностей.

2. *Частотная область*, в которой используются синусоидальные входные тестовые сигналы. Достоинством этого способа является то, что для проверки РЭУ необходима распространенная аппаратура: генератор синусоидальных колебаний, вольтметры и спектроанализаторы. Существует несколько подходов к диагностике схем в частотной области:

- подход Сешу и Уоксмэна используется для линейных частотно-зависимых цепей. Справочник содержит информацию об амплитудных характеристиках на входах и выходах;
- билинейное преобразование используется для диагностики одиночных мягких или жестких неисправностей в линейных схемах. Справочник содержит годографы преобразования передаточной функции схемы в билинейную функцию, зависящую от параметров элементов схемы;
- методы распознавания с использованием разреженных матриц, которые получаются путем квантования сигнала по нескольким уровням и частоте. Уровень выбирается с учетом разброса параметров элементов и погрешностей элементов.

3. *Временная область*, в которой используются методы анализа с помощью произвольных тестовых сигналов. Основной сложностью при диагностировании в этой области является сравнение сигналов, хранящихся в справочнике, и измеренного сигнала. Например, для диагностики линейных динамических схем применяются методы динамических испытаний. Наиболее распространенные из них [2]:

- метод псевдослучайных сигналов. Для проверки цепи используется периодический псевдослучайный сигнал $\eta(t)$ [3]. При этом импульсная функция системы аппроксимируется взаимной корреляционной функцией входного и выходного сигналов. Для формирования справочника используются значения разности взаимно корреляционных функций при номинальном и неисправном режимах работы или результаты ее квантования (в случае, если взаимно корреляционная функция не может быть аппроксимирована импульсной функцией цепи);

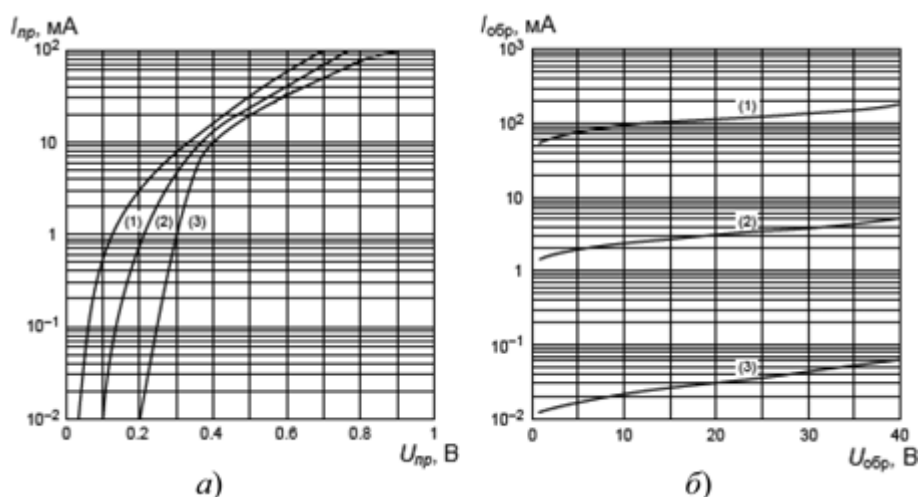


Рис. 1. Температурная зависимость ВАХ диода при прямом смещении (а), при обратном смещении (б):
1 — $T = 125^\circ\text{C}$, 2 — $T = 85^\circ\text{C}$, 3 — $T = 25^\circ\text{C}$

- метод конструирования тестового сигнала, предложенный Шрайбером, основан на использовании возможностей генерирования произвольных сигналов в современных испытательных комплексах [2]. При формировании справочника неисправностей берутся изменения уровней синтезированного сигнала и отклонения положения нулевого уровня отклика на временной оси для всех рассматриваемых случаев неисправностей.

Как говорилось выше, одним из достоинств метода справочников является возможность применения математического моделирования. Но при использовании моделирования для составления справочника необходимо обеспечить высокую точность расчета электрических характеристик. На примере диода видно, что при повышении температуры изменяются вольт-амперные характеристики (ВАХ) (рис. 1).

В аналоговых схемах на параметры элементов значительное влияние оказывает температура. Из приведенного на рис. 1 примера видно, что пренебрежение температурным фактором при использовании математического моделирования для составления справочника приводит к тому, что значения выходных характеристик, полученных при моделировании, будут отличаться от реальных. Существующие реализации метода справочников при расчете выходных характеристик не учитывают отличие номинальной температуры от рабочей. Это может привести к постановке неверного диагноза.

Таким образом, предлагается усовершенствовать метод справочников путем учета индивидуальных температур электрорадиоэлементов, что позволит повысить достоверность диагностирования. Блок-схема такого метода приведена на рис. 2.

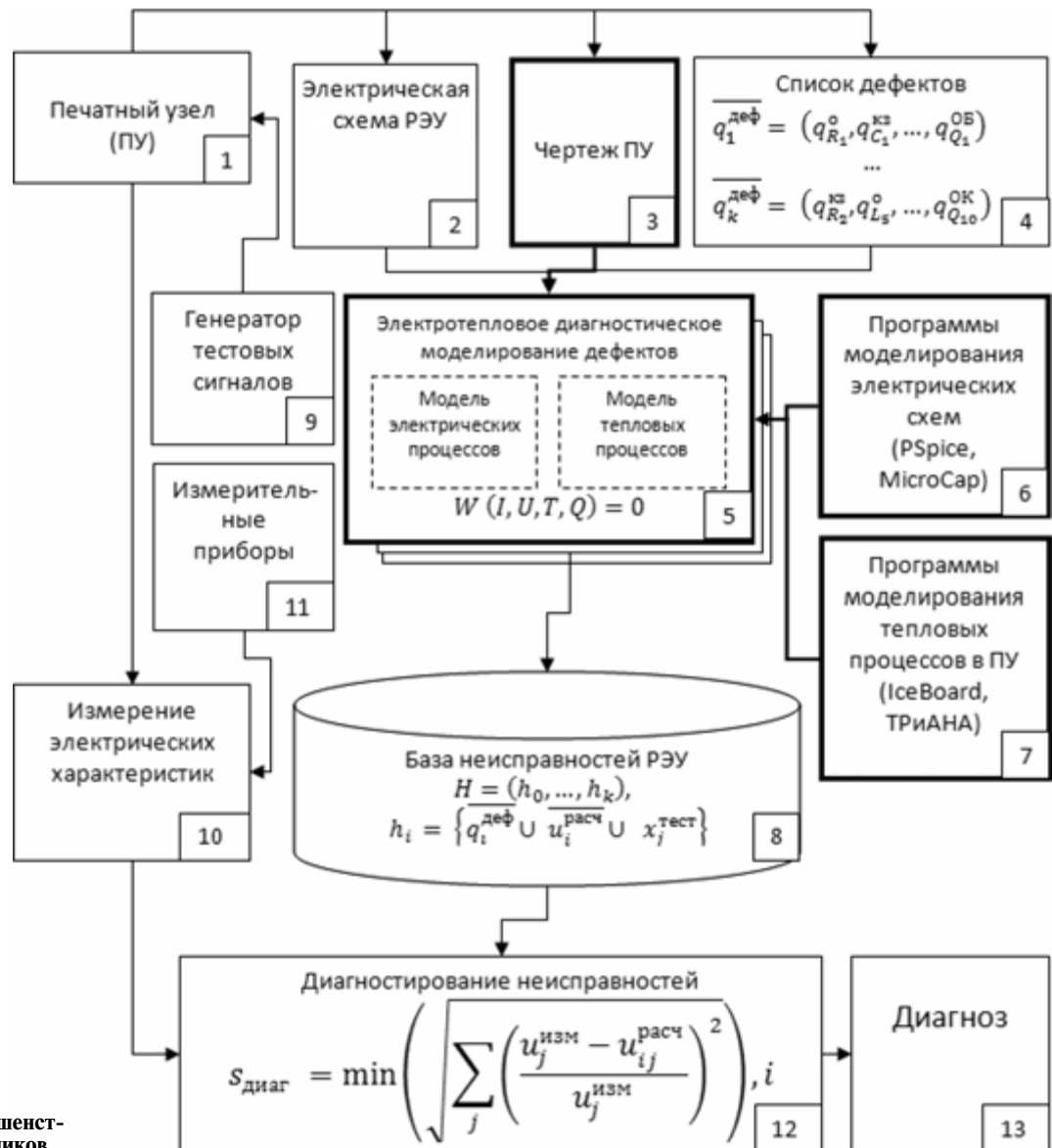


Рис. 2. Блок-схема усовершенствованного метода справочников

Выделенные блоки и связи отличают предложенный метод от известных.

Исходными данными для создания справочника для диагностирования печатного узла (блок 1) данным методом служат электрическая схема (блок 2), чертеж конструкции (блок 3) и список возможных неисправностей исследуемого объекта (блок 4). При этом рассматриваются следующие виды дефектов: катастрофические одиночные и множественные неисправности, типа "обрыв" и "короткое замыкание".

На следующем этапе на основе исходных данных синтезируется множество комплексных электротепловых диагностических моделей (блок 5). Для каждой модели существует система уравнений, связывающая между собой множество токов I , напряжений U , температур T , а также множество Q электрических, геометрических и теплофизических параметров печатного узла (ПУ).

Для расчета электрических подмоделей (блок 6) могут быть использованы системы PSpice, WinSpice, MicroCap и т. п. При моделировании тепловых процессов в ПУ (блок 7) могут быть использованы следующие программные комплексы: IceBoard, ТРИАНА, АСОНИКА-ТМ, ВЕТАsoft Board, PCAnalyze и др. В соответствии с описанной комплексной электротепловой диагностической моделью между программными комплексами организуется передача данных из результатов расчетов.

Результаты расчета комплексной диагностической модели для каждой неисправности при различных тестовых воздействиях сохраняются в базе неисправностей (блок 8). При диагностировании неисправности с помощью генератора тестовых сигналов (блок 9) создаются входные воздейст-

вия. С помощью измерительного оборудования (блок 11) измеряются напряжения в контрольных точках (блок 10).

Для каждой неисправности из списка проводится расчет критерия локализации (блок 12). В качестве диагноза (блок 13) выбирается та неисправность, у которой критерий локализации минимален. В случае, если вычисленные значения критерия локализации нескольких неисправностей близки друг к другу и к минимальному значению, то возникает множество неоднозначно определяемых неисправностей.

В заключение отметим, что основное достоинство предложенного метода — повышение достоверности диагностирования РЭУ по сравнению со стандартным методом справочников, что особенно заметно проявляется при диагностировании теплонагруженных РЭУ, а также РЭУ, содержащих компоненты, параметры которых сильно зависят от температуры. При этом объем вычислений при составлении справочника (на стадии проектирования РЭУ) несколько увеличивается из-за необходимости проведения комплексного электротеплового моделирования, но при диагностировании РЭУ на стадии эксплуатации объем работ не отличается от стандартного метода.

Список литературы

1. **Kacand U.** Extending IEEE Std. 1149.4 Analog Boundary Modules to Enhance Mixed-Signal Test / U. Kacand, F. Novak, F. Azas // Test and Measurements World. 2006. N 12.
2. **Schreiber H.** Fault dictionary based upon stimulus design Circuits and Systems // IEEE Transactions. 1979. Vol. 26. Issue 7. P. 529—537.
3. **Merrill H. M.** Failure diagnosis using quadratic programming // IEEE Truns. Reliuhiliv. 1973. Vol. R-22. P. 207—213.

Новая книга

Корсаков С. Н. Начертание нового способа исследования при помощи машин, сравнивающих идеи / Пер. с франц. под ред. А. С. Михайлова. — М.: МИФИ, 2009. — 44 с.

В 2009 году в издательстве Московского инженерно-физического института вышел перевод трактата С. Н. Корсакова "Начертание нового способа исследования при помощи машин, сравнивающих идеи" от 1832 года, написанный по традиции того времени на французском языке. Семен Николаевич Корсаков (1787—1853) — русский дворянин, пионер российской кибернетики и искусственного интеллекта. В первой половине XIX века Корсаков выдвинул концепцию усиления возможностей разума посредством разработки научных методов и устройств. В 1832 году он опубликовал описание своего метода и пяти механических "интеллектуальных машин" для информационного поиска и классификации. В работах Корсакова также содержится целая плеяда новых для того времени идей, как то: применение перфорированных карт в информатике, многокритериальный поиск, обработка больших массивов данных, предтеча современных экспертных систем. В трактате содержится подробное описание устройства и принципов функционирования всех изобретенных машин. В сопроводительной статье дается ретроспектива повторного открытия незаслуженно забытых изобретений Корсакова и попытка интерпретации функционирования предложенных устройств в современных терминах операций с множествами. Издание может свободно распространяться и использоваться без всяких ограничений. Предназначено для всех интересующихся историей кибернетики и информатики.

Электронная версия издания трактата С. Н. Корсакова доступна на сайте кафедры кибернетики МИФИ: <http://cyber.mephi.ru/emportal/h.csp?emI = МЕРХ8KRJ3&emK = K8X&emL = AS0H17>

В. Г. Секаев, канд. техн. наук, доц.,
Новосибирский государственный
технический университет,
e-mail: sekaev@mail.ru

Использование алгоритмов комбинирования эвристик при построении оптимальных расписаний

Рассмотрена задача составления оптимальных расписаний с использованием генетических алгоритмов комбинирования эвристик. Предложенный подход позволяет составлять квазиоптимальные расписания для большого класса многостадийных обслуживающих систем.

Ключевые слова: генетический алгоритм, метод комбинирования эвристик, расписание, время технологического цикла, набор эвристик, распределение работ.

Качество функционирования современного производства во многом определяется решениями, принимаемыми на этапах календарного планирования (КП) и оперативного управления. Наряду с улучшением качеств плановых решений все более жесткими становятся требования к сокращению сроков их выработки, повышению оперативности и гибкости управления.

Постановка задачи [1]

Имеется конечное множество $N = \{1, 2, \dots, n\}$ требований (деталей, партий) и конечное множество $M = \{1, 2, \dots, M\}$ приборов (станков). Процесс обслуживания требования $i \in N$ включает r_i стадий. При этом каждому требованию $i \in N$ и каждой стадии q , $1 \leq q \leq r_i$, его обслуживания соответствует некоторое множество приборов $M_q^i \subseteq M$. В зависимости от типа обслуживающей системы требование i на стадии q либо может обслуживаться любым из приборов $L \in M_q^i$ (но не более чем одним одновременно), либо должно одновременно обслуживаться всеми приборами из множества M_q^i . Обычно предполагается, что каждый прибор одновременно может обслуживать не более одного требования.

В этих системах с последовательными приборами для каждого требования $i \in N$ задается своя, специфическая для этого требования последовательность $L^i = (L_1^i, L_2^i, \dots, L_{r_i}^i)$ его обслуживания приборами. Требование i сначала обслуживается прибором L_1^i , затем прибором L_2^i и т. д., пока оно не будет обслужено прибором $L_{r_i}^i$. Последо-

вательности обслуживания могут быть различными для разных требований и содержать повторения приборов.

Если требование i на стадии q должно быть обслужено прибором $L = L_q^i$, то предполагается заданной длительность $t_{iL} \geq 0$ его обслуживания этим прибором. Процесс функционирования обслуживающей системы может быть описан заданием расписания (календарного плана, временного графика), т. е. некоторой совокупности указаний относительно того, какие именно требования, какими именно приборами обслуживаются в каждый момент времени. Обычно предполагается, что каждое требование не может одновременно обслуживаться двумя и более приборами и каждый прибор не может одновременно обслуживать более одного требования. При этих предположениях расписание можно рассматривать как совокупность $\{s_1(t), s_2(t), \dots, s_M(t)\}$ кусочно-постоянных непрерывных слева функций, каждая из которых задана на интервале $0 \leq t \leq \infty$ и принимает значения $0, 1, \dots, n$. Если $s_L(t') = i \neq 0$, то в момент времени t' прибор $L \in M$ обслуживает требование $i \in N$.

Оценка качества расписания осуществляется следующим способом.

Каждое допустимое расписание s однозначно определяет вектор $\mathbf{t}(s) = (t_1(s), t_2(s), \dots, t_n(s))$ моментов завершения обслуживания требований. Задается некоторая действительная неубывающая по каждой из переменных функция $F(x) = F(x_1, x_2, \dots, x_n)$, и качество расписания s оценивается значением этой функции при $x = \mathbf{t}(s)$. Из двух расписаний лучшим считается то, которому соответствует меньшее значение $F(x)$. При построении оптимального по быстродействию расписания $F(x) = \max_{1 \leq i \leq n} \{x_i\}$. В этом случае $F(\mathbf{t}(s)) = \mathbf{t}_{\max}(s)$, где $\mathbf{t}_{\max}(s) = \max_{1 \leq i \leq n} \{t_i(s)\}$.

Генетические алгоритмы относятся к наиболее универсальным подходам к решению задач обслуживаемого класса [2, 3].

Генетический метод комбинирования эвристик

Главная проблема постановки задачи синтеза расписаний, как отмечается в работах [4–6], заключается в том, как отобразить в хромосоме структурные параметры, т. е. в случае синтеза расписаний — данные о распределении заданного множества работ во времени и между обслуживающими аппаратами. Так, в известном методе РМХ, в котором гены соответствуют номерам работ, возникает проблема коррекции дочерних

хромосом, образующихся после операции кроссовера, поскольку каждая работа должна фигурировать в расписании, причем только один раз. Эта коррекция нарушает складывающиеся полезные сочетания (хромосомные блоки) аллелей (значений генов), что снижает эффективность поиска.

Преодолеть этот недостаток можно, если использовать идею отображения в хромосоме информации о расписании в неявной форме, при которой гены представляют не номера работ, а правила генерации очередного варианта расписания. Эта идея и была реализована в методе, названном методом комбинирования эвристик (НСМ — *Heuristics Combination Method*) [4—6].

Метод комбинирования эвристик может быть применен к широкому кругу задач планирования.

Метод НСМ [6] свободен от необходимости корректировки дочерних хромосом, так как аллелями служат не значения проектных параметров, а имена эвристик, используемых для определения этих значений.

В методе НСМ i -й ген соответствует i -й подзадаче, а i -й аллель есть номер (имя) процедуры (эвристики), реализующей локальную целевую функцию и ограничения i -й подзадачи.

Метод комбинирования эвристик в задачах синтеза расписаний

Для конкретной задачи планирования необходимо было определить такое расписание, в котором общая длительность выполнения всех операций была бы минимальной. Для решения поставленной задачи использовался генетический метод комбинирования эвристик, эффективность которого подтверждалась публикациями [2—7]. Специфика рассматриваемой задачи отражена в эвристиках построения расписания. Используя всевозможные эвристики (правило выбора очередной операции и правило ее назначения на определенную машину) для данного типа задач из опубликованных источников [2—7], был составлен следующий их перечень:

- выбирается операция с наименьшим временем окончания обслуживания на предыдущей стадии;
- выбирается машина, на которой обслуживание данной операции закончится раньше, чем на других машинах;
- выбирается операция с наибольшим временем окончания обслуживания на предыдущей стадии;
- выбирается операция с наименьшим временем окончания обслуживания на текущей стадии;
- выбирается операция с наибольшим временем окончания обслуживания на текущей стадии;
- выбирается операция с наименьшим временем длительности всех работ;

- выбирается операция с наибольшим временем длительности всех работ;
- выбирается операция с наименьшим временем длительности оставшихся работ;
- выбирается операция с наибольшим временем длительности оставшихся работ.

При этом учитывались следующие технологические условия:

- технологические операции не имеют приоритетов;
- время выполнения операций является детерминированной величиной;
- на каждом модуле в любой момент времени может выполняться операция, относящаяся только к одному виду работ;
- маршрут выполнения работ каждой детали известен;
- каждый модуль одновременно не может выполнять более одной операции;
- каждая операция выполняется на модуле без перерыва.

Экспериментальные результаты

Программная реализация метода проверялась на **модельных** задачах из работ [1] и [8], где предлагались квазиоптимальные алгоритмы решения задач календарного планирования. Для указанных в работах [1, 8] задач были приведены расписания, которые были сразу же получены нами при программной реализации НСМ. Дополнение вышеперечисленных эвристик еще одной — **"отложить выбор детали на следующий шаг"** — позволило получить лучшее решение, чем в представленных источниках. Решения были представлены диаграммами Ганта.

Пример 1 [1].

На стендовом полигоне необходимо провести испытания двух новых изделий 1 и 2. Для каждого изделия известен порядок прохождения приборов. На полигоне имеется по одному испытательному стенду А, В, С и D. Порядки и длительности испытаний на стендах первого и второго изделий заданы соответственно в табл. 1 и 2.

При расписании без прерываний моменты завершения обслуживания требований принимают значения (решение из источника [1]), представленные на рис. 1:

$$t_1(s_1) = 23; \quad t_2(s_1) = 29.$$

Полученное в данной работе решение представлено на рис. 2:

$$t_1(s_1) = 26; \quad t_2(s_1) = 22.$$

Пример 2 [8]. Расчет календарного плана для четырех станков (Р, Т, Ф, Ш) и четырех партий деталей (А, Б, В, Г).

Таблица 1

Испытания изделия 1

Порядок испытаний изделия 1	A	B	C	D	A
Длительность испытаний	2	2	8	3	6

Таблица 2

Испытания изделия 2

Порядок испытаний изделия 2	A	B	A	C	A	C
Длительность испытаний	1	3	2	4	4	2

Таблица 3

Исходные данные

	P		T		Ф		Ш	
Б	2	1	7	3	5	2		
В			6	1	4	2	9	3
Г	5	2	4	4	8	1	7	3
А	8	1	4	2			2	3

Таблица 4

Трудоёмкость обработки плиты № 1 по каждой операции

№ пп	Номер операции	Операция	Трудоёмкость, час.
1	1	Фрезерная	40,45
2	1	Фрезерная	11,45
3	2	Расточная	6,15
4	3	Разметка	1,8
5	2	Расточная	42,45
6	1	Фрезерная	16,6
7	4	Термозакалка	10
8	1	Фрезерная	38,6
9	5	Шлифовальная	5,3
10	2	Расточная	19
11	2	Расточная	48,6
12	3	Разметка	2
13	6	Сверлильная	7,45
ИТОГО			249,85

Исходные данные приведены в табл. 3, где первая цифра — время обработки детали на указанном станке, вторая цифра — технологический маршрут обработки детали.

Решение из источника [8] представлено на рис. 3:

$$t_{\max}(s) = 32.$$

Полученное в данной работе решение представлено на рис. 4:

$$t_{\max}(s) = 31.$$

Программная реализация метода была использована для получения КП производства станочных плит одного из машиностроительных предприятий.

Известен технологический маршрут каждой станочной плиты, которые обрабатываются на станочных модулях за конечное число операций, выполняемых в заданной последовательности.

Задана трудоёмкость обработки каждой плиты на каждой операции.

Исходные данные по каждой плите приведены в табл. 4—6.

Общее непрерывное время производства первой плиты составляет 249,85 ч. Общее непрерыв-

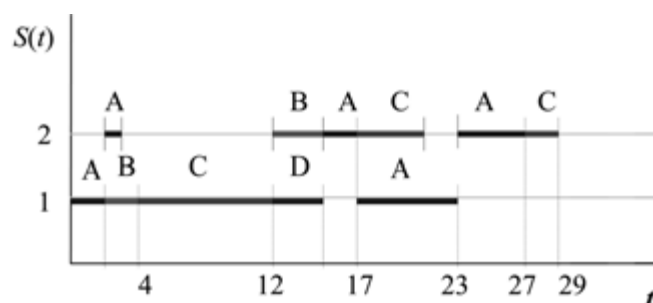


Рис. 1. Расписание для примера 1 из источника [1]

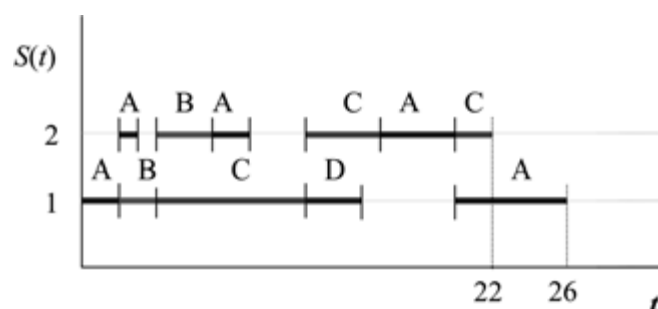


Рис. 2. Расписание для примера 1, полученное в данной работе

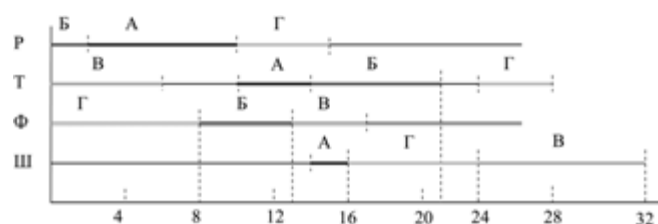


Рис. 3. Расписание для примера 2 из источника [8]

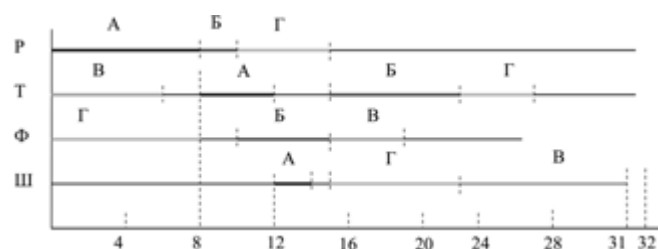


Рис. 4. Расписание для примера 2, полученное в данной работе

Таблица 5

Агрегированная трудоемкость обработки плиты № 2 по каждой операции

№ пп	Номер операции	Операция	Трудоемкость, ч
1	1	Фрезерная	7,9
2	3	Разметка	1
3	1	Фрезерная	7,3
4	2	Расточная	8,85
5	3	Разметка	0,4
6	1	Фрезерная	3,1
7	3	Разметка	0,5
8	1	Фрезерная	3
9	3	Разметка	1
10	2	Расточная	40,5
11	3	Разметка	0,55
12	2	Расточная	12,3
13	3	Разметка	1,5
14	1	Фрезерная	8,7
15	1	Фрезерная	44,5
16	5	Шлифовальная	3,6
17	2	Расточная	6,5
18	2	Расточная	53,3
19	2	Расточная	15,3
20	2	Расточная	28,35
21	3	Разметка	0,6
22	6	Сверлильная	5,5
ИТОГО			254,25

Таблица 6

Агрегированная трудоемкость обработки плиты № 3 по каждой операции

№ пп	Номер операции	Операция	Трудоемкость, час.
1	3	Разметка	1,7
2	1	Фрезерная	28
3	2	Расточная	34,7
4	4	Термозакалка	10
5	1	Фрезерная	18,4
6	1	Фрезерная	9,4
7	3	Разметка	2,4
8	2	Расточная	100,5
9	3	Разметка	0,5
10	2	Расточная	45
11	3	Разметка	1,5
12	1	Фрезерная	12,5
13	4	Термозакалка	10
14	1	Фрезерная	32,5
15	5	Шлифовальная	4,2
16	2	Расточная	26,65
17	2	Расточная	90
18	2	Расточная	30,15
19	2	Расточная	18,2
20	3	Разметка	1
21	6	Сверлильная	4,05
ИТОГО			481,35

ное время выполнения второй и третьей плиты составляет 254,25 и 481,35 ч соответственно. В результате решения задачи календарного планирования совокупная длительность производственного цикла составила 680,19 ч.

При расчете календарного плана производства станочной плиты трех типов с 22 видами работ потребовалось представление информации не только в виде графика Ганта, но и в матричном виде, так как длительности различных операций отличались на порядок.

СПИСОК ЛИТЕРАТУРЫ

1. Танаев В. С., Сотсков Ю. И., Струсевич В. А. Теория расписаний. Многостадийные системы. — М.: Наука, Гл. ред. физ.-мат. лит., 1989. 328 с.
2. Норенков И. П. Генетические алгоритмы синтеза расписаний и планов. // Техника, экономика. Сер. Автоматизация,

проектирование. М.: ВНИИ межотраслевой информации. 1995. Вып. 1—2. С. 11—15.

3. Батищев Д. И., Власов С. Е. Решение задачи "слепого" упорядочивания с помощью генетических алгоритмов // Труды конференции по генетическим алгоритмам. М.: 1996. — С. 32—35.

4. Батищев Д. И., Гудман Э. Д., Норенков И. П., Прилуцкий М. Х. Метод декомпозиций для решения комбинаторных задач упорядочения и распределения ресурсов // Информационные технологии. 1997. № 1. С. 29—33.

5. Норенков И. П. Генетические методы структурного синтеза проектных решений // Информационные технологии. 1998. № 1. С. 9—13.

6. Норенков И. П. Эвристики и их комбинации в генетических методах дискретной оптимизации // Информационные технологии. 1999. № 1. С. 2—7.

7. Норенков И. П. Методы оптимизации в задачах концептуального проектирования и логистики // Информационные технологии. 2000. — № 9. С. 9—14.

8. Петров В. А., Масленников А. Н., Осипов Л. А. Планирование гибких производственных систем. Л.: Машиностроение, 1985.

УДК 004: 656.61.073

В. М. Дорожко, канд. физ.-мат. наук, ст. науч. сотр.,
Институт автоматики и процессов управления
ДВО РАН, г. Владивосток,
e-mail:bendor@iacp.dvo.ru

Информационная технология идентификации начального этапа разгона морского судна

На основе исследования уравнений динамики движения разработана информационная технология идентификации начального этапа экстренного разгона морского судна. Установлено, что его идентификационным информационным признаком является максимум линейного ускорения судна. Численным моделированием показано, что причиной формирования информационного признака является конкуренция сил разнонаправленного действия: инерционных сил, определяющих характер нарастания скорости судна, и сил упора винта, развитие которых предопределено характеристикой двигателя. Выявленный информационный признак может быть использован в системах управления движением судов, а также судоводителями встречных судов и направлен на предупреждение морских катастроф, вызванных поздним выявлением ошибочного маневрирования судов.

Ключевые слова: информационная технология, идентификация маневра судна, максимум ускорения судна, упор винта, предупреждение морских катастроф.

Введение

В современных условиях морские перевозки являются наиболее эффективными с экономической точки зрения, а для большинства стран Азиатско-Тихоокеанского региона — это единственное доступное средство обеспечения грузооборота. В связи со строительством нефтепровода "Восточная Сибирь — Тихий океан" проблемы эффективности морских перевозок затронут также и российские порты Дальнего Востока. Акцент на проблеме безопасности обусловлен тем, что потоки судов создают на подходах и непосредственно на акваториях морских портов сложную динамическую картину пересекающихся насыщенных судопотоков, а также в связи с тем, что наблюдающаяся тенденция к повышению тоннажа судов влечет за собой увеличение частоты мор-

ских происшествий по причине ошибочного маневрирования [1].

В настоящее время суждение о маневренных режимах движения судна формируется в системах управления движением на основе радиолокационных наблюдений за траекторией и вектором скорости точки, отождествляемой с судном. Формирование на основе наблюдений вывода о переходе судна к экстремному разгону осуществляется при достижении им заметных изменений скорости. Это приводит к запаздыванию получения информации о действиях судов и, как следствие, невозможности принятия как системами управления движением судов (УДС), так и судоводителями своевременных превентивных мер по исключению недопустимого сближения и столкновения судов.

Проблеме маневренных режимов движения судов посвящено значительное число работ. Результаты исследований и перечень основной литературы представлены как в справочном издании [2], так и в публикациях, например [3], [4]. Экстремное изменение скорости движения рассматривается как маневр "последней минуты", применяемый для уклонения от столкновения судов в аварийных ситуациях, и в связи с этим имеет исключительное значение для обеспечения безопасности судоходства. Тем не менее, параметры движения судна на начальном этапе разгона не используются как информационные предвестники перехода судна к экстремному изменению скорости движения, так как не известны связанные с изменением угловой скорости вращения винта и доступные для дистанционного наблюдения идентификационные признаки начала разгона судна.

Настоящая работа посвящена разработке информационной технологии идентификации начального этапа экстренного разгона судна в интересах повышения безопасности мореплавания в насыщенных судопотоках для распознавания случаев ошибочного маневрирования, а также принятия своевременных и эффективных мер по предупреждению столкновения судов, что обуславливает актуальность предлагаемых исследований.

Постановка задачи и основные модельные представления

Отметим комплекс недостатков, характерных для применяемых в настоящее время системами

УДС методов распознавания маневренных режимов судна:

- интервал оценки вектора скорости движения судна в системах УДС составляет до 20 с; для формирования вывода о тенденции изменения скорости необходимо несколько таких циклов. Учитывая, что расстояние между судами в условиях высокой плотности морского движения составляет всего несколько длин корпусов, применение указанной технологии связано с риском запаздывания выявления случаев ошибочного маневрирования;
- системы УДС не позволяют идентифицировать в реальном масштабе времени динамику угловой скорости вращения гребного винта наблюдаемого судна и развиваемой им продольной силы.

Предлагаемая информационная технология свободна от указанных недостатков в связи с применением установленного в рамках настоящей работы информационного признака обнаружения начального этапа экстренного разгона судна и выявлением его связи (признака) с динамикой угловой скорости вращения гребного винта наблюдаемого судна и развиваемой им продольной силы.

Суть предлагаемой информационной технологии состоит в том, что на основе радиолокационного наблюдения за положением судна выполняется оценка и мониторинг за линейным ускорением движения судна. Заключение о начале разгона формируется на основе обнаружения максимума ускорения судна. Ранняя диагностика разгона судна позволяет в системе УДС сформулировать рекомендации судоводителю по предупреждению столкновения. На рис. 1 представлена структурная схема модели рассматриваемой информационной технологии, содержащая основные информационные объекты и отношения информационной зависимости между ними. В связи с тем, что функционирование элементов действующих систем УДС отработано, в данной статье рассматриваются аспекты технологии, имеющие непосредственное отношение к формированию максимума линейного ускорения движения судна.

Обоснование корреляции между моментом времени достижения максимума ускорения и выхода винта на установившиеся обороты позволит принять момент времени достижения максимума ускорения в качестве идентификационного признака начала экстренного маневрирования. В связи с этим перейдем к формальному описанию задачи обоснования оригинальной части предложенной информационной технологии.

Экстренное изменение скорости судна представляет собой типичный неустановившийся режим движения, обусловленный инерционностью вращающихся и поступательно движущихся масс судна в сопротивляющейся его движению среде.



Рис. 1. Основные информационные объекты

Введем неподвижную правую систему координат $ox_1x_2x_3$, связанную с морской акваторией, таким образом, что ось ox_1 совпадает с направлением движения судна, а ось ox_3 направлена вверх.

На начальном этапе разгона судна примем следующие традиционные допущения [5]:

- судно, руль которого установлен в нейтральное положение, продолжает двигаться неизменным прямым курсом;
- присоединенные массы воды корпуса судна и его винта на начальном этапе разгона остаются неизменными;
- сопротивление воды движению судна квазистационарно, т. е. зависит только от скорости судна;
- продольная сила, развиваемая винтом, и момент сил сопротивления жидкости вращению винта квазистационарны, т. е. зависят только от мгновенных значений частоты вращения винта и скорости движения судна.

Движение судна описывается следующей системой уравнений [6], [7]:

$$(J + j)\dot{\omega} + M_p(\omega, v_p) + M_f(\omega) = [1 - \exp(-t/\tau_e)]M_e(\omega); \quad (1)$$

$$(M_s + m_s)\dot{v} + R(v) = (1 - t_p)F_p(\omega, v_p), \quad (2)$$

где $J = J_p + J_e$ — суммарный момент инерции винта и движущихся частей вала (J_p) и двигателя (J_e), отнесенных к оси винта; $j = 0,032(\rho D^5/z_p) \times (\alpha_p \beta_p)^2$ — присоединенный момент инерции винта; ρ — плотность жидкости; D — диаметр винта; z_p — число лопастей винта; ω — угловая

скорость вращения винта, α_p — отношение площади лопастей винта к площади круга, описываемой лопастями винта; β_p — отношение смещения винта за один оборот в твердой среде к его диаметру; $M_p = K_M \rho \omega^2 D^5$ — момент сил сопротивления жидкости вращению винта (момент винта); K_M — коэффициент момента винта; v — скорость движения судна; $v_p = v(1 - w)$ — скорость жидкости в корме судна, т. е. на входе винта; w — коэффициент, отражающий относительное снижение скорости жидкости в корме судна; τ_e — постоянная времени реакции двигателя на изменение подачи топливной смеси; t_p — коэффициент, отражающий уменьшение продольной силы винта за счет снижения давления за кормой судна, вызванного работой винта; $M_f = (c_e + c_s) \times \left(\frac{1}{3} + \frac{2}{3} \frac{\omega}{\omega_n}\right) M_n$ — момент потерь в двигателе судна и механической системе передачи движения, $c_e = 0,1 \dots 0,2$, $c_s = 0,02 \dots 0,03$; ω_n — угловая скорость вращения винта при номинальной (полной) мощности M_n двигателя; M_e — крутящий момент, развиваемый двигателем судна; M_s — масса судна; $m_s = \left(0,465 \frac{B}{L} - 0,028\right) \frac{2 T_M}{B}$ — присоединенная масса судна; B — ширина, L — длина, T_M — осадка судна в его средней части; $R = C_d \frac{\rho v^2}{2} S$ — сила сопротивления жидкости движению судна; C_d — коэффициент силы сопротивления; S — площадь смоченной поверхности судна; $F_p = K_F \rho \omega^2 D^4$ — движущая судно продольная сила (упор винта), создаваемая вращающимся винтом; K_F — коэффициент упора винта.

Уравнение моментов (1) описывает динамику комплекса, включающего в себя судовую двигатель, механическую систему передачи крутящего момента на винт и, собственно, сам вращающийся в воде винт. Движение судна под действием силы упора винта описывается уравнением (2).

Концепция формального решения

Для решения системы уравнений (1), (2) требуются численные значения коэффициентов K_M , K_F и C_d , которые обычно получают экспериментальными методами. В данной задаче значения указанных коэффициентов получены методом вычислительной гидродинамики (CFD — *Computational Fluid Dynamics*), который с развитием информационных технологий в последние годы на-

чинает составлять достойную альтернативу сложным и затратным экспериментальным методам.

В основу CFD-моделирования положены уравнения неразрывности несжимаемой жидкости и сохранения импульса (RANSE — *Reynolds-averaged Navier-Stokes equations*) [8], которые с учетом принятых допущений могут быть записаны в тензорном виде соответственно как (3) и (4):

$$\rho \frac{\partial u_i}{\partial x_i} = 0; \quad (3)$$

$$\rho \frac{\partial}{\partial x_j} (u_i u_j) = -\frac{\partial p}{\partial x_i} + \mu \frac{\partial}{\partial x_j} \left[\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} - \frac{2}{3} \delta_{ij} \frac{\partial u_l}{\partial x_l} \right] + \rho \frac{\partial}{\partial x_j} (-\overline{u_i' u_j'}), \quad (4)$$

где x_1, x_2, x_3 — декартовы координаты в абсолютной системе $ox_1 x_2 x_3$ и соответствующие им u_1, u_2, u_3 — мелкомасштабные осреднения значений абсолютной скорости потока жидкости, а также u_1', u_2', u_3' — флуктуации абсолютной скорости, ответственные за диссипацию энергии; μ и ρ — вязкость и плотность жидкости, соответственно; p — давление; $\rho \overline{u_i' u_j'}$ — напряжения Рейнольдса, полученные осреднением флуктуационных компонент скорости.

Для вычисления напряжений Рейнольдса применяли две известные модели турбулентности:

- модель турбулентности $k-\varepsilon$, RNG (*renormalization group mathematical technique*) [9], которая позволяет получить расчетные значения C_d , близкие к их экспериментальным значениям [10];
- модель турбулентности $k-\omega$ [11], верификация которой [12] показала в сравнении с другими моделями расчетные значения K_M и K_F , наиболее близкие к их экспериментальным значениям.

В связи с этим уравнения (3), (4) дополнены уравнениями [9], [11] переноса турбулентной кинетической энергии k и скорости диссипации ε или специфической скорости диссипации ω турбулентной энергии для моделей $k-\varepsilon$, RNG и $k-\omega$, соответственно.

Вычисление коэффициентов сопротивления $C_d = f(v)$, момента $K_M = f(\omega, v_p)$ и силы $K_F = f(\omega, v_p)$ выполнено в обращенном потоке жидкости: движение судна и винта имитировалось встречным потоком жидкости с заданной скоростью. Моделирование C_d и K_M, K_F выполнено отдельно.

Полномасштабная аналитическая модель судна устанавливалась неподвижно в канале прямоугольной формы, ребра которого ориентированы параллельно осям системы координат $ox_1 x_2 x_3$, наибольшая сторона канала и курс судна параллельны оси ox_1 ; оси ox_1 и ox_2 лежат на поверхно-

сти жидкости, при этом на входе канала задавались компоненты скорости $u_1 = v$, $u_2 = u_3 = 0$. При вычислении коэффициента C_d измерялась скорость v_p в кормовой части судна, что позволило вычислить коэффициент w и, соответственно, установить точное значение скорости натекающего на винт потока.

Вычисление коэффициентов K_M и K_F выполнено в канале цилиндрической формы, ось которого совпадала с осью вращения винта. Вычисления повторялись по сетке значений, представляющих собой матрицу ($m \times n$) дискретных значений линейной скорости на входе канала $u_1 = v_p$, $u_2 = u_3 = 0$ (m — строки матрицы) и угловой скорости ω вращения винта (n — столбцы матрицы). Нормальные составляющие скорости жидкости на поверхностях судна и винта приняты равными нулю.

Численное интегрирование системы уравнений (1), (2) выполнено методом Рунге—Кутты четвертого порядка точности [13].

В результате интегрирования системы уравнений (1), (2) должны быть вычислены скорость движения судна $v(t)$, его ускорение $\dot{v}(t)$, угловая скорость $\omega(t)$, момент $M_p(t)$ и сила упора $F_p(t)$ винта. Анализ корреляций указанных переменных позволит выработать правило идентификации начала экстренного маневрирования судна.

Численные результаты и их верификация

В качестве модельного судна выбрано небольшое судно (буксир) с высокой маневренностью, что позволило снизить требования к вычислительному ресурсу.

Основные параметры судна: $L = 26,1$ м; $B = 7,9$ м; $T_M = 2,9$ м; $M_s = 355,4$ т; $S = 224$ м².

Параметры винта: тип — тяговый винт фиксированного шага; $z_p = 4$; $D = 1,65$ м; $\alpha_p = 0,83$; $\beta_p = 1,15$; $J_p = 78,14$ кг·м².

Параметры двигателя: дизель 64РН36/45; $J_e = 760$ кг·м²; номинальная мощность 490 кВт; максимальная частота вращения 375 мин⁻¹; $\tau_e = 1,5$ с.

Плотность жидкости принята равной 1000 кг/м³.

Геометрические размеры каналов для CFD-моделирования: прямоугольного — длина 140 м, ширина 80 м, высота 50 м; цилиндрического — длина 20 м; диаметр 10 м.

На рис. 2 приведена диаграмма коэффициента сопротивления C_d , который в системе уравнений (1), (2) представлен аналитическим выражением в виде оценки вычисленных значений методом наименьших квадратов (МНК) [14] полиномиальным рядом третьей степени.

На рис. 3 представлена МНК-оценка вычисленных значений коэффициентов K_M и K_F полиномами первой степени.

На рис. 4 представлены диаграммы, отражающие результаты интегрирования системы уравне-

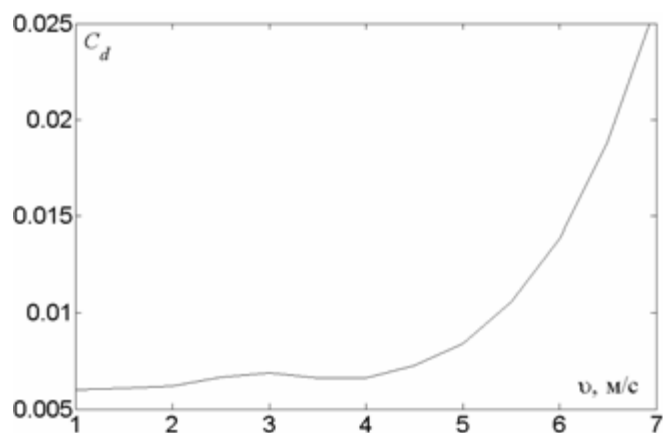


Рис. 2. Диаграмма коэффициента сопротивления корпуса судна

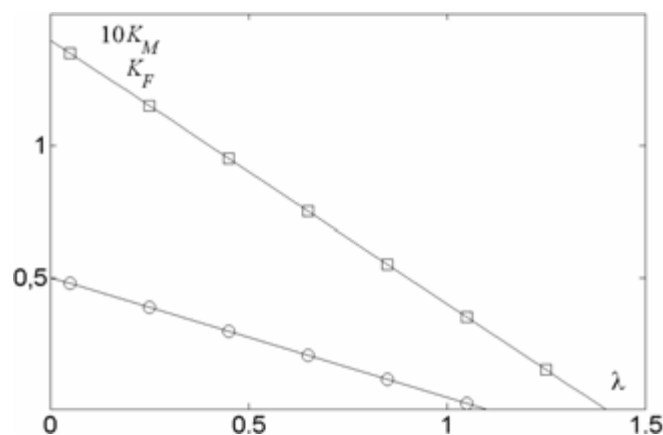


Рис. 3. Диаграммы коэффициентов: —□— K_M и —○— K_F ; $\lambda = 2\pi v_p(D\omega)^{-1}$

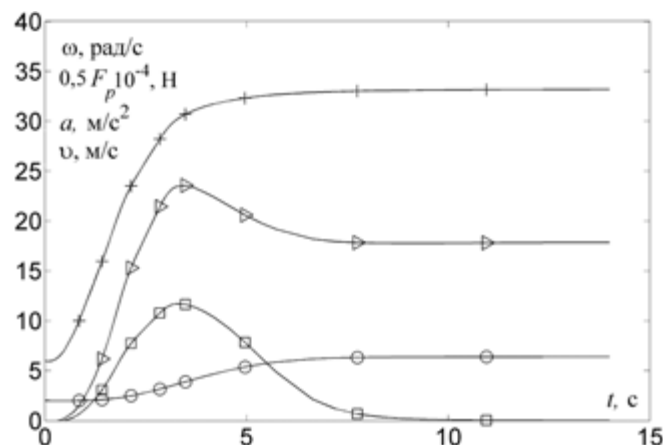


Рис. 4. Диаграммы: —+— — угловая скорость винта ω ; —▷— — упор винта F_p ; —□— — ускорение движения судна $a = \dot{v}$; —○— — скорость движения судна v

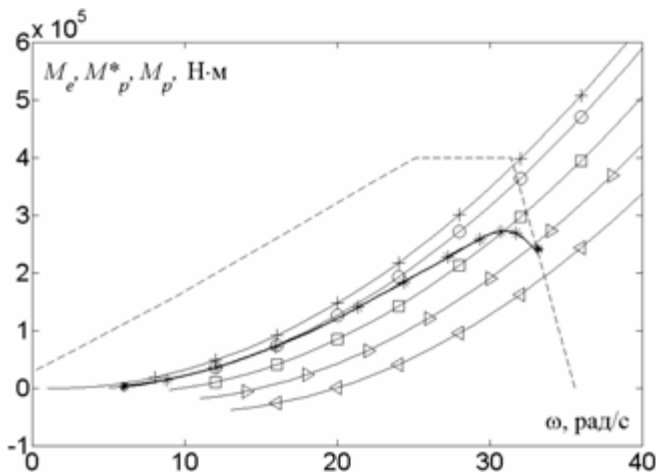


Рис. 5. Характеристика крутящего момента M_e двигателя (---); семейство винтовых характеристик момента M_p^* для различных скоростей обтекания винта встречным потоком: —+— $v_p = 1$ м/с, —○— $v_p = 2$ м/с, —□— $v_p = 4$ м/с, —▷— $v_p = 6$ м/с, —◁— $v_p = 8$ м/с; диаграмма момента M_p винта (-*-)

ний (1), (2). До разгона рассматриваемое судно двигалось с постоянной начальной скоростью 2 м/с и угловой скоростью вращения винта 6 рад/с. В момент времени $t = 0$ на двигатель была подана топливная смесь, соответствующая экстренному разгону судна (характеристика крутящего момента двигателя, соответствующая уровню топливной смеси для случая экстренного разгона судна показана на рис. 5).

Формально можно отметить совпадение положений максимума линейного ускорения судна, максимума силы, развиваемой винтом и завершающего этапа достижения винтом максимальной частоты вращения при заданном уровне подачи топливной смеси. При этом следует отметить, что положение указанных максимумов соответствует начальному этапу экстренного разгона судна.

Для выяснения причин формирования максимума ускорения рассмотрим совместно (рис. 5) характеристику крутящего момента двигателя, семейство характеристик момента винта и диаграмму изменения момента винта в процессе экстренного разгона судна.

Типичная характеристика крутящего момента M_e двигателя состоит из трех частей: ограничительная — восходящая прямая (предотвращает перегрузку двигателя), внешняя — горизонтальная прямая (предельный уровень крутящего момента при заданном расходе топлива) и регуляторная — спадающая прямая (ограничивает максимальную угловую скорость винта). Семейство характеристик момента винта (винтовые характеристики момента обозначены как M_p^*) построено по фор-

муле $M_p^* = K_M \rho \omega^2 D^5$ с учетом вычисленных (см. рис. 3) значений K_M . При этом для каждой кривой M_p^* задавалось постоянное значение скорости, а значение угловой скорости вращения винта варьировалось в заданных пределах.

Диаграмма момента M_p винта получена в результате интегрирования системы уравнений (1), (2). Из анализа диаграмм (см. рис. 5) следует, что в процессе увеличения угловой скорости вращения винта на интервале $6 < \omega < 20$ рад/с мгновенное значение момента M_p винта движется по винтовой характеристике, соответствующей 2 м/с, т. е. несмотря на то, что угловая скорость вращения винта увеличивается, судно за счет инерционных свойств продолжает двигаться со скоростью, которую оно имело до начала разгона.

В процессе дальнейшего увеличения угловой скорости вращения винта ($20 < \omega < 28$ рад/с) мгновенное значение момента M_p винта медленно переходит на другие винтовые характеристики, соответствующие более высоким значениям скорости, т. е. судно под действием сил упора винта, преодолевая инерцию, увеличивает свою скорость.

При достижении угловой скорости винта $\omega > 28$ рад/с кривая момента M_p винта резко отклоняется в сторону винтовых характеристик с высокими значениями скорости (4...6 м/с) судна, после чего увеличение скоростей вращения винта и движения судна замедляется по причине резкого снижения крутящего момента двигателя за счет его перехода с внешней на регуляторную характеристику. В итоге при $\omega \approx 32$ рад/с формируются максимумы ускорения движения судна и момента винта.

Отмечавшееся ранее (см. рис. 4) совпадение положений максимумов силы упора винта и ускорения движения судна объясняется тем, что характер зависимости K_M и K_F от $\lambda = 2\pi v_p (D\omega)^{-1}$, а следовательно, и от ω одинаков (см. рис. 3); отсюда — положения максимумов упора F_p и момента M_p винта также должны совпадать. В итоге положения всех трех максимумов — ускорения движения, момента и упора винта совпадают. При этом с физической точки зрения (2) формирование максимума ускорения обусловлено наличием максимума силы упора винта.

Максимум упора винта и обусловленный им максимум ускорения движения судна совпадают с завершающим этапом установления максимальной угловой скорости винта (см. рис. 4). Причина формирования такой корреляции объясняется достижением угловой скорости винта области действия "регуляторной" прямой, которая предотвращает дальнейшее увеличение угловой скорости винта.

Результаты численного моделирования подтверждаются полунатурными данными, полученными на морском тренажере [15], математическая

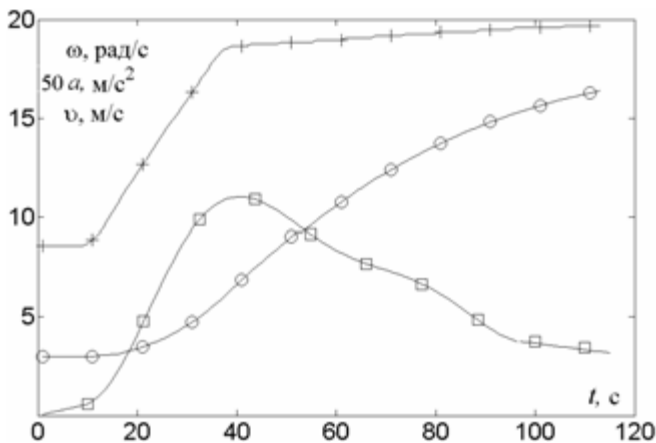


Рис. 6. Диаграммы экстренного разгона рыболовного судна: —+— угловая скорость винта ω ; —□— ускорение движения судна $a = v$; —○— скорость движения судна v

модель которого верифицирована на основе экспериментальных данных. На рис. 6 представлены диаграммы экстренного разгона рыболовного судна водоизмещением 3893 т, длиной 93 м, шириной 14,9 м, осадкой 5,8 м, двигатель которого имеет мощность 4416 кВт. Оценка линейного ускорения движения судна выполнена с помощью алгоритма Калмана [16].

Из диаграмм следует, что максимум ускорения при экстренном разгоне судна совпадает с завершающим этапом выхода винта на максимальную скорость вращения. Сравнивая диаграммы буксира и рыболовного судна (рис. 4 и рис. 6) можно отметить, что с увеличением размеров и соответственно инерционных свойств судна увеличивается наглядность вывода о том, что максимум ускорения движения судна является идентификационным признаком начального этапа экстренного разгона судна.

К основным существенным свойствам модели предложенной информационной технологии можно отнести следующие: выраженное прикладное значение, связанное с обеспечением безопасности мореплавания; полноту и достаточность предложенных процедур идентификации разгона судна; целостность и непротиворечивость, исключающие неоднозначность и ошибочность формирования выводов о движении судна; адекватность, подтвержденную численными исследованиями и натурными данными; высокую восприимчивость к реализации в существующих системах УДС.

Заключение

Причиной формирования информационного признака начала экстренного маневрирования в

виде максимума линейного ускорения судна является изменение во времени соотношения сил разнонаправленного действия: инерционных сил, определяющих характер нарастания скорости судна, и сил упора винта, характер развития которых предопределен характеристикой (типичной для морских судов) двигателя.

Информационный признак предоставляет возможность системам УДС с помощью радиолокационных средств выявлять формирование максимума линейного ускорения наблюдаемого судна и, тем самым, сформулировать вывод о выходе его гребного винта на максимальные обороты, т. е. о переходе судна к экстренному маневрированию. Выявленный информационный признак может быть использован не только в системах УДС, но также и судоводителями встречных судов и направлен на предупреждение морских катастроф, вызванных поздним выявлением ошибочного маневрирования встречных судов.

Список литературы

1. Romer H., Petersen H., Haastrup P. Marine accident frequencies-review and recent empirical results. // The J. of Navigation. 1995. Vol. 48. N 3. P. 410—424.
2. Справочник по теории корабля: В трех томах. Т. 3. / Под ред. Я. И. Войткунского. Л.: Судостроение, 1985. 544 с.
3. Васильев М. В. Определение динамических свойств судовых дизельных установок в условиях нормальной эксплуатации // Вестник МГТУ. 2006. Т. 9. № 2. С. 301—303.
4. Дорожко В. М. Оценка информационного признака начального этапа поворота морского судна // Мехатроника, автоматизация, управление. 2008. № 8. Приложение. С. 13—18.
5. Гофман А. Д. Движительно-рулевой комплекс и маневрирование судна. Справочник. Л.: Судостроение, 1988. 360 с.
6. Першиц Р. Я. Управляемость и управление судном. Л.: Судостроение, 1983. 272 с.
7. Камкин С. В., Возницкий И. В., Шмелев В. П. Эксплуатация судовых дизелей. М.: Транспорт, 1990. 344 с.
8. FLUENT 6.3 user's guide. Lebanon. Fluent. Inc., 2006. 2501 p.
9. Yakhot V., Orszag S. A. Renormalization Group Analysis of Turbulence: Basic Theory // Journal of Scientific Computing. 1986. Vol. 1. N 1. P. 1—51.
10. Hsu K. L., Chen Y. J., Chau S. W., Chien H. P., Kouh J. S. Ship Flow Computation of DTMB 5415 // CFD Workshop Tokyo 2005. Tokyo. Japan. March 9—11.
11. Wilcox D. C. Turbulence Modeling for CFD. DCW Industries Inc. La Canada. California, 1998.
12. Rhee S. H., Joshi S. Computational validation for flow around a marine propeller using unstructured mesh based Navier-Stokes Solver // JSME International Journal. Series B. 2005. Vol. 48. N 3. P. 562—570.
13. Мэтьюз Д., Финк Г., Куртис Д. Численные методы. Использование MATLAB. 3-е изд.: Пер. с англ. М.: Издательский дом "Вильямс", 2001. 720 с.
14. Дорожко В. М. Ортогональные многочлены в моделях регрессии наблюдений // Математическое моделирование. 2003. Т. 15. № 11. С. 45—50.
15. Navi-Trainer 4000. Mathematical Models. Technical Description. Transas. Marine. St. Petersburg, 2001. 102 p.
16. Калман Р., Фалб П., Арбиб М. Очерки по математической теории систем. М.: Мир, 1971.

Н. Б. Агаев, канд. техн. наук, доц.,
Национальная академия авиации, г. Баку,
e-mail:nadir_azisu@yahoo.com

Диагностирование предаварийного состояния систем газоснабжения на основе теории нечетких множеств

Предлагается методика прогнозирования и контроля предаварийной ситуации с применением математического аппарата нечетких множеств. Обсуждаемая методика рассчитана на ситуации, когда исходная информация, характеризующая состояние систем газопроводов, задается нечетко и для числовых, и для лингвистических переменных. Для описания нечетких значений параметра состояния выбрана экспоненциальная функция принадлежности с ограниченной областью.

Ключевые слова: газ, газопровод, аварийное состояние, нечеткая модель, нечеткие переменные, функция принадлежности.

Введение

Учитывая ведущую роль природного газа для экономического развития государства и необходимость снижения вредных воздействий на людей и окружающую среду, можно отметить, что в настоящее время остро стоят вопросы повышения безопасности и эффективности объектов системы транспортировки газа. Ситуация осложняется значительной погрешностью замеров технологических параметров, наличием различных видов неопределенностей. Описание таких неопределенностей с применением математического аппарата нечетких множеств позволяет создать высокоэффективную систему по прогнозированию и контролю предаварийной ситуации, координировать решения, принимаемые на каждом уровне системы, и получать четкие решения [1].

Постановка задачи

Для построения модели диагностирования предаварийной ситуации газопроводной системы принимаем следующую классификацию состояний: нормальное, предаварийное, аварийное. Как правило, эти состояния газопровода соответствуют вполне определенному характеру изменения параметров технологического процесса, т. е. каждый класс содержит в себе некоторые показатели, по значениям которых можно судить о состоянии анализируемого газопровода.

Так как при длительной эксплуатации оборудования и линейной части газопроводных систем изменяются их технические и технологические параметры, вследствие чего может нарушиться их нормальный режим функционирования, реакция трубопроводной системы на воздействия этих изменений неоднозначна и не всегда сопровождается нарушением общего технологического процесса в системе. В результате внешних воздействий различного типа могут образоваться утечки или закупорки на линейной части газопровода, а также могут неправильно срабатывать отключающие или регулирующие устройства. В большинстве случаев вид этого изменения установить практически невозможно, поэтому для первичной классификации состояний конкретных газопроводных систем требуется создать картотеку первичных состояний (текущих состояний). Картотеку можно использовать не только для диагностики отказов, имеющих место при эксплуатации системы, но и для определения причины ее аварийного состояния путем сравнения действительного изменения параметров с данными картотеки. Разработанный на этой основе метод расчета аварийных ситуаций отличается особой гибкостью и позволяет проводить анализ как нормального, так и аварийного режимов, исследовать работу трубопровода при неисправностях различных видов одного или нескольких типов одновременно. Каждая неисправность имеет свои особенности и требует определенных способов их учета в разработанной модели [2].

Как известно, появление в трубопроводе неисправностей любого вида вызывает изменения давления или пропускной способности на обоих концах или хотя бы на одном конце участка трубопровода. Например, полное или частичное перекрытие трубопроводов, которое происходит из-за отложения гидратов и конденсата или образования водяных пробок вследствие замерзания влаги в газопроводе и накопления различных частиц (пыли, песка, грязи, строительного мусора), обычно моделируют увеличением коэффициента гидравлического сопротивления определенного участка трубопровода, что отражает увеличение перепада давления на данном участке. При возникновении утечки вследствие коррозии или строительных неполадок газ вытекает в окружающую среду чаще всего с давлением, равным атмосферному, а иногда из-за достаточно высокого давления в газопроводе истечение становится критическим.

Нечеткая модель задачи

Для более детального описания предлагаемой методики и облегчения процесса диагностирования всю совокупность параметров делим на группы с общими признаками (табл. 1). Такие группы носят название "блок основных показателей". Как видно из табл. 1, некоторые признаки имеют количественные значения, некоторые из них выражаются терминами "низкое" или "недостаточное". Следует отметить, что предлагаемый набор показателей, характеризующих состояние газопроводной системы, не является полным и окончательным и может быть изменен за счет информации разной природы, таких как финансово-экономический показатель надежности, информация об уровне обеспечения ресурсов и т. д. Однако это не влияет на суть предлагаемой методики и только увеличивает надежность принятия решения.

Для описания нечетких значений параметра состояния выбираем экспоненциальную функцию принадлежности с ограниченной областью значений в виде [3]

$$\mu(x) = \exp[-k(x - r)^2], \quad (1)$$

Таблица 1

Основные признаки параметров газотранспортных систем по группам

Наименование показателей	Обозначения	Пороговые значения	
		Предаварийное	Аварийное
Блок технического состояния газопровода			
Срок эксплуатации	$T_{\text{экс}}$	$T_{\text{экс}}^{\text{п/ав}}$	$T_{\text{экс}}^{\text{ав}}$
Уровень загрузки	$U_{\text{загр}}$	$U_{\text{загр}}^{\text{п/ав}}$	$U_{\text{загр}}^{\text{ав}}$
Наличие защиты от коррозии	"Защита от коррозии"	Недостаточное	Низкое
Качество строительных работ	"Строит. работы"	Недостаточное	Низкое
Проведение ремонтных работ	"Ремонт"	Недостаточное	Низкое
Блок обеспечения технологического режима			
Давление в начале газопровода (минимальное)	$P_{<н}$	$P_{<н}^{\text{п/ав}}$	$P_{<н}^{\text{ав}}$
Давление в начале газопровода (максимальное)	$P_{>н}$	$P_{>н}^{\text{п/ав}}$	$P_{>н}^{\text{ав}}$
Давление в конце газопровода (минимальное)	$P_{<к}$	$P_{<к}^{\text{п/ав}}$	$P_{<к}^{\text{ав}}$
Давление в конце газопровода (максимальное)	$P_{>к}$	$P_{>к}^{\text{п/ав}}$	$P_{>к}^{\text{ав}}$
Частота изменения давления в начале газопровода в сутки	T_{P_n}	$T_{P_n}^{\text{п/ав}}$	$T_{P_n}^{\text{ав}}$
Частота изменения давления в конце газопровода в сутки	T_{P_k}	$T_{P_k}^{\text{п/ав}}$	$T_{P_k}^{\text{ав}}$

где k и r — коэффициенты, определяющие степень нечеткости и положение экспоненты на числовой оси.

При определении числового значения показателей, приведенных в табл. 1, эксперты указывают интервал соответствующих степеней неопределенностей. Например, минимальное давление в начале газопровода задается экспертом приблизительно так: "оно принимает значение от 17,5 атм. до 18,5 атм. с 5 %-ной шириной границ неопределенности". Учитывая это, для определения неизвестного коэффициента в (1) используем метод "построения функции принадлежности на основе экспертных оценок" [3]. За значение параметра r принимаем среднюю точку нечеткого интервала. Если обозначить уровень значимости α , процент с шириной значимого диапазона неопределенности β , то для каждой нечеткой границы интервала коэффициент k определяется следующей формулой:

$$k = -\frac{\ln \alpha}{(\beta r)^2}. \quad (2)$$

Например, функция принадлежности для минимального давления в начале газопровода при уровне значимости $\alpha = 0,1$ с 5 %-ной шириной границ неопределенности имеет вид:

- для левой границы интервала

$$\mu(x) = \exp[-3,007(x - 17,5)^2]; \quad (3)$$

- для правой границы интервала

$$\mu(x) = \exp[-2,691(x - 18,5)^2]. \quad (4)$$

Тогда функция принадлежности для высказывания "давление в начале газопровода принимает значение от 17,5 атм. до 18,5 атм. с 5 %-ной шириной границ неопределенности" представляется следующей формулой:

$$\mu(x) = \begin{cases} \exp[-3,007(x - 17,5)^2], & x \leq 17,5; \\ 1, & 17,5 < x < 18,5; \\ \exp[-2,691(x - 18,5)^2], & x \geq 18,5. \end{cases} \quad (5)$$

Если показатель выражается нечетким термом (табл. 1, показатель "защита от коррозии") с лингвистическим значением, то для построения функции принадлежности используется "параметрический подход к построению функции принадлежности" [3]. Необходимость использования этого метода обусловливается нечетким высказыванием эксперта при оценке показателя. Например, при оценке показателя "защита от коррозии" вместо основного лингвистического значения "недостаточно", эксперт может использовать термы типа "более-менее достаточно" или "не очень достаточно" и т. д. При этом один из термов представляется модификацией другого. Для этого эксперту предлагается на 100 %-ной шкале ука-

зять три точки, из которых две крайние точки по его мнению не принадлежат описываемому лингвистическому значению, а другая определенно принадлежит данному терму. Задача состоит в том, чтобы, используя параметры основных термов описать переход к не основному терму с помощью упорядоченного отношения "меньше". Как показано в [3], для решения этой задачи используется аппарат автоморфных функций:

$$v \rightarrow \frac{\alpha x + \beta}{\gamma x + \delta}. \quad (6)$$

Неизвестные коэффициенты α , β , γ , δ определяются, если известны три значения основного терма (z_1 , z_2 , z_3) и модифицированного терма (w_1 , w_2 , w_3):

$$\begin{aligned} \alpha &= z_1 z_2 (w_1 - w_2) + z_1 z_3 (w_3 - w_1) + z_3 z_2 (w_2 - w_3); \\ \beta &= w_1 w_2 z_3 (z_1 - z_2) + w_1 w_3 z_2 (z_3 - z_1) + w_3 w_2 z_1 (z_2 - z_3); \\ \gamma &= z_2 (w_1 - w_3) + z_1 (w_3 - w_2) + z_3 (w_2 - w_1); \\ \delta &= w_1 w_2 (z_1 - z_2) + w_1 w_3 (z_3 - z_1) + w_3 w_2 (z_2 - z_3). \end{aligned} \quad (7)$$

Используя эти коэффициенты, можно определить обратный переход от основного терма к модифицированному. Например, если эксперт, описывая состояние газопровода по отношению к показателю "защита от коррозии", использует лингвистический терм "низкое" и указывает три значения по процентам: $z_1 = 20$, $z_2 = 25$, $z_3 = 30$, из которых, по его мнению, z_1 и z_3 еще не принадлежат описываемому лингвистическому значению, то z_2 определенно принадлежит ему.

Переходим от этого терма к терму "очень низкое": $w_1 = 20$, $w_2 = 22$, $w_3 = 25$. Переход осуществляется по формуле (7). Коэффициенты для описания перехода:

$$\alpha = -255; \beta = 4000; \gamma = -5; \delta = 50. \quad (8)$$

Используя найденное преобразование, перейдем от терма "недостаточное" — $z_1 = 35$, $z_2 = 40$, $z_3 = 50$ к терму "очень недостаточное": $w_1 = 35$, $z_2 = 40$, $z_3 = 42,5$ или к терму "очень-очень недостаточное": $w_1 = 39,3$, $z_2 = 40$, $z_3 = 40,7$.

Для распознавания ситуации нужно создать базу знаний, которая содержит набор используемых высказываний, логических операций и правил. На начальном этапе для создания базы знаний используется общетеоретическая информация и суждения экспертов о рассматриваемом газопроводе в виде описания классов. При этом классы описываются такими высказываниями экспертов, как: "Если A и B и ..., ТО это класс C ". В этой записи A и B — условия, налагаемые на значения параметров; C — наименование классов. Например: "если давление в начале газопровода P_H меньше 20 атм. и больше 18 атм., то это класс нормальных состояний".

Для разработки методики распознавания ситуации необходимо создать базу знаний по всем блокам, указанным в табл. 1. Как отмечено выше, каждый параметр, входящий в блок, сформулирован с использованием пороговых значений. Согласно указанной выше классификации для определения класса вводим следующие правила:

- если значения всех параметров лучше предаварийного состояния, то ситуация принимается нормальной;
- если хотя бы один параметр хуже предаварийной ситуации, а остальные параметры характеризуются нормальными, то ситуация принимается предаварийной;
- если хотя бы один параметр хуже аварийной ситуации, то ситуация признается аварийной.

Поясним термины "лучше" и "хуже". Как указано выше, при определении классов эти термины используются для четкого нахождения границ классов.

- Для параметров "Блок обеспечения технологического режима" они вводятся следующим образом: если значения давления или расхода "не меньше минимального и не больше максимального порогового значения" и "частота давления меньше порогового значения", то ситуация "лучше", иначе "хуже".
- Для параметров "Блок технического состояния газопровода": если значения параметров "меньше своего порогового значения" или "модифицированный терм выражает превосходство по отношению к пороговому значению, то ситуация "лучше", иначе "хуже". Если в базе знаний имеются параметры других блоков, аналогично можно определить эти термины для них.

Следует отметить, что возможно применение другого определения ситуации, например, предаварийную и аварийную ситуации можно определить одновременно по нескольким значениям параметров.

Высказывания могут формулироваться с применением лингвистических переменных, нечетких значений параметров, нечетких отношений и нечетких логических связей. Для описания одного класса может быть использовано несколько подобных высказываний, которые связаны между собой логической связкой "ИЛИ". При этом используются правила операции объединения и пересечения нечетких множеств соответственно:

- для логической связки "И"
$$\mu(x) = \min[\mu_1(x), \mu_2(x)]; \quad (9)$$

- для логической связки "ИЛИ"
$$\mu(x) = \max[\mu_1(x), \mu_2(x)], \quad (10)$$

где $\mu_1(x)$ и $\mu_2(x)$ — функции принадлежности условий, налагаемых на значения параметров.

Состав показателей параметров газопровода "Говсан-Дигях" и их пороговые значения

Наименование показателей	Обозначения	Пороговые значения	
		Предаварийное	Аварийное
Блок технического состояния газопровода			
Срок эксплуатации, год	$T_{\text{эксп}}$	25	30
Уровень загрузки, %	$U_{\text{загр}}$	45	30
Наличие защиты от коррозии	"Защита от коррозии"	Недостаточное	Низкое
Качество строительных работ	"Строит. работы"	Недостаточное	Низкое
Проведение ремонтных работ	"Ремонт"	Недостаточное	Низкое
Качество эксплуатируемой среды	"Газ"	Жирные с мин. вод. паром	Жирные с макс. вод. паром
Блок обеспечения технологического режима			
Давление в начале газопровода (минимальное), атм.	$P_{<H}$	10	5
Давление в начале газопровода (максимальное), атм.	$P_{>H}$	19	22
Давление в конце газопровода (минимальное), атм.	$P_{<K}$	5	1
Давление в конце газопровода (максимальное)	$P_{>K}$	15	19
Частота изменения давления в начале газопровода в сутки, 1/сут	T_{P_H}	5	8
Частота изменения давления в конце газопровода в сутки, 1/сут	T_{P_K}	5	8

Если описание классов сформулировано в виде множества сочетаний высказываний по всем параметрам, то любая ситуация может быть отнесена к тому или иному классу. С этой целью для распознавания предъявленной ситуации вводим степень принадлежности ситуации к тому или иному классу. Для этого по всем параметрам и для всех классов определяем функцию принадлежности. Кроме того определяем функцию принадлежности по каждому высказыванию. При этом используем указанное выше правило операции над нечеткими параметрами. Степень принадлежности ситуации к тому или иному классу определяем по следующей формуле:

$$\pi_k = \max_i \{ \min_j [\sup_{x \in X_j} (\min(\mu_j(x), \mu_{kij}(x)))] \}, \quad (11)$$

где k — номер класса; π_k — степень принадлежности ситуации классу с номером k ; i — номер высказывания; j — номер параметра; X_j — область значений j -го параметра; $\mu_j(x)$ — функция принадлежности j -го параметра; $\mu_{kij}(x)$ — функция принадлежности i -го высказывания по j -му параметру внутри k -го класса.

По предлагаемой методике в зависимости от характера базы знаний можно решать задачи по оценке состояния газотранспортной системы в целом — по всем республикам, по отдельным регионам или по отдельным газопроводам. Например, если требуется оценить состояние газотранспортной системы в целом по республике, то в базу знаний должны быть включены блок обеспеченности газовым топливом, блок использования основных производственных фондов, экологический, финансово-экономический и другие блоки.

Численный эксперимент

Применим предлагаемую методику диагностирования по оценке состояния газопровода "Говсан-Дигях". Отметим, что подобным образом можно использовать аналогичные методики для любых газотранспортных систем.

Опишем классы исследуемого газопровода (табл. 2).

- Если $T_{\text{эксп}} < 25$ И $U_{\text{загр}} > 45$ И Защита от коррозии = "высокое" И Строит. работы = "высокое" И Ремонт = "высокое" И газ = "сухое" и $10 < P_H < 19$ И $5 < P_K < 15$ И $T_{P_H} \leq 4$ И $T_{P_K} \leq 4$ ТО это класс "нормальное".
- Если $25 < T_{\text{эксп}} < 30$ И $5 < P_H < 10 \vee 19 < P_H < 22$ И $2 < P_K < 5 \vee 15 < P_K < 17$ ИЛИ $30 < U_{\text{загр}} < 45$ И Защита от коррозии = "недостаточное" И $5 < P_H < 10 \vee 19 < P_H < 22$ И $2 < P_K < 5 \vee 15 < P_K < 17$ ИЛИ $25 < T_{\text{эксп}} < 30$ И Строит. работы = "недостаточное" И $5 < P_H < 10 \vee 19 < P_H < 22$ И $2 < P_K < 5 \vee 15 < P_K < 17$ ИЛИ $25 < T_{\text{эксп}} < 30$ И Ремонт = "недостаточ-

ное" И $5 < P_H < 10 \vee 19 < P_H < 22$ И $2 < P_K < 5 \vee 15 < P_K < 17$ ИЛИ газ = "Жирные с мин. вод. паром" И $5 < P_H < 10 \vee 19 < P_H < 22$ И $2 < P_K < 5 \vee 15 < P_K < 17$ И $5 < T_{P_H} < 8$ И $5 < T_{P_K} < 8$ ТО это класс "предаварийное".

- Если $30 < T_{\text{эксп}}$ И $P_H < 5 \vee P_H \geq 22$ И $P_K < 2 \vee P_K \geq 19$ ИЛИ $U_{\text{загр}} < 30$ И Защита от коррозии = "низкое" И $P_H < 5 \vee P_H \geq 22$ И $P_K < 2 \vee P_K \geq 19$ ИЛИ Строит. работы = "низкое" И $P_H < 5 \vee P_H \geq 22$ И $P_K < 2 \vee P_K \geq 19$ ИЛИ Ремонт = "низкое" И $P_H < 5 \vee P_H \geq 22$ И $P_K < 2 \vee P_K \geq 19$ ИЛИ газ = "Жирные с макс. вод. паром" ИЛИ $P_H < 5 \vee P_H \geq 22$ И $P_K < 2 \vee P_K \geq 19$ И $T_{P_H} \geq 8$ И $T_{P_K} \geq 8$ ТО это класс "аварийное".

Текущее состояние исследуемого газопровода: газопровод "Говсаны-Дигях" построен в 1974 г., диаметр газопровода 700 мм, длина 30 144 м. В период исследования давление в начале газопровода принимает значение от 17,5 до 18,5 атм., а в конце от 12,5 до 17,5 атм. Уровень загрузки более 50 %, защита от коррозии не очень достаточная, строительные работы высокие, ремонт — низкий, газ — сухой.

По указанной выше методике построим соответствующие функции принадлежности для текущего состояния исследуемого газопровода. Уровень значимости принимаем 0,1, ширина границ неопределенности 5 %. Тогда функция принадлежности давления в начале газопровода описывается формулой (5), а функция принадлежности в конце газопровода формулой:

$$\mu_{P_k}(x) = \begin{cases} \exp[-5,895(x - 12,5)^2], & x \leq 12,5; \\ 1, & 12,5 < x < 17,5; \\ \exp[-3,007(x - 17,5)^2], & x \geq 17,5. \end{cases} \quad (12)$$

Аналогичным образом построим функцию принадлежности для уровня загрузки газопровода:

$$\mu_Z(x) = \begin{cases} \exp[-0,3684(x - 50)^2], & x \leq 50; \\ 1, & 50 < x < 100. \end{cases} \quad (13)$$

Для лингвистических параметров принимаем одинаковую шкалу с основными терминами. На основе вышеуказанной методике по формулам (6)–(7) определяются модифицированные термы.

Для определения степени принадлежности ситуации к тому или иному классу используем формулу (11). С этой целью для каждого класса нужно построить функцию принадлежности. Для примера построим функцию принадлежности давления в начале и в конце для нормальной, предаварийной и аварийной ситуаций. Как показано выше, при определении класса нормального состояния давление в начале присутствовало только один раз, поэтому число высказываний $i = 1$. Для класса "нормальный" принимаем $k = 1$. Класс "нормальный" по отношению к параметрам P_H и P_K задан нечетким интервалом, поэтому функция принадлежности строится аналогично (5). Учитывая это, по вышеуказанной методике можно вычислить функцию принадлежности по параметрам давления в начале газопровода (P_H):

$$\mu_{11P_H}(x) = \begin{cases} \exp[-9,210(x - 10)^2], & x \leq 10; \\ 1, & 10 < x < 19; \\ \exp[-2,551(x - 19)^2], & x \geq 19 \end{cases} \quad (14)$$

и функцию принадлежности по давлению в конце газопровода (P_K):

$$\mu_{12P_K}(x) = \begin{cases} \exp[-36,841(x - 5)^2], & x \leq 5; \\ 1, & 5 < x < 15; \\ \exp[-4,093(x - 15)^2], & x \geq 15. \end{cases} \quad (15)$$

Аналогичным образом определяем функцию принадлежности для других параметров: срока эксплуатации, уровня загрузки и др. Все найденные функции принадлежности учтем в (11):

$$\pi_1 = \max_i \{ \min [\sup_{T_{\text{эксп}}} (\min(\mu_{\text{эксп}}(x), \mu_{11\text{эксп1}}(x)))],$$

$$\min [\sup_Z (\min(\mu_Z(x), \mu_{11Z}(x)))];$$

$$\min [\sup_{\text{КОРР}} (\min(\mu_{\text{КОРР}}(x), \mu_{11\text{КОРР}}(x)))],$$

$$\min [\sup_{\text{СТ.РЕМ.}} (\min(\mu_{\text{СТ.РЕМ.}}(x), \mu_{11\text{СТ.РЕМ.}}(x)))];$$

$$\min [\sup_{\text{РЕМ}} (\min(\mu_{\text{РЕМ}}(x), \mu_{11\text{РЕМ}}(x)))],$$

$$\min [\sup_{\text{ГАЗ}} (\min(\mu_{\text{ГАЗ}}(x), \mu_{11\text{ГАЗ}}(x)))];$$

(16)

$$\min [\sup_{P_H} (\min(\mu_{P_H}(x), \mu_{11P_H}(x)))],$$

$$\min [\sup_{P_K} (\min(\mu_{P_K}(x), \mu_{11P_K}(x)))].$$

По формуле (16) найдем степень принадлежности текущего состояния газопровода "Говсаны-Дигях" к нормальной ситуации: $\pi_1 = 0,454$.

Рассмотрим класс "предаварийный": $k = 2$. Высказывания по данному параметру использовались пять раз с логической связкой "И", т. е. в формуле (11) $i = 1, \dots, 5$. Кроме того, каждое высказывание по этому параметру содержит логическую связку ИЛИ, поэтому при вычислении функции принадлежности следует использовать (10). Например, для логической связки $5 < P_H < 10 \vee 19 < P_H < 22$ функция принадлежности принимает вид

$$\mu_{21P_H}(x) = \max \left[\begin{array}{l} \left\{ \begin{array}{l} \exp[-36,841(x - 5)^2], \quad x \leq 5; \\ 1, \quad 5 < x < 10; \\ \exp[-9,210(x - 10)^2], \quad x \geq 10; \end{array} \right. \\ \left\{ \begin{array}{l} \exp[-2,551(x - 19)^2], \quad x \leq 19; \\ 1, \quad 19 < x < 22; \\ \exp[-1,903(x - 22)^2], \quad x \geq 22. \end{array} \right. \end{array} \right] \quad (17)$$

После некоторого преобразования окончательно получаем:

$$\mu_{21P_H}(x) = \begin{cases} \left\{ \begin{array}{l} \exp[-36,841(x - 5)^2], \quad x \leq 5; \\ 1, \quad 5 < x < 10; \\ \exp[-9,210(x - 10)^2], \quad 10 < x < 13,1; \end{array} \right. \\ \left\{ \begin{array}{l} \exp[-2,551(x - 19)^2], \quad 13,1 \leq x \leq 19; \\ 1, \quad 19 < x < 22; \\ \exp[-1,903(x - 22)^2], \quad x \geq 22. \end{array} \right. \end{cases} \quad (18)$$

Аналогичным путем можно определить функцию принадлежности для давления в конце газопровода с логической связкой $2 < P_K < 5 \vee 15 < P_K < 17$

$$\mu_{21P_K}(x) = \begin{cases} \left\{ \begin{array}{l} \exp[-230,259(x - 2)^2], \quad x \leq 2; \\ 1, \quad 2 < x < 5; \\ \exp[-36,814(x - 5)^2], \quad 5 < x < 7,5; \end{array} \right. \\ \left\{ \begin{array}{l} \exp[-4,093(x - 15)^2], \quad 7,5 \leq x \leq 15; \\ 1, \quad 15 < x < 17; \\ \exp[-3,187(x - 17)^2], \quad x \geq 17 \end{array} \right. \end{cases} \quad (19)$$

и функция принадлежности параметра "срок эксплуатации"

$$\mu_{12\text{эксп}}(x) = \begin{cases} \exp[-1,474(x - 25)^2], & x \leq 25; \\ 1, & 25 < x < 30; \\ \exp[-1,023(x - 30)^2], & x \geq 30. \end{cases} \quad (20)$$

Аналогичным образом определяем функцию принадлежности для других параметров. Все найденные функции принадлежности учтем в (11) и вычислим степень принадлежности текущего состояния газопровода "Говсаны-Дигях" к предаварийной ситуации: $\pi_2 = 1$.

Так же определяем степень принадлежности текущего состояния газопровода "Говсаны-Дигях" к аварийной ситуации: $\pi_3 = 0,7891$.

По найденным значениям степени принадлежности можно сформулировать суждение о принадлежности текущего состояния газопровода "Говсаны-Дигях": текущее состояние газопровода четко принадлежит к предаварийному состоянию, в какой-то мере принадлежит к аварийному состоянию и не принадлежит к нормальному состоянию.

Заключение

Обобщая результаты диагностирования состояния газопроводных систем по предлагаемой методике, можно заключить следующее:

- преимущество предложенной методики по сравнению с другими заключается в том, что она не требует более строгого обоснования пороговых значений и может быть применена при отсутствии точной информации о значениях параметров;
- предлагаемая методика не требует информации о конкретной топологической структуре исследуемого объекта, поэтому ее можно применить не только к отдельно взятым газопроводам, но и к любой газотранспортной системе, такой как газотранспортная система г. Баку или газотранспортная система в целом по республике.

Список литературы

1. Алтунин А. Е., Семухин М. В. Модели и алгоритмы принятия решений в нечетких условиях. Тюмень: Изд-во Тюменского государственного университета, 2000. 352 с.
2. Салихов И. А. Безопасность транспорта газа по трубопроводным системам. // Материалы Междунар. научно-техн. конф. "Трубопроводный транспорт — сегодня и завтра". Научные труды. Уфа, 27—29 ноября 2002. С. 243, <http://www.ogbus.ru/library>.
3. Борисов А. Н., Крумберг О. А., Федоров И. П. Принятие решений на основе нечетких моделей. Рига: Зинатне, 1990. 184 с.

ИНФОРМАЦИЯ

Лидеры Flash-индустрии встретились в Крыму

10—16 августа в Коктебеле (Украина) состоялась ежегодная конференция специалистов в области информационных технологий ITSea 2009. В этом году мероприятие было посвящено вопросам Flash- и Flex-разработки. В нем приняли участие специалисты из России, Украины, Белоруссии и ближнего зарубежья. Мероприятие было совместно организовано украинской и российской группами пользователей Adobe (Ukrainian and Russian Adobe Flash Platform User Groups).

Большинство докладов было посвящено игровой тематике. Виталий Хить (Украина), соучредитель компании Absolutist и один из авторитетных специалистов в области игровой индустрии, рассказывал о казуальных играх. Российская "Альтернатива" продемонстрировала возможности разработки игр на собственной платформе, в том числе о своей последней игре TankiOnline. А Валерия Маллаева (Украина), организатор международной конференции FlashGamm, затронула про-

блемы взаимодействия flash-программистов и аниматоров при разработке игр.

О корпоративных решениях говорили специалисты компании "Флексис" — генеральный директор Артемий Малков и технический директор Александр Гахов. Доклад Артемия Малкова был посвящен Flash Catalyst (средство для быстрой разработки интерфейсов на базе Flash-платформы). По его словам, использование данной технологии "примерно в три раза ускоряет процессы проектирования и графического оформления интерфейсов". Александр Гахов говорил об организации эффективной разработки Enterprise-проектов.

"Подобные мероприятия укрепляют взаимодействие между профессионалами из разных стран. Это очень важно для обмена опытом, контактами, идеями и контрактами, — комментирует Артемий Малков. — И что особенно приятно, пока политики делят границы, газ, черноморский флот и Крым, профессионалы России и Украины легко находят общий язык и замечательно проводят время вместе!"

Программа экономии и качества на базе современных IT-технологий

Новаторские проекты казахстанских железнодорожников активно внедряются в жизнь. Один из них — программа "АРМ (автоматизированные рабочие места) нарядчика экономического и технического отделов", разработанная Оксаной Баклюк, инженером-программистом Карагандинского локомотивного эксплуатационного депо — филиала АО "Локомотив". Над этой программой молодой новатор работала несколько месяцев. Впервые она презентовала ее руководству депо больше трех лет назад. Программа пришлась по вкусу не только самим нарядчикам, которые непосредственно ощутили ее плюсы перед устаревшей программой, но и вызвала интерес на высшем уровне — в АО "Локомотив" (дочернее предприятие АО "НК "Казахстан темир жолы"). В данный момент программа, как пилотный проект, внедрена на "родном" Оксанином депо.

По словам Оксаны Баклюк, программа позволит нарядчикам полностью исключить ручное ведение отчетности, сэкономить их время и усилия, а также проводить более качественно приемку и обработку маршрутных листов, которые вручаются локомотивной бригаде перед поездкой.

Главная задача программы "АРМ нарядчика" — автоматизировать работу нарядчиков локомотивных эксплуатационных депо. С ней она справляется. Если ранее работникам самим нужно было вручную рассчитывать, какого числа, какая локомотивная бригада отправляется в поездку, то теперь программа сама делает все нужные расчеты, отмечает маршруты и точно выставляет числа поездок. В карагандинском депо с помощью "АРМ нарядчика" одним нажатием кнопки можно получить расчетные данные по девятистам работникам депо и распечатать их для планирования работ. В то же время при ручных расчетах данные по каждому машинисту и помощнику машиниста требуют долгих поисков и обработки. При этом не только тратится время и повышается утомляемость нарядчика, но и снижается его внимательность и, как следствие, появляются ошибки при расчетах. А это может привести к подаче недостоверной информации, — говорит она.

Программа Оксаны Баклюк имеет массу преимуществ. В частности, вся информация с маршрутного листа "вбивается" в программу, поэтому можно отказаться от большого количества журна-

лов. Согласно регламенту отработанные журналы необходимо хранить более трех лет, из-за этого кабинеты нарядчиков превращались в архивы пожелтевших "канцелярских книг". Теперь этот вопрос кардинально пересмотрен. Ежемесячно все данные автоматизированной программы аккумулируются в электронном архиве, создаются их резервные копии. Они хранятся необходимый период времени на электронных носителях, которые занимают небольшую площадь в письменном столе нарядчика.

Программа, разработанная карагандинским программистом, весьма удобна и проста в применении. Она позволяет буквально в считанные минуты раскрыть полную картину: когда локомотивная бригада приступила к работе, сколько ночей она пробыла в пути, сколько времени уделено для отдыха. Ежемесячно программа дает отчет по каждому работнику: индивидуально рассчитывается отработанная норма за месяц, сколько дней работник пробыл на больничном, в ежегодном отпуске или без содержания, в командировке, на учебной сессии, сколько времени замещал коллегу. На основании дневной и месячной нормы часов программа введет подсчет переработки и недоработки на текущий момент для каждого работника. Программа позволяет также вести учет количества случаев и сверхнормативного времени в поездке с момента отправления из депо и прибытия на конечный пункт маршрута и с момента прибытия в депо из конечного пункта. Кроме того, программа содержит журналы закрепленных участков обслуживания бригад, графики закрепленных плечей, журналы раскрепления бригад по колоннам, а также фамилии инструкторов и телефоны сотрудников. А это позволяет оперативно найти нужную информацию.

Уже сейчас нарядчики из других регионов, которые еще работают по старинке, вооружившись ручкой и бумагой, проявляют интерес и желание внедрить новаторский опыт Караганды на своих рабочих местах. По словам начальника резерва локомотивных бригад Астанинского локомотивного депо Екатерины Кирик, новая программа будет востребована в столичном депо. В данный момент астанинские нарядчики работают по "гомельской" программе, внедренной 25 лет назад, которая уже не отвечает современным требованиям. Поэтому они с нетерпением ждут перемен. Как заверили в АО "Локомотив", до конца года эта программа будет внедрена еще в пяти депо.

Любовь Артымович

CONTENTS

Vasenin V. A. *Critical Energetics Infrastructure: Means of Counteraction to Cyberterrorism* 2

The present publication continues the topic, present in the publication "Critical Energetic Infrastructure: Cyberterrorist Threat". It is dedicated to problems of protecting critically important energetic infrastructure from destructive computer impacts with terrorist aims. Methods and means of protecting such objects based on complex approach to information security are highlighted.

Keywords: energetics infrastructure, fuel and hydro-energy complexes, nuclear energetics, destructive impact, critically important object, methods and means of protection, information security, complex approach.

Remizov A. V., Filippov M. V., Chichvarin N. V. *Methods of the Protection of Information in the Acoustic Files* 10

This paper examines information hiding methods using MP3 audio files. MP3 coding scheme is shown. New hiding methods are proposed, both for hiding large amounts of information and for hiding information with high resistance against detection. Comparative results of existing and proposed hiding methods' speed, capacity and resistance against detection are shown.

Keywords: steganography, steganalysis, steganographic key, message, lossy compression, bitrate, quantization, entropy encoding, Huffman coding, system of linear equations.

Shapchenko K. A. *An Approach to Verifying Security Properties in Access Control Models with a Hierarchy of Access Objects* 13

In this paper we discuss mathematical problems of specifying and verifying security properties in access control models. We propose a model checking approach to verifying a class of such properties, namely information flow constraints. Applications of this approach are described in the context of analyzing access control models in Unix-like operating systems considering the tree-like hierarchy induced on access objects.

Keywords: information security, access control model, model checking.

Mochalov V. A. *Algorithms of Transit Nodes Accommodation in Sensor Networks* 18

The proposed algorithms accommodate transit nodes (T-nodes) in the sensor network (SN). At the outset, describes the basic algorithms for placing T-nodes and an overall performance of such algorithms are given. Then into the structure of SN excess T-nodes are added according to one of the fault-tolerant algorithms of T-nodes accommodation, which increase probabilities of functional nodes connectivity with a server. In this paper we consider two fault-tolerant algorithms of T-nodes accommodation: the first is based on the use of genetic algorithms, the second puts T-nodes under a given degree of functional nodes connectivity with a server.

Keywords: sensor network, fault-tolerance, connectivity, algorithms.

Ognev V. A., Ivanov S. R. *GLONASS/GPS Receiver Mathematical Model* 24

In this paper the generic structure of GLONASS/GPS receivers are shown and the simplified mathematical model of such devices is presented. This model allows to conduct impact evaluation of the Gaussian interference on the receiver operability. In paper it is shown that the receiver anti jam capabilities are determined by performance of tracking loops which used for signal parameter estimation. The various types of tracking loops are evaluated and comparison of their qualities is made. The numerical modeling results are presented.

Keywords: global navigation satellite system (GNSS), antijam capability.

Petunin A. A. <i>Automation of Nesting: the Basic Methods of Approach to Development of CAD/CAM System for the Nesting and the Preparation of NC Programs for Cutting Machines of Sheet Material</i>	29
---	----

This article describes the structure of the software of modern CAD/CAM system for automated cutting of sheet materials in complicated geometrical shapes, and for preparation of control programs for the equipment with numerical program control (NC units) which uses thermal or waterjet technology of cutting. The new approach to creation of the subsystem for automatic irregular cutting is offered. The requirements to the basic modules of the system are formulated and justified.

Keywords: CAD/CAM system, nesting, cutting of material, subsystem of computer-aided system, NC machines, control program.

Nemtinov V. A., Egorov S. Ya., Pahomov P. I. <i>Fuzzy-Set Theory and Expert Systems Application to Perform Automatically Selection of an Engineering System Element</i>	34
--	----

The article examines various aspects of expert systems' theory application at development of automated choice procedural model of isolation pipeline valves for water-supply and heat networks.

Keywords: theory of fuzzy sets, automated choice, procedural model, isolation pipeline valves.

Aliguliyev R. M. <i>Document Summarization by Sentence Ranking</i>	39
---	----

The method automatic document summarization through sentence ranking is proposed. For ranking of sentences a modified variant of the PageRank algorithm is applied.

Keywords: document summarization, sentence ranking, PageRank algorithm, resistance distance.

Danchul A. N., Novikov A. P. <i>Representation of Problem Area Knowledge on the Basis of Gnoseological Model</i>	44
---	----

Conceptions of representation of problem area knowledge on the basis of gnoseological model are stated. Formation of network structure in gnoseological model which leads to representation of problem area knowledge in a kind quazigraph which concept is entered in article is considered. The created shell of the expert system realising offered concepts is noticed, that, possesses necessary efficiency of actualisation of knowledge.

Keywords: gnoseological model, knowledge representation, network structure.

Mukhacheva E. A., Zaminova A. R. <i>Evolution Algorithms' Efficiency Research for Solving Tasks of Discrete Optimization by the Example of Travelling Salesman Problem</i>	51
---	----

The methods of evolution algorithms' using for solving travelling salesman problem are presented in this work. We regard genetic algorithm and evolution strategies by the example of $(\mu + \lambda, \lambda)$ -ES, (μ, λ) -ES and in particular case — strategy $(1 + 1)$ -ES as evolution algorithms. Numerical results presented show the efficiency of simple evolution strategy.

Keywords: genetic algorithm, evolution strategies $(\mu + \lambda, \lambda)$ -ES, (μ, λ) -ES and $(1 + 1)$ -ES, travelling salesman problem.

Volovikova E. V., Uvaisov S. U. <i>The Method of Electronics Test Founded On Complex Electro Thermal Simulation</i>	57
--	----

There is modify fault dictionary method for analog devices tests founded on complex electro thermal simulation in the article. Advanced method allows improving reliability of test because of tacking in account temperature influence on elements parameters.

Keywords: analog electronics, electronics fault test, fault dictionary method, electro thermal simulation.

Sekaev V. G. <i>Use of Algorithms of a Heuristics Combination at Construction of Optimum Schedules</i>	61
---	----

The problem of drawing up of optimum schedules with use of genetic algorithms of a heuristics combination is considered. The offered approach allows making quasioptimal schedules for the big class of multiphase serving systems.

Keywords: genetic algorithm, heuristics combination method, schedule, technological cycle period, heuristics set, task assignment.

Dorozhko V. M. *Information Technology of Initial Stage Identification of the Sea Vessel Racing* 65

On the basis of research of dynamical motion equations the information technology of initial stage identification of emergency racing of an ocean vessel is developed. It is revealed, that its informational indication is the maximum of linear vessel acceleration. By numerical modeling it is shown, that a reason of informational indication formation is the concurrence of various forces: the inertial forces defining character of vessel acceleration, and forces of screw thrust which character of development is predetermined by an engine performance. The revealed informational indication can be used in vessel traffic systems, and also by captains of the counter vessels and is directed on the warning of the marine catastrophes caused by late detection of erratic maneuvering of the vessels.

Keywords: information technology, maneuver vessel identification, maximum vessel acceleration, screw thrust, warning of marine catastrophes.

Agaev N. B. *Diagnosing of Preemergency Condition of Systems of Gas Supply by Method of the Theory of Fuzzy Sets* 71

With application of mathematical apparatus of fuzzy theory is offered the forecasting and the control of the preemergency case. The offered technique is calculated on situations when the initial information characterising condition of systems of gas pipelines, is fuzzy variables both on numerical, and on linguistic variables. Thus simple relations between variables are described by means of fuzzy statements, and difficult relations — by fuzzy algorithms. The following classification of conditions is accepted: normal, preemergency, emergency. For the description of fuzzy values of parametre of fitting function of an accessory with the limited area is chosen exponential function.

Keywords: gas, gas pipeline, emergency cases, fuzzy model, fuzzy variables, fitting function.

Адрес редакции:

107076, Москва, Стромынский пер., 4

Телефон редакции журнала **(499) 269-5510**

E-mail: it@novtex.ru

Дизайнер *Т.Н. Погорелова*. Технический редактор *О. А. Ефремова*.
Корректор *Т. В. Пчелкина*

Сдано в набор 07.08.2009. Подписано в печать 17.09.2009. Формат 60×88 1/8. Бумага офсетная. Печать офсетная.
Усл. печ. л. 9,8. Уч.-изд. л. 11,62. Заказ 822. Цена договорная.

Журнал зарегистрирован в Министерстве Российской Федерации по делам печати,
телерадиовещания и средств массовых коммуникаций.
Свидетельство о регистрации ПИ № 77-15565 от 02 июня 2003 г.

Отпечатано в ООО "Подольская Периодика"
142110, Московская обл., г. Подольск, ул. Кирова, 15