

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ УКРАИНЫ
Одесский национальный политехнический университет



**ИНФОРМАЦИОННЫЕ
УПРАВЛЯЮЩИЕ СИСТЕМЫ
И ТЕХНОЛОГИИ.
ПРОБЛЕМЫ И РЕШЕНИЯ**

МОНОГРАФИЯ

ОДЕССА
«Экология»
2019

УДК 681.518.54
И74

Авторский коллектив:

Васянин В., Востров Г., Вычужанин В., Бобок І., Грикун Д., Дрозд О., Сгольников О., Сгошина Г., Журиленко Б., Казимиренко Ю., Калбазов Д., Козлов А., Кірьязов Ю., Кобозева А., Коновалов С., Кораблев В., Кораблев Н., Лисецкий Ю., Мазурок Т., Михайлов С., Николаева Н., Опята Р., Отрадская Т., Петров И., Рудниченко Н., Ситніков В., Соловьев Д., Теплечук А., Трофимчук А., Ухіна Г., Ушакова Л., Фомичев А., Хорошко В., Хриненко А., Чуприна А., Чупринка В., Чупринка Н., Шевцов Ю., Шибав Д., Шибав Н., Шпінарева І., Юхименко Б.

Под научной редакцией
проф. В. Вычужанина

Рецензенты:

Александр САВЧЕНКО, доктор физико-математических наук, профессор, Херсонский филиал Национального университета кораблестроения имени адм. С. О. Макарова;

Александр СТРИЖАК, доктор технических наук, профессор, Институт телекоммуникаций и глобального информационного пространства Национальной академии наук Украины;

Игорь ЧМЫРЬ, доктор технических наук, профессор, профессор кафедры «Информатика» Одесского государственного экологического университета

Рекомендовано в печать ученым советом
Одесского национального политехнического университета.
Протокол № 3 от 29.10.2019 г.

Информационные управляющие системы и технологии.
И74 Проблемы и решения : монография / авт. кол.: В. Васянин, Г. Востров, В. Вычужанин [и др.] ; под науч. ред. проф. В. Вычужанина — Одесса : Экология, 2019. — 244 с. : ил., табл.
ISBN 978-617-7046-88-1

В монографии отражены результаты научных исследований в области информационных интеллектуальных технологий для автоматизированных систем обработки данных и управления, моделирования и оптимизации в информационных управляющих системах, управления защитой информационных систем.

Материалы монографии будут полезными для аспирантов, магистрантов, преподавателей высших учебных заведений, специализирующихся в области IT-технологий.

УДК 681.518.54

ISBN 978-617-7046-88-1

© Одесский национальный
политехнический университет, 2019

Информационные управляющие системы и технологии. Проблемы и решения

СОДЕРЖАНИЕ

Предисловие.....	5
<i>Лисецкий Ю., Калбазов Д.</i> Облачные технологии как основа виртуальных управляемых сервисов.....	6
<i>Вычужанин В., Козлов А.</i> Оптимизация параметров настройки цифрового ПИД – регулятора с помощью алгоритма дифференциальной эволюции.....	19
<i>Рудниченко Н., Вычужанин В., Шibaева Н., Шibaев Д., Отрадская Т., Петров И.</i> Применение методов машинного обучения для автоматизации процессов классификации массивов текстовых данных большого объема.....	31
<i>Хорошко В., Кобозева А., Бобок I.</i> Удосконалення методу виявлення блокової обробки цифрового зображення.....	47
<i>Журиленко Б., Николаева Н.</i> Математическое представление вероятностной надежности защиты информации в зависимости от направления взлома.....	60
<i>Шніарева I.</i> Інтелектуальна система виявлення атак в локальних мережах.....	74
<i>Михайлов С., Шевцов Ю.</i> Структура защиты информации в международной гидрографической организации.....	86
<i>Кораблев Н., Фомичев А., Соловьев Д., Чуприна А.</i> Гибридные модели принятия решений с использованием иммунного подхода.....	100
<i>Vostrov G., Opiata R.</i> Computer modeling of the processes of development of information technology in dynamic processes of the formation of classes of the generalized Artin's hypothesis.....	117

Информационные управляющие системы и технологии. Проблемы и решения

- Ухіна Г., Теплечук А., Кірьязов Ю., Ситніков В.*
Управління смуговим частотно-залежним компонентом інформаційно-управляючою системою для усунення детонації двигуна внутрішнього згорання.....137
- Казимиренко Ю., Дрозд О., Єгольніков О.*
Інформаційна підтримка технічного обслуговування систем залізнично-водного сполучення.....151
- Коновалов С., Єгошина Г.*
Діагностика ризиків аварійних ситуацій за допомогою гібридних експертних систем.....162
- Трофимчук А., Васянин В., Ушакова Л.*
Алгоритмы полного перебора в схемах ветвления решения комбинаторных задач с псевдодулевыми функциями.....175
- Кораблев В., Мазурок Т.*
Інформаційна технологія побудови поведінкових моделей мультиагентних робототехнічних систем.....187
- Чупринка В., Грикун Д., Чупринка Н.*
Автоматизована підготовка раціональних схем розкрою матеріалів зі складною конфігурацією зовнішніх контурів на плоскі геометричні об'єкти.....198
- Юхименко Б.*
Некоторые информационные проблемы комбинаторной линейной оптимизации.....211
- Vostrov G., Khrinenko A.*
Dimensionality Reduction and Extremal Parameter Grouping for Weighted Graph Problems.....226

ПРЕДИСЛОВИЕ

В коллективной монографии изложены результаты научных исследований в области информационных интеллектуальных технологий, моделирования в информационных управляющих системах, управления защитой информационных систем, совершенствования информационно-ресурсного обеспечения науки.

Монография составлена в виде научных статей-разделов, соответствующих решаемым проблемам в области информационных управляющих систем и технологий, в которых рассмотрены основные вопросы в разделах: совершенствования информационно-ресурсного обеспечения науки, техники и социальной сферы; способы и методы защиты информации; информационные интеллектуальные технологии для автоматизированных систем обработки данных и управления; математическое моделирование и оптимизация в информационных управляющих системах; информационные технологий управления проектами. Результаты выполнения работ по перечисленным разделам позволяют решить некоторые проблемы информационных управляющих систем и технологий.

В коллективной монографии авторами уделено внимание решению проблем в области информационных управляющих систем и технологий применением методов машинного обучения для автоматизации процессов классификации текстовых данных большого объема, интеллектуальной системы выявления атак в локальных сетях.

Значительное внимание уделено компьютерному моделированию процессов развития информационных технологий в динамических процессах, математическому представлению вероятностной надежности защиты информации в зависимости от направления взлома, облачным технологиям, как основе виртуальных управляемых сервисов и другим актуальным темам. Благодаря этому читатели получают полезный объем знаний, необходимых для лучшего понимания проблем и их решений в области информационных управляющих систем и технологий.

Материалы монографии будут полезными для аспирантов, магистрантов, преподавателей высших учебных заведений, специализирующихся в области IT-технологий.

В коллективной монографии приняли участие ученые в составе 11 докторов наук, 15 кандидатов наук, 15 соискателей.

ОБЛАЧНЫЕ ТЕХНОЛОГИИ КАК ОСНОВА ВИРТУАЛЬНЫХ УПРАВЛЯЕМЫХ СЕРВИСОВ

Лисецкий Ю., Калбазов Д.

Рассмотрены облачные технологии, модели их развертывания и обслуживания. Представлена облачная коммуникационная платформа по модели SaaS. Показано, что облачные технологии являются основой для предоставления виртуальных управляемых сервисов. Проведен анализ отечественного рынка телекоммуникационных услуг для определения рыночных сегментов и основных видов услуг, которые могут быть предоставлены как vMS. Предложены алгоритмы развертывания vMS конечным потребителем и сервис-провайдером.

Ключевые слова. Облачные технологии, облачные модели, коммуникационная платформа, инфраструктура, виртуальная среда, виртуальные управляемые сервисы, vMS, сервис-провайдер, телеком-провайдер.

ВВЕДЕНИЕ

В последние годы во всем мире быстрыми темпами развиваются облачные технологии или, как еще их называют, – облачные вычисления [1]. Это модель предоставления повсеместного и удобного сетевого доступа к общему пулу конфигурируемых вычислительных ресурсов (сетям передачи данных, серверам, системам хранения данных, приложениями сервисам), которые могут быть оперативно предоставлены и освобождены с минимальными усилиями по управлению и необходимости взаимодействия с провайдером.

Существует четыре модели развертывания облачных технологий [2]: частное облако; публичное облако; общественное облако и гибридное облако. Также есть четыре основных модели обслуживания облачных вычислений [3]:

- программное обеспечение как услуга – SaaS (Software as a Service);
- платформа как услуга – PaaS (Platform as a Service);
- инфраструктура как услуга – IaaS (Infrastructure as a Service);
- рабочий стол как услуга – DaaS (Desktop-as-a-Service).

Основное преимущество данной технологии заключается в том, что, не имея собственной инфраструктуры, операционной системы и программного обеспечения, пользователь получает возможность доступа к своим данным и работы с ними. Также к преимуществам относятся: доступность, мобильность, экономичность, гибкость, высокая технологичность и надежность. Именно благодаря этому увеличивается количество пользователей коммуникационными сервисами облачной телефонии, видео-конференцсвязи, унифицированных коммуникаций и совместной работы над проектами, предоставляемыми по модели SaaS, при которой сервис-провайдер предлагает подписчикам готовое прикладное программное обеспечение (ПО) как услугу [4].

Учитывая востребованность таких сервисов и достаточно большую конкуренцию на рынке коммуникационных услуг, компании-производители предлагают различные решения, как например, объединение всех коммуникационных служб в одном приложении.

ОБЛАЧНАЯ КОММУНИКАЦИОННАЯ ПЛАТФОРМА ПО МОДЕЛИ SAAS

Облачные и мобильные технологии дают возможность доступа к таким приложениям с любого рабочего места (веб-браузер, ноутбук, мобильный телефон и др.). В этих приложениях можно вести переписку с группой людей, организовывать аудио- и видеоконференции, делать аудио- и видеозвонки пользователям приложений, а также на номера ТфОП и IP-телефонии, интегрировать необходимые для работы инструменты. Одно из таких приложений Spark разработано компанией Cisco – лидером среди производителей on-premise систем коммуникационных сервисов, и провайдера одной из самых популярных облачных служб WebEx [5].

Cisco Spark Service – основная служба облачной платформы, представляющая собой набор инструментов для совместной работы. В сервисе для общения выделено три базовых инструмента: Message для обмена сообщениями и файлами; Meeting для проведения голосовых и видеоконференций; Call для телефонных звонков. Реализация и настройка этих инструментов осуществляется через приложение Cisco Spark App и веб-портал Cloud Collaboration Management Portal (CCMP). С помощью приложения можно обмениваться сообщениями и делиться файлами с возможностью предварительного просмотра как с командой, так и друг с другом, создавать или присоединяться к голосовым и видеоконференциям с демонстрацией на экран, совершать голосовые и видео звонки по SIP-протоколу. К комнатам через веб-портал Cisco Spark Depot возможно подключать ботов и интегрировать различные облачные приложения: ZenDesk, GitHub, Zapier и др.

Для работы с Cisco Spark используется клиент на десктопе под управлением Mac или Windows, веб-клиент или мобильный клиент для Android или iOS. Автоматизация приложения возможна с помощью службы Spark for Developers. С ее помощью Cisco Spark Service может использоваться как REST API приложение для интеграции разработчиками в службу своих решений. Настройка сервиса производится с помощью CCMP, который предоставляет удобный интерфейс администрирования пользователями, службами и лицензиями, а также отчеты по использованию служб и мониторингу сервисов. С его помощью к сервису можно «привязать» DNS компании, настроить SIP URI и интегрировать свой SingleSignOn-сервис. На портале возможно приобрести контракт на предоставление услуг ТФОП у Preferred Media Partner (PMP) Cisco, после чего платформа становится облачной АТС, с возможностью регистрации телефонов моделей 7800 и 8800 или подключения системы для проведения конференций SX10. Также можно использовать Spark Hybrid Services для интеграции облачного сервиса с on-premise сервисами заказчика. Сервис включает в себя такие инструменты, как Hybrid Call, Hybrid Calendar, Hybrid Directory. При подключении Hybrid Directory к Directory Connector на стороне клиента возможны синхронизация списка пользователей облачного сервиса с списком пользователей компании в службе каталогов Active Directory и использование SSO-службы компании для работы с Cisco Spark App и CCMP. Подключение Hybrid Call, Hybrid Calendar к клиентскому Cisco Expressway позволяет использовать календарь MS Exchange с приложением Cisco Spark для создания запланированных конференций и как программный клиент для Cisco Unified Communication систем.

Надо отметить, что не только компания Cisco активно развивает свои облачные коммуникационные платформы по модели SaaS, но и ее конкуренты в этой области – Microsoft, Facebook, Google, так как с развитием облачных и мобильных технологий появилась возможность изменить традиционные механизмы общения между людьми, объединения все коммуникационные сервисы в одной облачной службе, а также на их базе развивать виртуальные управляемые сервисы (Virtual Managed Services, vMS) [6].

ВИРТУАЛЬНЫЕ УПРАВЛЯЕМЫЕ СЕРВИСЫ

В современном мире информационных технологий все больше внимания уделяют новым концепциям и одна из них – виртуальные управляемые сервисы. Называем их концепцией, так как это понятие не является новой технологией, а представляет собой, скорее, новый подход к построению ИТ-инфраструктуры и предоставлению сервисов.

Первое наше практическое знакомство с подходом к построению ИТ-инфраструктуры с точки зрения предоставления сервисов состоялось

в 2012 году во время проектирования ИТ-инфраструктуры для проведения футбольного чемпионата «Евро-2012» в Украине.

Подготовка ИТ-инфраструктуры для проведения мероприятий чемпионата была поручена национальному оператору связи «Укртелеком». Для реализации этой задачи был разработан и предоставлен на рассмотрение комиссии УЕФА проект, включавший в себя спецификацию необходимого оборудования и ПО для построения ИТ-инфраструктуры, соответствующей требованиям ТЗ заказчика. Но представители УЕФА не стали рассматривать этот проект и спецификации оборудования и ПО, так как их не интересовало построение ИТ-инфраструктуры, а интересовало получение необходимых сервисов, как услуги. После этого была разработана сервисная модель ИТ-инфраструктуры для «Евро-2012», которая включала в себя перечень необходимых сервисов: «Public Internet», «VIP-Internet», «Corporate IP-telephony», «Virtual Privet Network», «Virtual Privet Network with Internet» и пр. (рис. 1). Определены характеристики каждого из сервисов – резервирование сервиса, надежность, скорость передачи данных, скорость реакции на проблемы и пр.

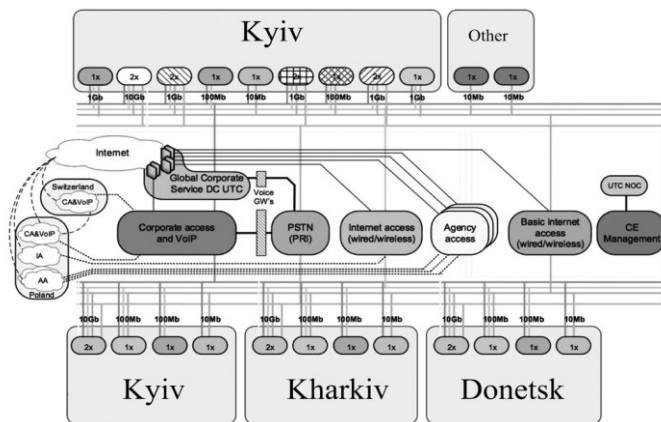


Рис. 1 – Сервисная модель ИТ-инфраструктуры для «Евро-2012»

Далее был согласован список объектов, на которых необходимо предоставлять сервисы и перечень самих сервисов, которые необходимы на каждом из объектов. Таким образом, конечный потребитель платил не за ИТ-инфраструктуру, а за необходимый ему сервис с определенным уровнем обслуживания. Это был первый практический опыт работы по модели «управляемые сервисы».

Надо отметить, что подходы к представлению сервисов существовали на ИТ-рынке и до появления технологий виртуализации [7–10], но благодаря им получили возможность эволюционировать в концепцию vMS. Соответственно появилось понятие Service Provider, которое постепенно вытесняет другое понятие – Telecom/Internet Provider. Таким образом, концепция vMS органично выросла из технологий SDN (Self Defined Network), NFV (Network Function Virtualization) и технологий оркестрации виртуальной инфраструктуры [11], под которой понимается технология автоматизации разворачивания и настройки виртуальной среды, в соответствии с конкретными требованиями. Технологии SDN позволяют консолидировать всю логику принятия решений по работе сетевого оборудования в одной точке, что, в свою очередь, позволяет получить более полный контроль над инфраструктурой, а NFV дают возможность создавать виртуальные сетевые элементы, для выполнения узконаправленных задач (фильтрация, мониторинг и изучения пользовательского трафика, антивирус). Ранее одно физическое устройство могло выполнять несколько функций, сейчас с помощью NFV появилась возможность разграничивать сетевые функции на необходимое количество виртуальных устройств, что позволяет упростить схему подключения и устранить аппаратные сбои.



Рис. 2 – Совмещение технологий оркестрации, SND, NFV и традиционной ИТ-инфраструктуры в vMS

Именно на стыке этих технологий находится vMS – автоматизация процесса разворачивания и управления ИТ-сервисами (рис. 2).

С точки зрения конечного пользователя, ИТ-сервис – это любая функция с определенным набором характеристик, реализация которой

обеспечивается ИТ-инфраструктурой. Например, возможность печатать на сетевом черно-белом принтере – это ИТ-сервис (назовем его «Basic Printer Access»). Если пользователь, является руководителем определенного уровня, ему должен быть предоставлен сервис «VIP Printer Access» – доступ к цветному принтеру. Если пользователь – рядовой сотрудник, ему должны быть предоставлены сервисы: «Basic Printer Access» и «Plotter Printer Access» – печать на обычном принтере и на плоттере. Приведенный пример, хотя и весьма прост, тем не менее иллюстрируют сервисный подход к построению ИТ-инфраструктуры. Другие примеры сервисов изображены на рис. 3.



Рис. 3 – Примеры ИТ-сервисов

Таких примеров можно привести весьма много, учитывая, что каждый тип сервиса может иметь свои подтипы, в зависимости от характеристик предоставляемого сервиса, как, например, скорость передачи данных, резервируемость, доступность, период действия, качество передачи данных (применимо для видео), приоритетность данных и пр.

Вопрос востребованности vMS надо рассматривать как с точки зрения сервиса провайдера (продажа сервисов в концепции vMS), так и с точки зрения владельца бизнеса (использование vMS для обеспечения нужд бизнеса).

Подход к проектированию ИТ-инфраструктуры с точки зрения предоставления сервисов позволяет изначально проектировать ИТ-инфраструктуру исходя из требований бизнеса, что, в свою очередь, позволяет владельцам бизнеса лучше понять свои потребности, формализовать и систематизировать их.

В каких же случаях надо внедрять ИТ-инфраструктуру, позволяющую предоставлять виртуальные управляемые сервисы (vMS)? С точки зрения владельца ИТ-инфраструктуры – внедрение vMS

оправдано для организаций корпоративного уровня, имеющих территориально-распределенную филиальную сеть. В этом случае гораздо выгоднее вместо большого количества разрозненных программно- аппаратных систем создать виртуальные сервисы, которые можно точно также использовать, но уже в единой виртуальной среде. Разворачивание любого требуемого сервиса будет занимать небольшое время, а расходы на его предоставление в каждой из точек значительно уменьшаться.

Также, внедрение vMS может быть привлекательным для организаций у которых динамично изменяется филиальная сеть. В таких случаях для предоставления корпоративных сервисов в новом отделении, нужна только организация надежного канала передачи данных. Все сервисы, которые необходимы этому отделению, разворачиваются автоматически, в виртуальной ИТ-инфраструктуре и в кратчайшие сроки, так как нет необходимости приобретать оборудование, доставлять в отделение, обслуживать, хранить запасные модули и т.д.

Для ответа на вопрос: «Зачем vMS нужны провайдерам?», необходимо провести небольшой анализ рынка телекоммуникационных услуг.

Сегодня наибольшее распространение получили три основных вида услуг: сотовая связь, интернет, VPN доступ (рис. 4).



Рис. 4 – Услуги современных телеком-операторов

Если рассмотреть более детально каждый из видов услуг, то окажется, что количество абонентов сотовой связи в Украине – 60 млн., при населении в 42 млн. Поэтому рынок предоставления услуг доступа к сети интернет настолько конкурентный и насыщенный, что повышение стоимости услуг даже на 1 USD может привести к потере клиента. Рынок же услуг VPN достаточно небольшой и специфический, поэтому можно не выделять его в отдельную категорию.

С учетом вышеизложенного, очевидно, что отечественным сервис-провайдерам необходимо выбрать одну из следующих стратегий

дальнейшего развития: развитие услуг – предоставление новых сервисов на существующем рынке; развитие рынка – поиск новых рынков с существующими продуктами; диверсификация – поиск новых рынков с новым продуктом.

Реализацию таких стратегий мы можем наблюдать на примере тройки лидеров отечественного рынка телекоммуникаций:

- Киевстар – внедрение интернет-магазина, сотрудничество с Viasat для предоставления услуг телевидения и видео по запросу. Такой подход можно рассматривать как пример стратегии диверсификации.

- Vodafone – сотрудничество с Word of Tanks, возможность покупки авиабилетов, сотрудничество с сетью АЗС WOG – и это пример стратегии развития услуг;

- Lifecell – открытие интернет магазина, предоставление услуг «СИНЕМА», «ИМПЕРИЯ ИГОР» и др. – также является примером стратегии развития услуг.

Если проанализировать целевую аудиторию нынешних телеком-провайдеров, ее можно разделить на следующие сегменты:

- пользователи мобильных устройств (смартфоны, планшетные компьютеры, ноутбуки, часы и пр.) – передача голоса, видео, интернет, мобильные игры, музыка и т.д.;

- домохозяйства – домашний интернет и дополнительные сопутствующие сервисы;

- малый и средний бизнес – доступ к сети интернет, каналы связи. Реже аренда оборудования или хостинг.

- большой бизнес – доступ к сети интернет, каналы связи.

В различных источниках оценки емкости и относительные доли рынка несколько отличаются друг от друга, но общая картина примерно такова, как на рис.5.



Рис.5 – Сегменты рынка телеком-провайдеров

Если рассмотреть каждый из этих сегментов, то:

1. Сегмент мобильных пользователей перенасыщен. Новые сервисы здесь нужны скорее для удержания существующих абонентов, чем для привлечения новых.

2. Сегмент домохозяйств весьма конкурентный. Локальные игроки («домашние» интернет провайдеры) показывают себя гораздо лучше в этом сегменте, так как умеют работать в своих нишах быстрее и качественней, и в их бюджете нет расходов на поддержание большого управленческого аппарата, маркетинговых программ и пр.

3. Малый и средний бизнес – объединены в одном сегменте рынка, так как имеют общие характеристики. Представителям этого бизнеса уже нужна минимальная инфраструктура, но у них еще нет понимания, как они будут ее обслуживать и развивать.

4. Большой бизнес – как правило, характеризуется наличием своей собственной инфраструктуры, своим штатом профессионалов и пониманием требований к инфраструктуре. Их основная потребность заключается в передаче большого объема данных.

Из анализа предоставляемых услуг телеком-провайдерами можно сделать вывод, что в их портфеле практически отсутствуют услуги для решения инфраструктурных задачи в сегменте «малый и средний бизнес», представители которого сталкиваются со следующими типовыми проблемами:

- какое оборудование установить?
- где его взять и как настроить?
- где хранить данные и как защитить их от несанкционированного доступа?
- как совместить несколько отделений в одной инфраструктуре?
- кто будет эксплуатировать, обслуживать и оказывать техническую поддержку?
- у кого получить необходимые консультации?

У мобильных пользователей и домохозяйствах такие проблемы отсутствуют, так как у них нет собственной ИТ-инфраструктуры, а большой бизнес вынужден решать эти проблемы собственными силами. Что касается малого и среднего бизнеса, то здесь подход к решению аналогичных проблем, как правило, ситуативный и спонтанный. Именно в этих отраслевых сегментах отечественные телеком-провайдеры могут себя показать в новой роли – провайдера услуг или Service Provider, так как концепция vMS сделала возможным предоставление услуг в виде сервисов, а автоматизация всех процессов делает это экономически выгодным для провайдера и доступным для конечного потребителя. Особенно с учетом того, что даже небольшая организация нуждается в

минимально необходимом наборе сервисов, как например: доступ к сети интернет; обмен данными между пользователями; корпоративная почта; сетевой экран (Firewall).

По мере роста потребностей организации к этому списку могут еще добавиться: доступ к Wi-Fi; Anti-SPAM; гостевой доступ к сети; Веб-сервер; хранилище данных (резервируемое или не резервируемое); корпоративная телефония и/или видеоконференции; база данных.

Также vMS позволяет автоматизировать процесс заказа и разворачивания сервиса для предоставления услуги конечному потребителю. При этом необходимо реализовать алгоритм, включающий следующие шаги (рис.6).

Шаг 1 – на портале самообслуживания заказчик заполняет форму, в которой указывает место подключения, количество пользователей и набор сервисов, которые он хочет заказать и выполняет оплату.

Шаг 2 – в течении определенного времени (до нескольких часов) в виртуальной инфраструктуре сервис-провайдера автоматически создаются и настраиваются необходимые виртуальные машины, после чего сервисы готовы к использованию (например, почтовый сервер и сетевой экран).

Шаг 3 – на объект заказчика выполняется доставка оборудования для подключения рабочих станций к сети (это может быть коммутатор или точка доступа).

Шаг 4 – оборудование подключается к сети интернет, после чего можно использовать заказанные сервисы.



Рис. 6 – Алгоритм развертывания сервисов конечным потребителем

С точки зрения сервис провайдера этот процесс представляет собой алгоритм, включающий следующие шаги (рис. 7).

Шаг 1 – в виртуальной инфраструктуре провайдера создается «область» под выделенного заказчика.

Шаг 2 – в этой области создаются виртуальные машины, необходимые для предоставления заказанного набора сервисов.

Шаг 3 – в автоматическом режиме создаются связи между виртуальными машинами, права доступа, конфигурационные параметры.

Шаг 4 – настраивается канал связи в точке включения оборудования пользователя.

Шаг 5 – подключается оборудование на объекте заказчика и устройства получают доступ к инфраструктурным элементам, созданным в виртуальной среде провайдера.

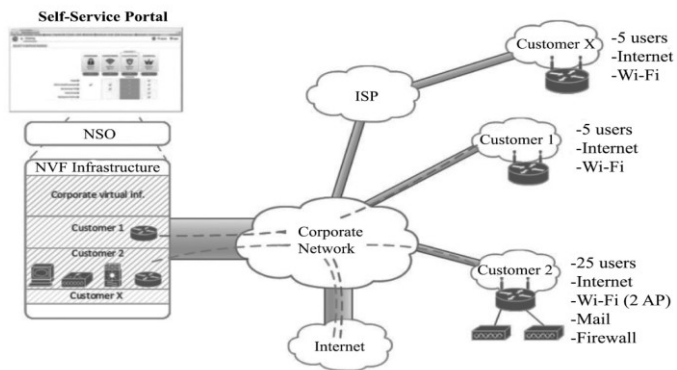


Рис. 7 – Алгоритм развертывания сервисов сервис-провайдером

Кроме этого заказчики могут пользоваться сервисами, не имея физического подключения от этого же провайдера. Иными словами – имея любой доступ к сети интернет, можно получить доступ к заказанным сервисам.

Преимущества данной модели предоставления услуг для сервис провайдера следующие:

- возможность повторно использовать ресурсы виртуальной инфраструктуры;
- минимизация количества использования аппаратного обеспечения, что ведет к уменьшению расходов на техническое обслуживание и хранение запчастей;
- предоставление услуг через любой канал связи;
- автоматизация процессов с помощью vMS и портал самообслуживания;

- расширение рынка и спектра услуг.

При этом конечный потребитель получает следующие преимущества:

- фокусировка на бизнесе, а не на инфраструктуре;
- сохранность информации и резервирование;
- отсутствие расходов на оборудование и его обслуживание;
- скорость разворачивания нового отделения определяется

только временем доставки оборудования доступа;

- открытие и/или закрытие отделения не влечет за собой дополнительных затрат на оборудование (покупку, модернизацию, монтаж, демонтаж и пр.)

- эксплуатация и техническая поддержка осуществляется соответствующей службой сервис-провайдера.

Необходимо отметить, что на отечественном рынке уже присутствуют глобальные провайдеры, которые предоставляют сервисы по такой же модели. Примером могут служить компании Google, Microsoft, Amazon и др. Но конкуренция с ними не столь сложна, так как они не имеют представительств в каждом городе Украины и локальной команды технической поддержки. Также большим недостатком этих сервис-провайдеров является практически полное отсутствие ответственности за нарушение условий предоставления сервисов.

ВЫВОДЫ

Таким образом, использование vMS дает возможность сформировать принципиально новые наборы услуг, которые могут быть востребованы на рынке Украины и тем самым придать достаточно мощный импульс развитию ИТ-отрасли в целом. Учитывая, что существующие возможности отечественных телеком-провайдеров позволяют внедрить сервисную модель предоставления услуг без существенных дополнительных затрат, это дает им хорошие шансы в перспективе стать полноценными и конкурентными сервис-провайдерами.

СПИСОК ЛИТЕРАТУРЫ

[1] Облачные технологии. URL: <https://kontur.ru/articles/225> (дата обращения: 15.07.2019).

[2] Модели развёртывания. URL: https://ru.wikipedia.org/wiki/%D0%9E%D0%B1%D0%BB%D0%B0%D1%87%D0%BD%D1%8B%D0%B5_%D0%B2%D1%8B%D1%87%D0%B8%D1%81%D0%BB%D0%B5%D0%BD%D0%B8%D1%8F (дата обращения: 15.07.2019).

- [3] Что такое IaaS, PaaS, SaaS и DaaS? URL: <https://3data.ru/services/cloud/iaas-paas-saas-daas> (дата обращения: 05.08.2019).
- [4] Что такое модель SaaS, ее преимущества и примеры. URL: <https://www.kasper.by/blog/model-saas/> (дата обращения: 05.08.2019).
- [5] Cisco Spark: весь функционал для совещаний в одном решении. URL: https://www.cisco.com/c/ru_ru/about/press/press-releases/2017/01-25b.html (дата обращения: 05.08.2019).
- [6] Управляемые сервисы. URL: https://ko.com.ua/upravlyaemye_servisy_117807 (дата обращения: 18.08.2019).
- [7] Анализ современных технологий виртуализации. URL: <https://habr.com/company/southbridge/blog/212985/> (дата обращения: 18.08.2019).
- [8] Тормасов А.Г. Виртуализационные технологии и их возможности// ВУТЕ, 2005. – № 3. – С. 35 –45.
- [9] Лисецкий Ю.М. Виртуализация: динамика развития и перспективы. *Інформаційні управляючі системи та технології: 3-я міжнар. наук.-практ. конф. (Одеса, 23–25 вересня 2014 р.)*. Одеса, 2014. – С. 271 – 273.
- [10] VMware. История виртуализации. URL: <http://www.vmware.com/ru/virtualization/virtualizationbasics/history.html> (дата обращения: 18.08.2019).
- [11] Лисецкий Ю.М., Татаренко М.А. Технологии виртуализации от Oracle. *Математичні машини і системи*, 2018. – № 4. – С. 34 – 43.

CLOUD TECHNOLOGIES AS A BASIS OF VIRTUAL MANAGED SERVICES

Lisetskiy Yu., Kalbazov D.

Cloud technologies, models of their deployment and maintenance are considered. Introduced SaaS cloud communications platform. It is shown that cloud technologies are the basis for the provision of virtual managed services. The analysis of the domestic telecommunication services market is carried out to identify market segments and the main types of services that can be provided as vMS. Algorithms for deploying vMS by the end consumer and service provider are proposed.

Keywords. Cloud technologies, cloud models, communication platform, infrastructure, virtual environment, virtual managed services, vMS, service provider, telecom provider.

ОПТИМИЗАЦИЯ ПАРАМЕТРОВ НАСТРОЙКИ ЦИФРОВОГО ПИД – РЕГУЛЯТОРА С ПОМОЩЬЮ АЛГОРИТМА ДИФФЕРЕНЦИАЛЬНОЙ ЭВОЛЮЦИИ

Вычужанин В., Козлов А.

В работе приведены результаты синтеза цифрового регулятора для систем автоматического управления теплообменными аппаратами центральных систем кондиционирования воздуха. Разработанный регулятор призван обеспечить заданное качество регулирования, изменения параметров настройки регуляторов с учетом условий эксплуатации объекта регулирования. Реализация требований к регулятору осуществлена в цифровом типом ПИД – регуляторе с оптимизацией его параметров настройки с помощью алгоритма дифференциальной эволюции. ПИД – регулятор реализован в ПЛИС. Исследования цифрового ПИД - регулятора с оптимизацией его параметров позволили установить, что регулятор удовлетворяет требуемому качеству регулирования. Обладает возможностью изменения параметров настройки с учетом условий эксплуатации объекта регулирования.

Ключевые слова. Цифровой ПИД - регулятор, дифференциальная эволюция, программируемая логическая интегральная схема, теплообменный аппарат.

ВВЕДЕНИЕ

Проблема качественного управления объектами, функционирующих при переменных существенных возмущающих воздействиях, является одной из важных, которую необходимо учитывать при разработке соответствующих систем автоматического управления (САУ) [1-4].

Основные требования, предъявляемые к разрабатываемым регуляторам подобных САУ, заключаются в обеспечении заданного качества регулирования (малые время регулирования, допустимая величина перегулирования), в возможности изменения параметров настройки регулятора с учетом условий эксплуатации объекта регулирования, а также в относительной простоте схемного решения регуляторов. Неправильная настройка параметров регуляторов может привести к циклическому и медленному их восстановлению, плохой устойчивости и потери управляемости объектом регулирования.

К известным методам настройки параметров регуляторов относятся методы Циглера-Николса, Чина-Хронеса-Ресвика, Козна-Куна и другие методы [5,6]. Большинство из этих методов применимы как к непрерывным, так и дискретным системам регулирования. Предложенные Циглером-Николсом, Козном-Куном экспериментальные

способы настройки регуляторов основаны на методах проб и ошибок. Однако при использовании этих методов могут возникнуть проблемы при настройке регуляторов, например для нелинейных объектов высокого порядка, при наличии временных задержек, нелинейных процессах в системе регулирования.

Для преодоления названных проблем применяются различные методы для получения рациональных параметров настройки регуляторов, в том числе методы, основанные на применении эволюционных вычислений - генетического алгоритма (GA) и дифференциальной эволюции (DE) [7-9]. Недостатки в работе GA: ложная сходимость, потеря наилучшего найденного решения, отсутствует поддержка оптимального значения [10]. Поэтому в настоящее время для получения рациональных параметров настройки регуляторов находит все большее применение дифференциальная эволюция. Главная особенность DE заключается в использовании схемы генерации векторов пробных параметров. Обычно DE [11] реализация нуждается в трех параметрах: CR (определяющий взаимоисключающие операции кроссовера и мутации); F (коэффициент масштабирования разности двух особей); NP (размер популяции) для генерации эволюционного процесса для n-мерной задачи. Алгоритм DE прост, а его производительность сопоставима или даже превосходит GA [12,13]. Исходя из перечисленных преимуществ DE, такой алгоритм может быть выбран для оптимизации параметров настройки регуляторов управления объектами, функционирующими при переменных существенных возмущающих воздействиях.

Обеспечить реализацию настройки параметров цифровых регуляторов с учетом условий эксплуатации объекта регулирования возможно в программируемых логических интегральных схемах (ПЛИС). Развитие ПЛИС позволяет реализовать цифровые регуляторы с изменяемыми параметрами настройки с минимальными материальными затратами и сокращенным временем на проектирование. Производители предлагают программируемые простые, матричные и сложные логические устройства (SPLD, PAL, CPLD); программируемые пользователем базовые матричные микросхемы (FPGA) [14]. В настоящее время FPGA широко применяются в приложениях развивающегося аппаратного обеспечения с использованием интеллектуальных вычислительных методов при проектировании цифровых систем управления [15,16]. Цифровая САУ на FPGA обладает особенностями, отличающими ее от аналоговых систем: законы управления реализуются в форме алгоритмов, запрограммированных с помощью аппаратных или программных средств; обрабатываются квантованные (дискретные во времени) сигналы [17-19].

Семейства FPGA позволяют реализовать эволюционные алгоритмы, начиная от выделенной системы только на одном чипе и заканчивая кластером FPGA [20] для выполнения параллельных вычислений, которые могут быть полезны для различных приложений. Метод DE позволяет в реальном времени осуществить числовую оптимизацию параметров настройки регуляторов. DE подходит для точной минимизации числовых параметров регуляторов. При этом FPGA являются желательными устройствами для использования из-за их массивного параллелизма [21, 22].

В [23] исследован ПИД - регулятор при использовании его представления с фиксированной запятой в программе для FPGA. Производительность сравнивается с представлением с плавающей запятой с представлением с фиксированной запятой. Представление с фиксированной запятой оценивалось с использованием методов анализа длины слова. Было установлено, что представление с фиксированной точкой позволяет сэкономить значительные ресурсы в схеме на FPGA (потеря мощности и энергопотребление, сократить время разработки). Алгоритм DE по своей природе всегда использует кодирование с плавающей запятой и несколько поколений случайных чисел. Поэтому при реализации DE на FPGA необходимо внедрение в FPGA нескольких генераторов случайных чисел. Из [9] следует, что вопросы оптимизации параметров настройки ПИД - регуляторов, реализованных в FPGA, с использованием алгоритма DE недостаточно отражены с точки зрения оценки обеспечения стабильности работы регуляторов в режиме реального времени при минимальной среднеквадратичной и интегральной абсолютной ошибках.

Таким образом, поиск новых схемных решений и принципов функционирования САУ для обеспечения заданного качества управления объектами, функционирующих при переменных существенных возмущающих воздействиях предполагает дальнейшую разработку и исследования ПИД – регуляторов с оптимизацией параметров настройки с DE.

ЦЕЛЬ РАБОТЫ

Целью работы является обеспечение стабильности работы регулятора в режиме реального времени при минимальном времени регулирования и величине перегулирования при автоматизации объектов подверженных воздействию существенных возмущений.

РАЗРАБОТКА ЦИФРОВОГО ПИД - РЕГУЛЯТОРА В FPGA С ОПТИМИЗАЦИЕЙ ПАРАМЕТРОВ НАСТРОЙКИ НА ОСНОВЕ DE

В качестве объекта автоматизации выбран теплообменный аппарат (воздухонагреватель) центральной системы кондиционирования воздуха (СКВ), функционирующей при переменных существенных возмущающих воздействиях [23-25]. Для данного объекта управления из-за нелинейности характеристик агрегатов процесс настройки параметров управления регуляторов является трудоемкой задачей. Характеристики воздухонагревателя определяют процессы тепловой обработки воздуха в кондиционере, влияют на выбор управляющих воздействий и на качество переходных процессов в используемой системе регулирования. Тепловые процессы, протекающие в теплообменном аппарате, характеризуются распределенностью параметров и поэтому динамика их описывается в общем случае нелинейной системой уравнений [26]. В воздухонагревателях возмущающими воздействиями является температура и относительная влажность воздуха на его входе. Управляющими воздействиями могут быть расход пара или температура пара на входе, расход воздуха (если аппарат работает при переменном расходе). Регуляторы, используемые для поддержания заданных параметров на выходе теплообменных аппаратов центральной СКВ, реализуют П, ПИ и ПИД законы регулирования. Однако в условиях переменных существенных возмущающих воздействий регуляторы не обеспечивают требуемое качество регулирования. Параметры регуляторов нуждаются в настройке, что не всегда возможно выполнить в условиях эксплуатации СКВ.

При разработке регуляторов с перестраиваемыми параметрами настройки учитывалась математическая модель воздухонагревателя [27], выбранного в качестве объекта автоматизации. Оптимизация параметров настройки ПИД - регулятора осуществлена в соответствии с алгоритмом его функционирования и алгоритмом DE (рис.1). В соответствии с алгоритмом DE [28] используется набор решений (популяцию) и на каждом шаге преобразовывают их последовательным применением операций селекции, мутации скрещивания.

Целевая функция при оптимизации параметров настройки ПИД - регулятора в соответствии с алгоритмом DE

$$\min f(x) = f(x_1, x_2, \dots, x_n) \quad (1)$$

где x - n -мерный вектор, а f - действительная функция вещественных аргументов.

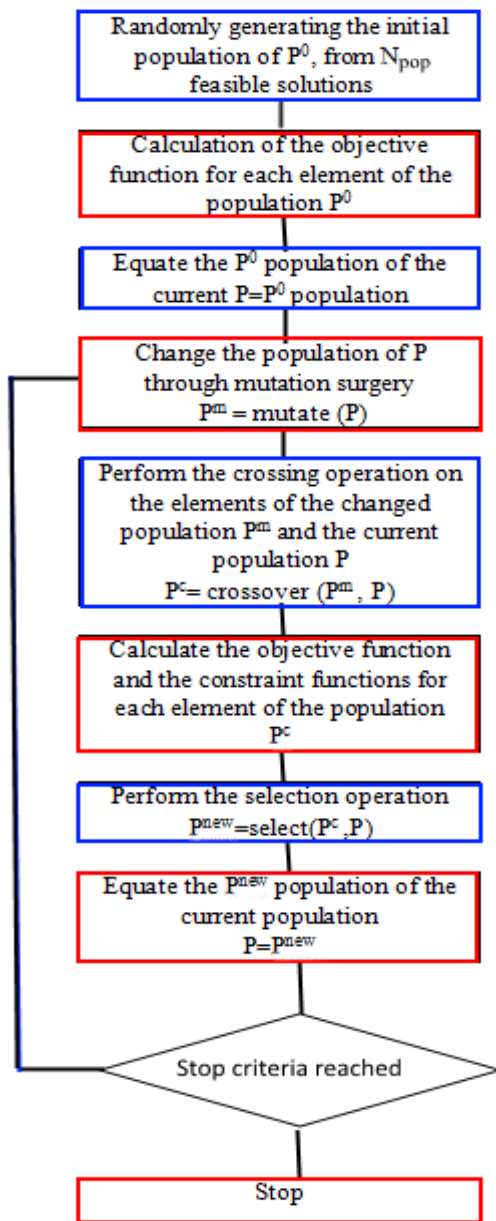


Рис.1 – Алгоритм дифференциальной эволюции

Для DE используется [29] требуемый размер популяции (NP) - $NP=10 \times D$, где D – это номер популяции, константа кроссовера (CR) - $CR \in [0,1]$, вес, примененный к случайному дифференциалу (коэффициент мутации) (F) - $F \in [0.5,1]$ и количество хромосом для соответствия контрольным параметрам регулятора.

Популяция инициализируется генерацией случайных особей от 1 до NP хромосом равномерно между 1 и D. Разностный вектор генерируется случайным образом из индивидуального выбора. Взвешенный разностный вектор формируется из разностного вектора, умноженного на весовой коэффициент мутации F. Цифровая структура ПИД - регулятора реализуется на уровне модулей. На константу коэффициента мутации влияет длина шага мутации. При уменьшении разностного вектора также уменьшается длина шага мутации в популяции. Вектор рождается из суммы параметра разностного вектора и индивидуального параметра. Целевой вектор пересекается с разностным вектором. Скрещивание выполняется по поколениям между параметром кроссовера CR и случайным числом для каждой хромосомы. В дальнейшем целевой и пробный вектора выбирались на основе значения целевой функции для размещения целевых векторных хромосом по поколениям. Выбор осуществляется в соответствии с целевой функцией. Из вектора мутации выбирается минимум одна хромосома. Если целевой вектор равен пробному вектору, то пробный вектор выбирается для продолжения в роде.

Беззнаковые оптимизированные параметры импортируются из DE - оптимизации в ПИД - регулятор. Пропорциональные, интегральные и производные составляющие рассчитываются отдельно в соответствующих программных модулях. Выходное значение ПИД - регулятора получается суммирование всех слагаемых. Оптимизация ПИД - регулятора реализуется уменьшением эталонной и измеряемой величины или количеством поколений. Параметры ПИД - регулятора подбираются сравнением значений пригодности разных поколений. В соответствии со значением пригодности для ПИД - регулятора выбираются наилучшие параметры управления.

Функциональная схема цифрового ПИД- регулятора с DE представлена на рис.2.

Цифровой ПИД - регулятор традиционно описывается

$$\frac{U(z)}{E(z)} = \frac{b_0 + b_1 \cdot z^{-1} + b_2 \cdot z^{-2}}{a_0 + a_1 \cdot z^{-1} + a_2 \cdot z^{-2}} \quad (2)$$

Рекуррентная процедура вычисления управления, соответствующего работе цифрового ПИД - регулятора с учетом [9] имеет вид

$$u[n] = k_1 \cdot u[n-1] + k_2 \cdot u[n-2] - k_3 \cdot e[n] - k_4 \cdot e[n-1] - k_5 \cdot e[n-2] \quad (3)$$

Схема ПИД - регулятора, интегрированного в FPGA, обычно состоит из блоков регистров (REG_1- REG_4), используемых для хранения текущих и предыдущих значений ошибок ($E(n)$, $e(n - 1)$, $e(n - 2)$, $u(n - 1)$ и $u(n - 2)$), сумматора (ADD) и умножителей. Блок-схема ПИД - регулятора в FPGA, полученная в соответствии с (3) и учетом [30] приведена на рис.3.

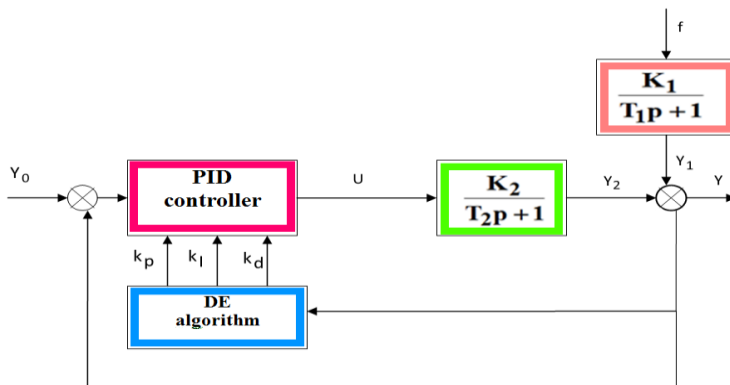


Рис.2 - Функциональная схема цифрового ПИД- регулятора воздушнонагревателя с DE

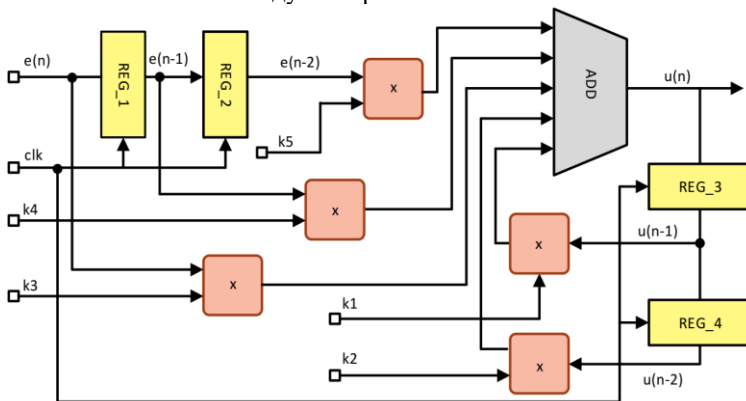


Рис.3 - Блок-схема ПИД - регулятора в FPGA

Для настройки параметров ПИД - регулятора использовалась отладочная плата DE2 (Altera) [30]. DE аппаратно реализован в модуле памяти PMem и модуле FXMem для хранения значений функций пригодности, генераторах случайных чисел и конечном автомате для управления последовательностью выполнения DE (рис. 4). Объем памяти определяется параметром размера популяции NP и размерностью D. Модуль FXMem реализован аналогично PMem, с той разницей, что размер FXMem определяется только параметром NP, так как индивидуум хранит только одно значение. Параметры DE при настройке ПИД – регулятора на FPGA: NP = 100; CR = 0,9; константа мутации F = 0,6; число поколений G = 50. D устанавливается на основе числа параметров, используемых в целевой функции.

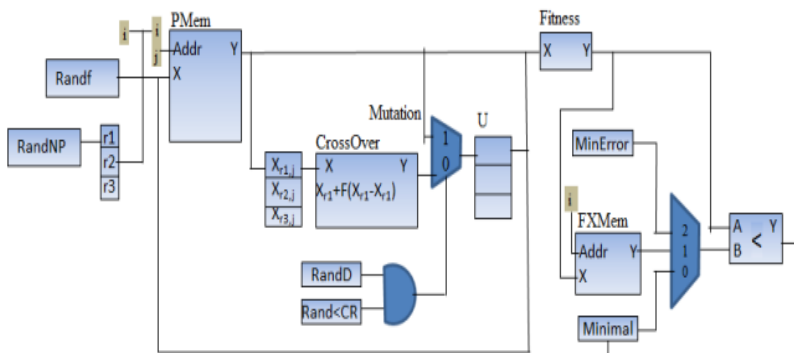


Рис.4 – Аппаратная реализация DE для конфигурации в FPGA

Ресурсы используемой программы FPGA для EP4CE115F29C7 (Cyclone IV E) определены в зависимости от количества логических элементов (4985), задействованных входов (71), регистров (1512), умножителей. В качестве языка для описания аппаратного обеспечения ПИД - регулятора используется язык описания аппаратных средств VHDL.

ИССЛЕДОВАНИЯ ЦИФРОВОГО ПИД - РЕГУЛЯТОРА В FPGA С ОПТИМИЗАЦИЕЙ ПАРАМЕТРОВ НАСТРОЙКИ НА ОСНОВЕ DE

Исследование качества регулирования ПИД – регулятора осуществлено моделированием в MATLAB. Результаты исследований для DE при использовании среднеквадратичной ошибки и интегральной абсолютной ошибки в качестве целевой функции при оптимизации параметров настройки ПИД – регулятора показаны на рис. 5, полученные при изменении ступенчатого возмущающего воздействия по теплосодержанию воздуха перед ОР.

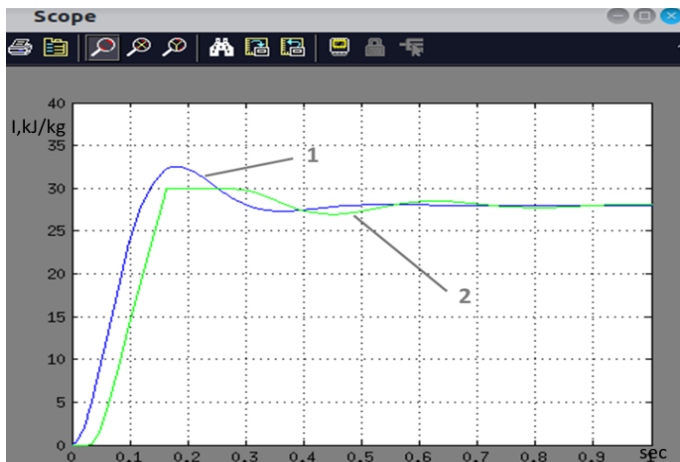


Рис.5 – Исследования ПИД – регулятора (1 – ПИД – регулятор; 2 – ПИД – регулятор с оптимизацией параметров настройки)

При моделировании использовались ранее полученные данные для воздухонагревателя центральной СКВ [31]: коэффициенты передачи $K_1 = 0,421 \pm 0,06$; $K_2 = 0,813 \pm 0,17$ кДж/(кг%); постоянные времени $T_1 = 1,4 \pm 0,3$; $T_2 = 6,1 \pm 1,2$ сек; период дискретизации $T_k = 1$, сек. Значения коэффициентов ПИД - регулятора k_p , k_i , k_d , оптимизированного по целевой функции среднеквадратичной ошибки соответственно составили 0,62 кДж/кг; 0,12 кДж сек/кг; 2,14 кДж/кг, по целевой функции интегральной абсолютной ошибки соответственно - 0,66 кДж/кг; 0,069 кДж сек/кг; 1,40 кДж/кг.

Из проведенных исследований: ПИД - регулятора время регулирования составило 0,48 сек, максимальное значение перерегулирования - 17,6%; ПИД - регулятора с оптимизацией параметров по целевой функции - среднеквадратичной ошибке равной 0,018 и по целевой функции - интегральной абсолютной ошибке равной 18,65 время регулирования составило 0,7 сек, максимальное значение перерегулирования - 11,2%.

Из результатов исследований, проведенных в условиях кратковременных ступенчатых возмущающих воздействий, установлено, что ПИД - регулятор с оптимизированными параметрами обладает на 6,4% меньшей величиной перерегулирования, в 1,45 раза меньшим временем установления управляемого параметра в сравнении с ПИД - регулятором с неоптимизированными параметрами, время установления управляемого параметра ПИД - регулятора с оптимизацией параметров

настройки на основе DE равно 0.48 сек. Оценка качества регулирования ПИД - регулятора с оптимизацией параметров по целевой функции позволила установить, что максимальное значение перерегулирования составило 11,2%.

ВЫВОДЫ

Таким образом, проведенные исследования цифрового ПИД – регулятора с оптимизированными параметрами настройки показали, что такой регулятор в сравнении с ПИД – регулятором с неоптимизированными параметрами настройки обладает лучшими динамическими характеристиками (меньшее перерегуливание). Позволили установить, что регулятор функционирует стабильно в режиме реального времени при минимальной величине перерегулирования и допустимом времени регулирования.

СПИСОК ЛИТЕРАТУРЫ

- [1] Richard, C. Dorf, & Robert, H. Bishop (2011). “Modern Control Systems”, Pearson, 1111 p. DOI: 10.1002/rnc.1054.
- [2] Katsuhiko, Ogata (2011). ”Modern Control Engineering, Fifth Edition”, Pearson, 905 p. DOI: 10.1093/benz /9780199773797.article.b00128780
- [3] Г. Олссон, Дж. Пьяни (2001). Цифровые системы автоматизации и управления. СПб, 2001. – 577 с.
- [4] Astrom, K. J., & Hagglund, T. (2006). “Advanced PID control”, *ISA The Instrumentation, System, and Automation Society*. – 460 p. DOI: 10.1016/s0967-0661(01)00062-4 .
- [5] Ziegler, J. G. & Nichols, N. B. (1942). “Optimum Settings for Automatic Controllers”. *Trans. ASME*. –64(8). – P. 759 –768.
- [6] Cohen, G. H. & Coon, G. A. (1953). «Theoretical Consideration of Retarded Control». *Trans. ASME*. 75. – P. 827 – 834.
- [7] Mahboobib, S. H., & Alasty, A. (2011). "Optimum Synthesis of Fuzzy Logic Controller for Trajectory Tracking by Differential Evolution", *Scientia Iranica*, Iran. – P. 261 – 267. DOI: 10.1016/j.scient.2011.03.021
- [8] Shing-Tai, Pan (2011). "Evolutionary Computation on Programmable Robust IIRFilter Pole-Placement Design", *Instrumentation and Measurement*. – Vol. 60. – P. 1469 – 1479. DOI: 10.1109/tim.2010.2086850
- [9] Mika, Hanhila, Timo, Mantere, & Jarmo, T. Alander (2018). «FPGA–implementation of PID-controller by differential evolution optimization», *Open Engineering*. –Volume 8, Issue 1. – P. 395 – 402, DOI: 10.1515/eng-2018-0038.
- [10] Rastegar, R., Hariri, A. (2006). "A Step Forward in Studying the Compact Genetic Algorithm", *Evolutionary Computation*. – Vol. 14. – No. 3. – P. 277 – 289. DOI: 10.1162/evco.2006.14.3.277

- [11] Storn, R. & Price, K. (1995)., "Differential Evolution - A Simple and Efficient Adaptive Scheme for Global Optimization over Continuous Spaces", *Journal of Global Optimization*. – Vol. 11. – P. 341 – 359.
- [12]. Chandra, A, & Chattopadhyay, S. (2012). "Role of Mutation Strategies of Differential Evolution Algorithm in Designing Hardware Efficient Multiplier-less Low pass FIR Filter", *Journal of Multimedia*. – Vol. 7. – No. 5. – P. 353 – 363. DOI: 10.4304/jmm.7.5.353-363
- [13]. Hiendro, A. (2011). "Multiple Switching Patterns for SHEPWM Inverters Using Differential Evolution Algorithms", *International Journal of Power Electronics and Drive System*. – Vol. 1. – P. 94 – 03. DOI: 10.11591/ijped.v1i2.101
- [14] Singh, R., Muzzio, F., Ierapetritou, M., & Ramachandran, R. (2015). «A Combined Feed-Forward/Feed-Back Control System for a QbD-Based Continuous Tablet Manufacturing Process», *Processes*, 3 (2), 339–356. DOI: 10.3390/pr3020339
- [15] Singh, R. (2015). «A Combined Feed-Forward/Feed-Back Control System for a QbD-Based Continuous Tablet Manufacturing Process», *Processes*. – Vol. 3. –, Issue 2. – P. 339 – 356. DOI: 10.3390/pr3020339
- [16] Besekerskiy, V. A., Popov, Ye. P. (2003). «Теория систем автоматического управления» [Theory of automatic control systems], SPb.: Profession. – 752 p.
- [17] Вычужанин В.В. (2005)/ Состояние рынка и динамика применений ПЛИС фирмы Alter CHIP NEWS. – №4. – С.22 – 26.
- [18] Вычужанин В.В. (2006). Рынок ПЛИС и их применение Электронные компоненты и системы. – №11. – С. 29 – 34.
- [19] Grout, Ian (2008). «Digital systems design with FPGAs and CPLD», Newnes. – 763 p. DOI: 10.1016/b978-0-7506-8397-5.x0001-3
- [20] Shyamala, G., M., Gurunadha, Babu, Muni, & Praveena Rela. (2014). «Digital PID Controller Implementation for Speed Control Applications Using FPGA», *International Journal of Scientific Engineering and Technology Research*. – Vol.03. –, Issue.48. – P.9745 – 9749.
- [21] Вычужанин В.В. (2011). Методика комплексного проектирования оптимальных систем управления с ПЛИС электродвигателями Современная электроника. – №7. – С.54 – 59.
- [22] Вычужанин В.В. (2011). Реализация цифровой обработки сигналов переменной точности на ПЛИС Цифровая обработка сигналов. – No.3. – С. 62 – 66.
- [23] Swarnalatha, A., & Shanthy, A. P. (2012). "Optimization of Single Variable Functions Using Complete Hardware Evolution", *Applied Soft Computing*. – Vol. 12. – P. 1322 – 1329. DOI: 10.1016/j.asoc.2011.12.001
- [24] Abdoalnasir, Almabrok, Mihalis, Psarakis & Anastasios, Dounis (2018). «Fast Tuning of the PID Controller in An HVAC System Using the Big Bang–

- Big Crunch Algorithm and FPGA Technology», *Algorithms Algorithms*. – Vol. 11. – P.1–19146, DOI:10.3390/a11100146
- [25] Hwang, K., & Cho, S. (2009). "Improving Evolvable Hardware by Applying the Speciation Technique", *Applied Soft Computing*. – Vol. 9. – P. 254 – 263. DOI: 10.1016/j.asoc.2008.01.011
- [26] Thoma, Y. & Sanchez, E. (2004)., "A Reconfigurable Chip for Evolvable Hardware", *GECCO, Springer-Verlag Berlin Heidelberg*. – Vol. – 3102. – P. 816 – 827. DOI: 10.1007/978-3-540-24854-5_84
- [27] Lima, João, Ricardo, Menotti, João, M. P. Cardoso & Eduardo, Marques (2006). «A methodology to design FPGA-based PID controllers». *Systems, Man and Cybernetics, IEEE Transactions*. – P. 2577-2583. DOI: 10.1109/icsmc.2006.385252
- [28] Вычужанин В.В. (2009). Повышение эффективности эксплуатации судовой системы комфортного кондиционирования воздуха при переменных нагрузках: Монография Одесса: ОНМУ.– 206 с.
- [29] Вычужанин В.В. (2007). Математические модели нестационарных режимов воздухообработки в центральной СКВ 2007, Вісник Одеського нац. морського ун-ту: Зб. наук. праць.– Одеса: ОНМУ. – №23. – С. 172–185.
- [30] Вычужанин В.В. (2007). Динамические свойства агрегатов центральной системы кондиционирования воздуха / Вісник Одеського національного морського університету, збірник наукових праць, випуск. –22. – С. 200 – 218.

OPTIMIZATION OF THE DIGITAL PID - CONTROLLER SETTING PARAMETERS USING THE DIFFERENTIAL EVOLUTION ALGORITHM

Vychujanin V.

The paper presents the results of the synthesis of a digital controller for automatic control systems of heat exchangers of central air conditioning systems. The developed regulator is designed to provide the specified quality of regulation, changing the settings of the regulators taking into account the operating conditions of the regulatory object. Implementation of the requirements for the controller is carried out in a typical digital PID - controller with optimization of its settings using the differential evolution algorithm. PID - controller is implemented in FPGA. Studies of a digital PID - controller with optimization of its parameters made it possible to establish that the controller satisfies the required quality of regulation. It has the ability to change settings, taking into account the operating conditions of the regulatory object.

Keywords. Digital PID - controller, differential evolution, programmable logic integrated circuit, heat exchanger.

ПРИМЕНЕНИЕ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ ДЛЯ АВТОМАТИЗАЦИИ ПРОЦЕССОВ КЛАССИФИКАЦИИ МАССИВОВ ТЕКСТОВЫХ ДАННЫХ БОЛЬШОГО ОБЪЕМА

Рудниченко Н., Вычужанин В., Шibaева Н., Шibaев Д.,
Отрадская Т., Петров И.

В статье приведены результаты исследования предложенного метода машинного обучения на базе искусственных нейронных сетей для автоматизации процессов классификации массивов текстовых данных большого объема. Проведен анализ применимости существующих методов машинного обучения для классификации текстовых объемов данных по ряду критериев. Формализована математическая модель и предложены процедуры повышения эффективности классификации текста на базе операций предобработки. Описаны ключевые этапы предложенного метода классификации текста на основе применения рекуррентных нейронных сетей.

Сформирована обобщенная схема входных и выходных данных ИС, разработана схема последовательности реализации ключевых функций, приведена диаграмма вариантов использования созданной системы для администратора и типового пользователя. Выполнена оценка точности классификации текста созданной моделью нейросети, рассчитаны значения метрик достоверности и функции потерь. Построены графики зависимостей значений оценок достоверности работы нейросети и функции потерь по эпохам обучения нейросети. Полученные результаты свидетельствуют о целесообразности и актуальности использования предложенного подхода к классификации текстовых данных

Ключевые слова. Классификация текста, интеллектуальный анализ данных, методы машинного обучения, сентимент-анализ, обработка естественного языка, анализ мнений, глубокое машинное обучение, искусственные нейронные сети

ВВЕДЕНИЕ

В настоящее время в информационной среде Internet наблюдается стремительный рост объемов разнородных данных, что связано с развитием и распространением социальных сетей, интернет-магазинов, тематических блогов и информационных веб-систем, что существенно отражается на активности развития разных направлений электронной коммерции и торговли различной электронной продукцией (EG) в частности [1]. В связи с регулярным появлением и активным развитием новых коммерческих и информационных ресурсов современные потребители виртуальных и физических товаров и услуг все чаще испытывают трудности в выборе компаний, организаций, производителей конкретных моделей технических гаджетов и средств

[2]. Это создает необходимость получения дополнительной информации о реальных функциональных возможностях и особенностях эксплуатации EG со стороны других пользователей и опытных экспертов. Дополнительные сложности вносит необходимость проведения фильтрации и анализа маркетинговых мероприятий конкурирующих компаний-производителей для выявления наиболее подходящих товаров и услуг для конкретных нужд пользователя, что требует выполнения большого количества вычислительных операций над данными [3]. С целью получения конкурентных преимуществ и лучшего понимания запросов клиентов организации вендоры также нуждаются в оперативном получении максимально достоверных и актуальных данных, извлекаемых из больших массивов информации, на базе проведения анализа мнений пользователей [4-6].

Частичное решение обозначенных проблем представляют существующие системы и информационные ресурсы, агрегирующие текстовые отзывы, комментарии и сравнительные видео-обзоры характеристик и специфики использования EG в разных условиях и режимах [7]. Однако, данные информационные площадки не всегда обладают гибким, удобным и информативным интерфейсом, системой тонкого поиска и поддержкой визуализации сводных статистических данных с формированием агрегированных и кросстабличных отчетов [8,9]. Анализ подобной информации, размещаемой на таких информационных ресурсах, часто бывает затруднен по причине необходимости просмотра интересующих отзывов и комментариев по товарам в ручном режиме, что сопряжено с большими временными затратами, т.е. процесс анализа формируемых пользователями мнений о предлагаемых товарах и оказываемых услугах является актуальным и трудоемким процессом [10,11]. В связи с этим целесообразным является автоматизация процесса оценки подходящих пользователю, по его индивидуальным предпочтениям, EG путем поиска и анализа собранных данных, характеризующих различные товары на основе решения задачи классификации их смыслового содержания по соответствующим группам.

Для решения обозначенной задачи на практике используются существующие подходы обработки естественного языка (NLP), в частности методы анализа тональности текста, морфологического анализа составляющих его сущностей и оценки эмоциональной окраски выражений [12]. Анализ тональности относится к использованию вычислительной лингвистики для идентификации и извлечения субъективной информации в исходных материалах [13]. Существующие подходы анализа тональности текста подразделяются на следующие основные категории: определение ключевых слов, лексическое сходство, статистические и концептуальные методы [14].

В общем виде задача определения типов пользовательских отзывов на приобретенные товары не является в полной мере четкой и однозначной, поэтому реализуется путем их классификации на отдельные группы в лингвистической форме. В различных работах по классификации пользовательских отзывов на различные современные товары на существующих информационных ресурсах часто применяются как стандартные методы классификации текстов, так и модифицированные методы, в которых учитывается возможная инверсия значений оценочных слов, синтаксическая структура предложений, зависимости между словами [15].

Спецификой и основной сложностью применения классических методов NLP для разных наборов пользовательских отзывов является необходимость сбора достаточного количества адекватных данных для обучения выбранной модели классификатора, выполнения ряда трудоемких подготовительных процедур по предобработке и очистке данных для обеспечения приемлемого уровня ее точности и скорости использования. В связи с этим целесообразно проведение анализа современных перспективных подходов по классификации текстов. В настоящее время на практике используется 2 подхода к решению поставленной задачи: методы на основе логических правил и машинного обучения [16].

Методы, базирующиеся на формировании логических правил, поддерживают возможность учета различных (семантических, структурных, пунктуационных) аспектов различных слов и самого языка, но их реализация сталкивается с рядом проблем:

- Требуется формировать определённый корпус различных лингвистических правил, который обязан учитывать обширную часть различных конструктивных языковых особенностей. Данный аспект требует привлечения групп экспертов в области лингвистики.

- Узкая сфера применения набора правил в связи с тем, что формат написания различных сообщений в сети Интернет достаточно сильно отличается от принятых норм русского языка в литературной форме. Сообщения, публикуемые в социальных сетях, отличаются тем, что они содержат ошибки пунктуационного и орфографического характера, имеют место для применения различных опечаток и словесного сленга, своеобразной пунктуации, а также использование специальных символов и графических обозначений для усиления эмоциональной окрашенности текста.

- Привязка к языку анализируемого текста всегда связана с уникальной языковой структурой и не может быть перенесена и применена для другого языка. Использование подхода, основанного на лингвистических правилах, может обеспечить высокие показатели результативности лишь в тех случаях, когда анализируемые тексты

будут грамматически верны, а также если различные конструкции анализируемого языка будут покрыты корпусом правил

Применение методов машинного обучения, подразумевает наличие некоторого набора входных данных, применяемых для обучения классификатора, и в свою очередь реализует алгоритмы обучения с учителем (для обучения используются размеченные примеры) и без учителя (использующие методы автоматической классификации). Данные методы более перспективны и обладают широким спектром возможностей в решении задачи классификации текстовых отзывов пользователей. При применении метода обучения с учителем, требуется наличие текстового корпуса, который заранее размечается метками полярности, в свою очередь идентифицирующими полярность для каждого текста из корпуса, а определение класса отзыва производится непосредственно автором текста, либо экспертом или их группой [17-19].

Использование подходов, основываемых на методах машинного обучения, позволяет подстраиваться под языковые особенности, включать в учёт дополнительные признаки, производить обработку текстов, которые, с точки зрения принятых языковых правил, являются грамматически неверными. К минусам этих подходов можно отнести некоторый проигрыш в отношении качества обработки и интерпретирования различных конструкций языка с высоким уровнем сложности [20,21].

В настоящее время разработано большое количество алгоритмов машинного обучения для решения задач классификации текстов, к наиболее популярным из которых относят: алгоритм k-ближайших соседей; алгоритм построения деревьев решений; алгоритм на основе машин опорных векторов; байесовский классификатор на основе смеси многомерных нормальных распределений и смеси распределений фон Мизеса-Фишера; классификатор Роше; искусственные нейронные сети (ИНС). Результаты анализа методов классификации текстовых данных приведены в табл. 1.

По результатам сравнительного анализа алгоритмов, выбран метод ИНС, как один из наиболее используемых на практике и перспективных в реализации. Дополнительным преимуществом данного метода является высокая функциональность существующих библиотек реализации моделей нейросетей от компании Google, их постоянная поддержка и обновление, что обеспечит возможности совершенствования системы в дальнейшем. Существующие решения на рынке анализа текстового контента имеют существенные ограничения в объеме вводимых данных для обработки, не обеспечивают гибких настроек при сборе и обработке текста на разных языках и не позволяют оценить точность отзывов с учетом смысловой тематики. В связи с этим актуальной задачей является разработка собственной

информационной системы, реализующей функционал по оценке пользовательских отзывов на ЕГ.

Таблица 1 – Результаты анализа методов классификации текстовых данных

Метод	Сравнение методов				
	Длительность обучения	Простота использования	Эффективность обработки больших объемов данных	Масштабируемость (для других классов текстов)	Функциональность существующих библиотек (1-10)
Алгоритм ближайших соседей	низкая	высокая	низкая	средняя	8
Алгоритм на основе машин опорных векторов	низкая	средняя	средняя	низкая	5
Деревья решений	средняя	высокая	низкая	низкая	7
Байесовский классификатор	средняя	высокая	средняя	низкая	5
Классификатор Роше	низкая	средняя	средняя	низкая	3
Искусственные нейросети	средняя	средняя	высокая	средняя	9

ФОРМУЛИРОВКА ЦЕЛИ ИССЛЕДОВАНИЯ.

Цель работы заключается в исследовании возможностей применения аппарата искусственных нейронных сетей для оценки пользовательских предпочтений по группам приобретаемых товаров путем автоматизации процесса анализа их мнений на базе решения задачи классификации.

ИЗЛОЖЕНИЕ ОСНОВНОГО МАТЕРИАЛА ИССЛЕДОВАНИЯ

Задача классификации текстовой информации определяется следующим образом. Пусть существует описание документа $d \in X$, где X - векторное пространство документов, и фиксированный набор классов $C = \{c_1, c_2, \dots, c_m\}$. Из обучающей выборки (множества документов с заранее известными классами) $D = \{\langle d, c \rangle \mid \langle d, c \rangle \in X \times C\}$ с помощью метода обучения G необходимо получить классифицирующую функцию $G(D) = \gamma$, которая отображает документы в классы $\gamma: X \rightarrow C$.

Используемая на практике модель Bag-of-Words для векторного представления документа W представляет собой некую параметризованную функцию $W: words \rightarrow R^n$, которая преобразует документ на естественном языке в некоторый n -мерный вектор. Подобную структуру следует описать таким образом

$$W(\textit{word and nextword}) = (w_{j,1}, w_{j,2}, \dots, w_{j,n}) \quad (1)$$

где W – представляет собой векторное представление текстового документа, $w_{j,n}$, – является значением веса термина j в текстовых данных, n -общее число терминов, находящееся в пространстве.

Функция отображения W формируется путем использования таблицы поиска посредством создания матрицы M , позволяющей идентифицировать однозначные соответствия для каждого из анализируемых слов текста $W_M = (word_n) = M_n$.

Таким образом, данная модель формирует векторное представление текста на естественном языке, сами текстовые данные задаются в виде набора терминов в не упорядоченном формате, без указания конкретных сведений связях между ними.

В качестве признаков классификации текстовых отзывов могут использоваться: n -граммы, представляющие собой ключевые термины, различные словоформы и наборы символьных последовательностей.

Типовыми n -граммами являются наборы последовательностей слов, которые расположены друг за другом и их суммарная длина равна N . Если $n > 1$, то путем оценки и анализа n -грамма возможно проведение учёта семантического контекста слова. В случае, когда анализируемый текст состоит из n числа предложений, а также m числа уникальных терминов, то матрица текста M будет иметь размерность $m \times n$. Для каждого отдельного термина, входящего в состав словаря определяется вес $w_{i,j}$, где i – является значением порядкового номера соответствующего термина в используемом словаре, j – это номер предложения [16].

Качество рассматриваемой нами задачи классификации может быть увеличено путем проведения дополнительных процедур по текстовой обработке:

1. Приведение всех встречающихся в тексте символов к нижнему регистру с целью уменьшения общего уникального числа терминов в словаре.

2. Исключение из текста не буквенных символов. Подобная процедура значительно понижает число уникальных терминов в словаре,

в тех случаях, когда для текста характерно обилие пунктуации, не несущей принципиальной смысловой нагрузки.

3. Исключение повторяющихся символов. Это позволяет заменить имеющиеся в тексте последовательности одинаковых символов для снижения размерности словаря.

4. Выделение основы слова из набора входных текстовых данных (стемминг).

Перечисленные действия выполняются перед процессом классификации текста, чтобы увеличить скорость и уменьшить итерационную и логическую сложность обработки данных.

Формальное описание предлагаемого метода проведения классификации в схематическом виде декомпозиции приведено на рис.1.

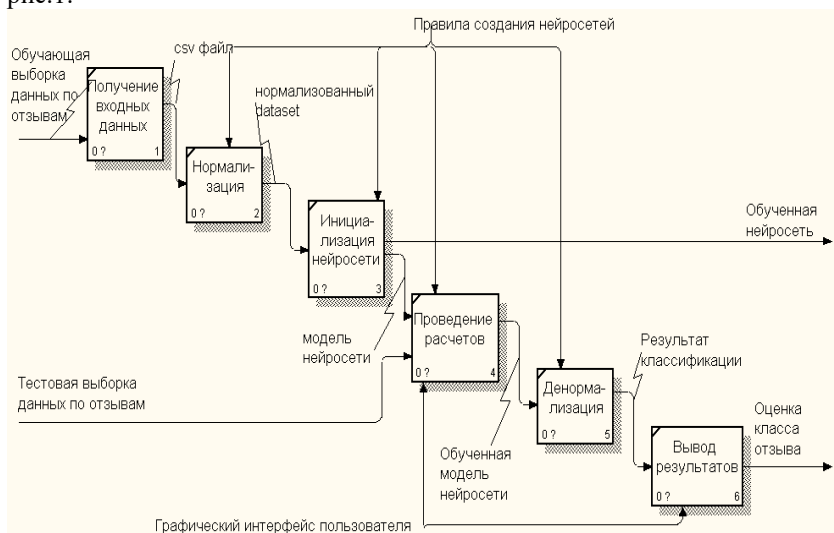


Рис.1 – Этапы реализации метода классификации на базе использования ИНС

Разрабатываемое программное обеспечение должно позволять проводить автоматический анализ входных текстов, предоставляя на выходе тип проанализированного текстового контента:

- положительный комментарий;
- негативный комментарий;
- нейтральный комментарий.

Для использования предложенного метода необходима разработка прикладного программного обеспечения (ПО) в виде информационной системы (ИС). Для обеспечения удобной и оперативной работы ИС, реализующей предложенный метод, необходимо внести ряд ограничений.

В связи с тем, что текстовые отзывы бывают различного размера и несут разную смысловую нагрузку, а обработка слишком больших фрагментов текста может быть трудоемкой и затратной с точки зрения расхода вычислительных ресурсов, целесообразно ограничить их объем. В частности, программой должна поддерживаться возможность проведения анализа текста на русском языке, общий объем текста должен составлять до 2000 символов, длительность анализа не должна превышать 10 секунд.

На вход ИС поступают текстовые данные пользовательских комментариев и отзывов, в результате обработки формируется таблица классов текста, рассчитывается уровень точности оценки (погрешность классификации модуль ИНС), форма сводной статистики и файл с выходными результатами классификации в формате *.xls (рис. 2).

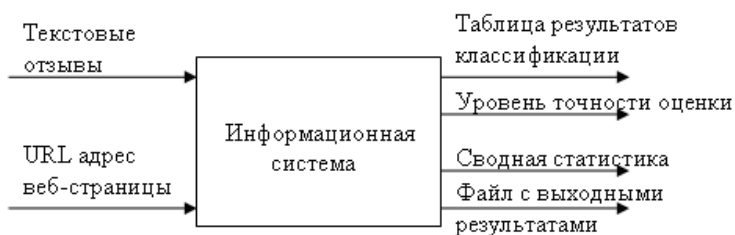


Рис. 2 – Обобщенная схема входных и выходных данных ИС

Основными этапами реализации проекта являются следующие:

1. Разработка программного модуля парсера для поиска, получения и сбора набора данных для формирования обучающей выборки нейросети.
2. Фильтрация данных по русскому языку и очистка посторонних символов, не несущих смысловой нагрузки в составе отзыва.
3. Экспорт полученной выборки в формат *.csv для импорта в структуру нейросети.
4. Создание и конфигурирование структуры нейросети, выбор алгоритмов обучения и оценки ее работы.
5. Разработка графического пользовательского интерфейса программного обеспечения, предусматривающего функции ввода текстового комментария и просмотра результата классификации.
6. Оценка принадлежности текста в одному из возможных классов отзыва.

Этап создания и конфигурации нейросети в более подробном виде подразделяется на ряд следующих задач:

1. Получение входной строки (массива строк), представляет собой процесс записи набора входных текстовых данных в переменную.
2. Нормализация входных данных, для преобразования всех элементов входного набора данных в двоичный код, который является

приемлемым для дальнейшей обработки искусственной нейронной сетью.

3. Инициализации модели ИНС, представляет собой процесс создания объекта нейросети, загрузки нормализованного набора данных и инициализации процесса обучения и сохранения модели искусственной нейронной сети.

4. Проведение расчётов, на базе подаваемых пользователем текстовых строк отзывов модель ИНС выполняет анализ входных данных и осуществляет их классификацию по доступным классам.

5. Преобразование результата проведённой классификации (денормализация), перевод полученных значений в текстовый вид, понятный пользователю человеку.

6. Вывод полученного значения во время выполнения данного этапа в интерфейсе пользователя отображается результат классификации.

Схематически, данные этапы отображены на рис.3.

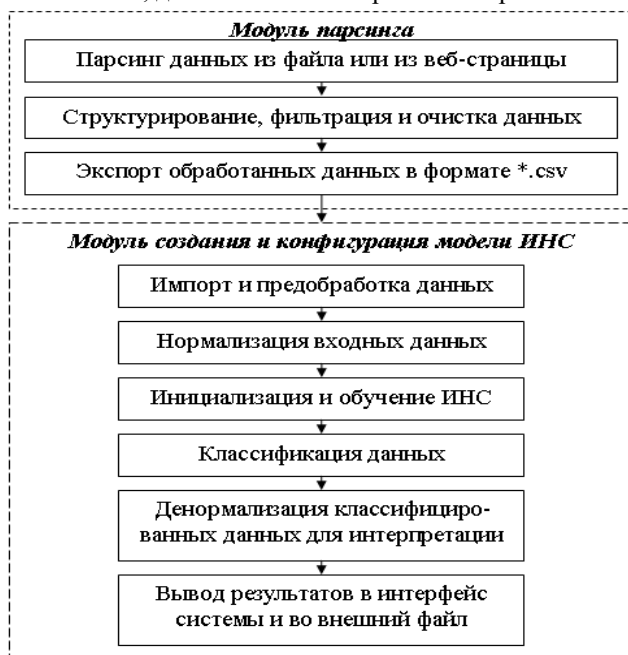


Рис.3 – Схема последовательности реализации ключевых функций ИС

В качестве языка разработки использован Python 3.7, который расширен средствами следующих библиотек обработки структур данных:

- Numpy, для поддержки использования многомерных массивов данных и реализации необходимых ряда математических функций по их обработке;
- Pandas, для имплементации функций моделирования и анализа при обработке и нормализации данных.

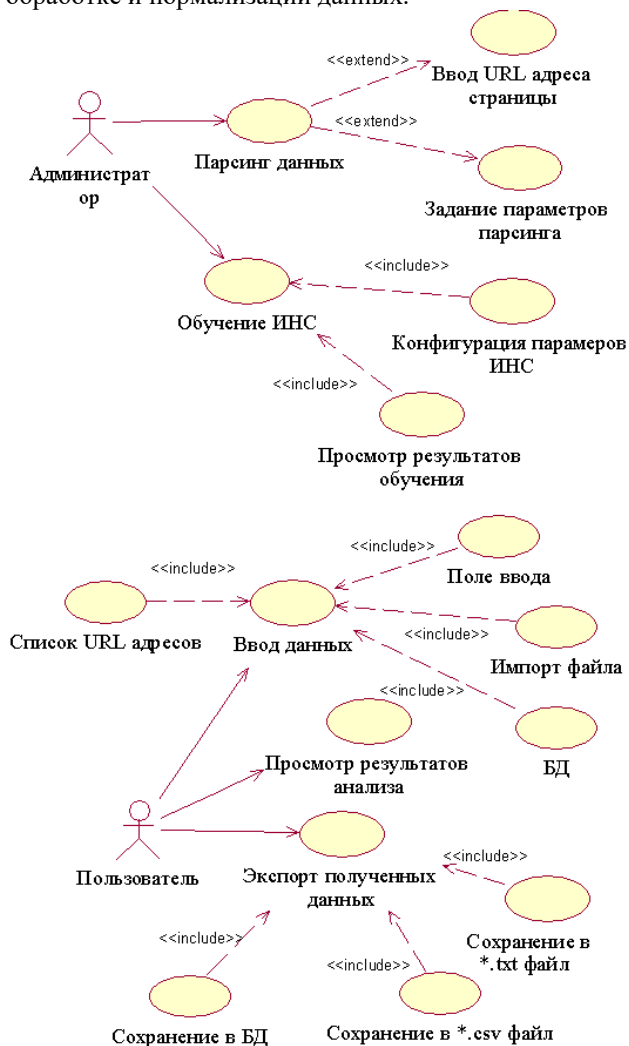


Рис.4 – Диаграмма вариантов использования ИС пользователем и администратором

Для нормализации и денормализации данных, создания, конфигурации и обучения модели ИНС применяется библиотека keras и

ее составляющие (модули tokenizer, TensorBoard, LSTM). Для создания графического пользовательского интерфейса, компоновки необходимых виджетов и элементов формы программы использована библиотека PyQt, а также модуль QtDesigner. При формировании требований к создаваемой системе разработана диаграмма прецедентов (рис. 4), в соответствии с которыми формализованы требования к ролям пользователей (представлены типовым пользователем и администратором системы). Пользователь должен имеет следующие возможности взаимодействия с ИС через графический интерфейс (форму): ввод и редактирование соответствующего текстового отзыва в рамках соответствующего текстового поля, просмотр результата анализа класса отзыва (положительный, отрицательный, нейтральный), экспорт полученного результата в текстовый файл. Администратор имеет возможности осуществлять парсинг данных с указанного URL страницы и задавать дополнительные параметры осуществления парсинга, а также проводить конфигурацию и обучение ИНС с просмотром полученных результатов.

АНАЛИЗ РЕЗУЛЬТАТОВ ИССЛЕДОВАНИЯ

Для выполнения исследования специфики функционирования созданного ПО на базе использования искусственных нейронных сетей была подготовлена выборка текстов отзывов на электронные товары с сайта Розетка: 45000 текстов (по 15000 на каждый из возможных классов).

Выборка была получена посредством разработки специализированного парсера данных, осуществляющего фильтрацию и очистку данных. Присвоение типов классов для каждой записи осуществлялось вручную.

Весь объем полученной выборки текстовых отзывов был разделен на обучающее и тестовое множество (60% и 40% соответственно), чтобы оценить качество работы созданного ПО.

В рамках проведения процесса исследования работы разработанного ПО проводилась оценка точность классификации, т.е. число корректно классифицируемых текстовых отзывов пользователей. В качестве числовых характеристик оценки работы ПО использовались:

ACCURACY – метрика достоверности, позволяющая оценить точность классификации, т.е. определить долю корректно классифицируемых документов

$$Accuracy = \frac{t_p + t_n}{t_p + f_p + f_n + t_n} \quad (2)$$

где t_p (True Positive) – истинно положительный вариант оценки класса отзыва, т.е. его фактическое значение (положительный отзыв) совпадает с результатом использования нейросети (положительный отзыв).

f_p (False Positive) — ложно положительный вариант оценки класса отзыва, т.е. фактическое значение (отрицательный отзыв) не совпадает с результатом классификации нейросетью (положительный отзыв).

f_n (False Negative) — ложно отрицательный вариант оценки класса отзыва (ошибка 2-го рода, это возможно в случае, когда фактическое значение (положительный отзыв) не совпадает с результатом классификации нейросетью (отрицательный или нейтральный отзыв).

t_n (True Negative) — истинно отрицательный вариант оценки класса отзыва, т.е. его фактическое значение (отрицательный отзыв) совпадает с результатом использования нейросети (отрицательный отзыв).

LOSS – функция потерь при работе нейросети, данный показатель иллюстрирует зависимость точности обучения от коэффициентов весовой матрицы

$$MSE = \frac{1}{N} \sum_{i=1}^N (Z(t) - \bar{Z}(t))^2 \quad (3)$$

где $Z(t)$ – фактическое значение класса отзыва.

$\bar{Z}(t)$ – значение класса полученного в процессе использования нейросети.

N - общее число текстовых отзывов в используемой выборке.

Для проведения численных исследований использования созданной модели нейросети в рамках разработанной ИС и визуализации графических зависимостей полученных результатов было развернуто средство анализа данных Tensor Board. Зависимость значения оценки достоверности работы нейросети (ось ординат) по пройденным эпохам обучения (ось абсцисс) приведена на рис. 5.

Тонкой красной линией отмечены результаты обучающей выборки отзывов, а толстой красной линией отображены результаты использования нейросети на тестовой выборке. Общая точность работы созданной нейросети составила порядка 89%. Зависимость значений функции потерь от эпохи обучения нейросети приведена на рис.6.

Таким образом, в результате проведенного анализа работы ПО установлено, что созданная нейросеть смогла успешно классифицировать порядка 90% текстовых отзывов пользователей.

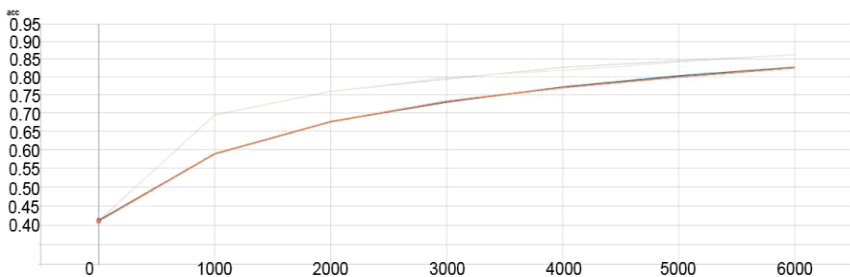


Рис. 5 – Зависимость значения оценки достоверности работы нейросети по пройденным эпохам обучения

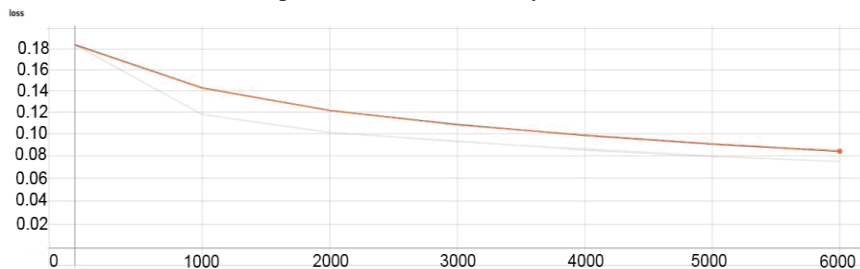


Рис. 6 – Зависимость значений функции потерь от эпохи обучения нейросети

ВЫВОДЫ

Разработанное программное обеспечение поддержки процессов оценки пользовательских отзывов является кроссплатформенным и обеспечивает достаточно высокую точность классификации, более 90%, что свидетельствует о достоверности решения задачи.

На базе полученных результатов классификации пользовательских отзывов становится возможным формирование агрегированного интегрального показателя оценки соответствующих товаров, который может быть использован для приоритетного представления предпочтений покупателей в ранжированном виде с целью поддержки и облегчения процессов принятия решений по выбору и покупке.

Крупные торговые площадки могут использовать полученные результаты оценки мнений пользователей для анализа и выбора наиболее авторитетных и надежных вендоров для дальнейшего сотрудничества или прекращения закупок у поставщиков, продукция которых регулярно критикуется покупателями.

Последующим логическим развитием предложенного подхода к классификации отзывов пользователей является интеграция механизмов анализа достоверности выборки данных с целью отсеивания шумовых и

не информативных данных, расширение типов классов и имплементация соответствующего им ряда количественных показателей для уточнения формируемых оценок.

СПИСОК ЛИТЕРАТУРЫ

- [1] Вычужанин В.В. Методы информационных технологий в диагностике состояния сложных технических систем. Монография / В.В. Вычужанин, Н.Д. Рудниченко. – Одесса: Экология, 2019. – 178 с.
- [2] Шibaев Д.С. Оптимизация отбора и анализа информации в разноструктурных хранилищах данных / Д.С. Шibaев, В.В. Вычужанин, Н.О. Шibaева, Н.Д. Рудниченко // Информатика и математические методы в моделировании, 2017. – №3. – С.318-324.
- [3] Шibaев Д.С. Оптимизация методов прогнозирования, обработки и анализа информации в разноструктурных хранилищах данных / Д.С. Шibaев, В.В. Вычужанин, Н.О. Шibaева, Н.Д. Рудниченко // Информатика и математические методы в моделировании, 2018. – №1. – С.78 – 85.
- [4] Шibaев Д.С. Повышение эффективности методов отбора и анализа информации в разноструктурных хранилищах данных / Д.С. Шibaев, В.В. Вычужанин, Н.Д. Рудниченко // 21-й міжнародний молодіжний форум «Радіоелектроніка та молодь у XXI столітті». Зб. Матеріалів форуму. – Харків: Хнуре, 2018. – Т.5. – С.221 – 222.
- [5] Вычужанин В.В. Распределенный программный комплекс на базе фреймворка Apache Spark для обработки потоков Big Data От сложных технических систем/ В.В. Вычужанин// Информатика и математические методы в моделировании, 2018. – Том 8. – №2. – С.146 – 155.
- [6] Vychuzhanin V.V. Big data mapping in the geopositioning systems for fishing industry / V.V. Vychuzhanin, D.S. Shibaev, V.D. Boyko, N.O. Shibaeva, N.D. Rudnichenko // International Scientific and Technical Conference on Computer Sciences and Information Technologies (CSIT). – 2017. – P. 28 – 31.
- [7] Адаскина Ю. В. Сентиментный анализ твитов на основе синтаксических связей / Ю. В. Адаскина, П. В. Паничева, А. М. Попов // Компьютерная лингвистика и интеллектуальные технологии: по материалам ежегодной Международной конференции «Диалог». – Москва : РГГУ, 2015. – С. 25 – 35.
- [8] Васильев В.Г. Классификация отзывов пользователей с использованием фрагментных правил / В.Г. Васильев, М.В. Худякова, С. Давыдов // Компьютерная лингвистика и интеллектуальные технологии: По материалам ежегодной Международной конференции «Диалог» (Бекасово, 30 мая–3 июня 2012 г.). – 2012. – Т1. – С. 66 – 78.

- [9] Гаршина В. В. Разработка системы анализа тональности текстовой информации / В. В. Гаршина, К. С. Калабухов, В. А. Степанцов, С. В. Смотров // Вестник ВГУ, серия: системный анализ и информационные технологии, 2017. – № 3. – С.185 – 194.
- [10]. Лысенко В. Д. Анализ тональности текста для прогнозирования цен на фондовом рынке // Молодой ученый, 2018. – №22. – С. 420 – 423.
- Павлов Ю. Н., Майструк К. А. Сравнение методов оценки тональности текста // Молодой ученый, 2016. – №12. – С. 59 – 64.
- [11] Loukachevitch N. SentiRuEval: testing object-oriented sentiment analysis systems in Russia / N. Loukachevitch, E. Kotelnikov , Y. Rubtsova // Proceedings of International Conference Dialog-2015, Moscow,2015. – 313 с.
- [12] Рубцова Ю. В. Построение корпуса текстов для настройки тонового классификатора / Ю. В. Рубцова // Программные продукты и системы. – Новосибирск: Научно-исследовательский институт «Центрпрограммсистем», 2015 – № 109 – С. 72 –78.
- [13] Меньшиков И.Л., Кудрявцев А. Г. Обзор систем анализа тональности текста на русском языке // Молодой ученый, 2012. – №12. – С. 140 – 143.
- [14] Котельников Е. В. Автоматический анализ тональности текстов на основе методов машинного обучения / Е. В. Котельников, М. В. Клековкина // Компьютерная лингвистика и интеллектуальные технологии: по материалам ежегодной Международной конференции «Диалог». – М : РГГУ, 2012. – С. 15 – 21.
- [15] Рудниченко Н.Д. Применение кластерного анализа данных для выделения меры схожести факторов влияния на работоспособность сложных технических систем / Н.Д. Рудниченко, В.В. Вычужанин, Д.С. Шibaев // Информатика и математические методы в моделировании, 2017. – №3. – С. 214 – 219.
- [16] Щербина А.Д. Порівняльний аналіз існуючих напрямів у інтелектуальному аналізі даних / А.Д. Щербина, Д.С. Шibaев, М.Д. Рудніченко, Н.О. Шibaєва // Project, Program, Portfolio Management The Third International Scientific-practical Conference, Odesa, ONPU 07–08 Dec 2018. – С.88 – 90.
- [17]. Рудниченко Н.Д. Разработка модели нейросети для прогнозирования риска отказов компонентов сложных технических систем / Н.Д. Рудниченко, В.В. Вычужанин // Информатика и математические методы в моделировании, 2016. – №4. – С. 333 – 338.
- [18] Сбоев А. Г. Продвинутые нейросетевые модели для решения задачи определения тональности / А. Г. Сбоев, И. Е. Воронина, Д. В. Гудовских, А. А. Селиванов// Вестник Воронеж. гос. ун-та. Сер. Системный анализ и информационные технологии, 2016. – № 4. – С. 178 – 183.
- [19] Горбань А.Н. Обучение нейронных сетей / А.Н. Горбань. – М.: "ParaGraph", 2010. – 160 с.

[20] Shybaiev D. S. Predicting system for the estimated cost of real estate objects development using neural networks / D. S. Shybaiev, T. V. Otradskaya, M. V. Stepanchuk, N. O. Shybaieva, N. D. Rudnichenko // Вісник ЖДТУ. Технічні науки, 2019. – №1 (83). – С.154 – 160.

[21] Рудниченко Н.Д. Применение нейронных сетей для оценки и классификации рисков элементов и межэлементных связей сложных технических систем / Н.Д. Рудниченко, В.В. Вычужанин, В.Д. Бойко, С.Н. Коновалов, Н.О. Шibaева // Интеллектуальные системы принятия решений и проблемы вычислительного интеллекта: Материалы международной научной конференции. – Херсон: Видавництво ПП Вишемирський В.С., 2016. – С.265 – 266.

APPLICATION OF MACHINE LEARNING METHODS FOR AUTOMATION OF CLASSIFICATION OF BULK TEXT DATA ARRAYS

Rudnichenko N., Vychujanin V., Shibaeva N., Shibaev D., Otradskaya T., Petrov I.

The article presents the results of a study the modern machine learning method based on artificial neural networks to automate the large text data arrays amounts classification. An analysis is made of the existing machine learning methods applicability for classifying textual volumes of data according to a number of criteria. A mathematical model is formalized and procedures are proposed for increasing text classification efficiency based on preprocessing operations. The key stages of the proposed method for text classification based on the use of recurrent neural networks are described. A generalized scheme of the the information system software input and output data has been made, a sequence diagram of the implementation of key functions has been developed, a diagram of the options for using the created system for the administrator and standard user is presented. The accuracy of the neural network model text classification has been evaluated, the values of reliability metrics and loss functions have been calculated. The dependencies graphs of the neural network reliability estimates values for the completed training eras and the loss function for the eras are constructed. The results obtained indicate the feasibility and relevance of using the proposed approach for the text data classification.

Keywords. Text classification, machine learning methods, sentiment analysis, natural language processing, opinion analysis, deep machine learning, artificial neural networks

УДОСКОНАЛЕННЯ МЕТОДУ ВИЯВЛЕННЯ БЛОКОВОЇ ОБРОБКИ ЦИФРОВОГО ЗОБРАЖЕННЯ

Хорошко В., Кобозєва А., Бобок І.

У даний момент обробка зображення часто носить блоковий характер, коли обробці послідовно/паралельно піддаються окремі блоки матриці зображення. Метою роботи є забезпечення можливості виявлення порушення цілісності зображення внаслідок його блокової обробки в умовах, коли місце розташування сітки, яка використовується при обробці, невідомо, шляхом удосконалення методу виявлення блокової обробки, розробленого раніше. Запропоноване вдосконалення використовує для аналізу параметр, що відображає зміну кількості непересічних блоків малого розміру, для яких кут між нормованим вектором квадратів сингулярних чисел і сингулярним вектором, що відповідає максимальному сингулярному числу, дорівнює куту між першим вектором стандартного базису відповідного простору й n -оптимальним вектором, при зсуві сітки розбивки. Установлено, що алгоритмічна реалізація вдосконаленого методу незначно поступається по ефективності методу, взятому за основу, порівнянню по ефективності з існуючими аналогами.

Ключові слова. Цифрове зображення, порушення цілісності, блокова обробка зображення, сингулярний вектор, сингулярне число, сітка розбиття.

ВСТУП

Цілісність будь-якого інформаційного контенту, зокрема цифрового зображення (ЦЗ), аналізу якого присвячена робота, є необхідною умовою для його використання з нерозважальними цілями, залишаючи задачу виявлення порушень цілісності актуальною й надзвичайно важливою на сьогоднішній день [1,2].

Нещодавно в [3] був запропонований, а в [4] удосконалений новий підхід до розв'язку проблеми виявлення порушень цілісності ЦЗ, відповідно до якого для більшості $l \times l$ – блоків малого розміру, отриманих шляхом стандартної розбивки [5] матриці оригінального ЦЗ має місце наступне співвідношення:

$$\angle(u_1, \bar{\sigma}) \approx \angle(v_1, \bar{\sigma}) \approx \angle(n^o, e_1), \quad (1)$$

де $\angle(a,b)$ – кут між векторами a,b ; u_1 і v_1 – відповідно лівий і правий сингулярні вектори (СНВ) $l \times l$ -блоку, що отримані шляхом його нормального сингулярного розкладання [6], які відповідають максимальному сингулярному числу σ_1 цього блока, $\sigma_1 \geq \dots \geq \sigma_l \geq 0$ – сингулярні числа (СНЧ) блока

$$\bar{\sigma} = (\sigma_1^2, \sigma_2^2, \dots, \sigma_l^2)^T / \left\| (\sigma_1^2, \sigma_2^2, \dots, \sigma_l^2)^T \right\| \in R^l, \quad (2)$$

$n^0 = (1/\sqrt{l}, 1/\sqrt{l}, \dots, 1/\sqrt{l})^T \in R^l$ – n -оптимальний вектор простору R^l , $e_1 = (1, 0, \dots, 0) \in R^l$ – перший вектор стандартного базису простору R^l .

На сьогоднішній день велике поширення отримала блокова обробка ЦЗ, яка використовується з різними цілями, коли обробці послідовно/паралельно піддаються окремі блоки, що отримуються в результаті розбивки матриці зображення [7]. Така обробка використовується при стиску ЦЗ з втратами (формати Jpeg, Jpeg2000), клонуванні й не тільки. Дуже часто порушення цілісності ЦЗ відбувається при його використанні в якості контейнера при організації прихованого (стеганографічного) каналу зв'язку. Виявлення такого каналу засобами стеганоаналізу є актуальною не вирішеною остаточно задачею захисту інформації. Найчастіше стеганоперетворення ЦЗ сучасними стеганографічними методами відбувається поблоково [8-10]. Це пов'язано з декількома причинами [11]:

- сучасні стеганографічні методи повинні бути стійкими до стиску з втратами для забезпечення можливості збереження ЦЗ-стеганоповідомлення в такому форматі (найчастіше - Jpeg), враховуючи те, що сьогодні передача й зберігання інформації відбувається, як правило, в форматах з втратами. Оскільки алгоритм стиску Jpeg розбиває матрицю на блоки, то процес стеганоперетворення доцільно проводити також поблоково для врахування особливостей зміни параметрів ЦЗ-стеганоповідомлення при стиску;

- сучасні стеганографічні методи повинні мати малу обчислювальну складність. Ця вимога стає критичною при необхідності забезпечення можливості їх роботи в режимі реального часу з потоковим контейнером, який усе частіше використовується на практиці. Обчислювальна складність будь-якого блокового алгоритму (методу) визначається кількістю блоків, тобто для $n \times n$ -матриці F ЦЗ-

контейнера складе $O(n^2)$ операцій, що робить такий алгоритм перспективним для використання в умовах потокового контейнера;

- обробка блоків зображення може відбуватися одночасно, що дозволяє розпаралелювати процес такої обробки, зменшуючи значно її час.

З врахуванням вище наведеного блокова обробка ЦЗ є розповсюдженою й широко використовуваною на сьогоднішній день, що робить задачу її виявлення *актуальною*.

В [11] показано, що блокова обробка ЦЗ має свої особливості: її природнім результатом найчастіше є зменшення кореляції значень яскравості для сусідів-пікселів, які виявляться на границях блоків, що використовуються при обробці. Це має наступні наслідки: відмінності в змінах кількостей блоків малого розміру, для яких виконується співвідношення

$$\angle(u_1, \bar{\sigma}) = \angle(n^o, e_1), \quad (3)$$

матриці оригінального ЦЗ й такого, цілісність якого була порушена в результаті блокової обробки, при зсуві сітки розбивки матриці при експертизі ЦЗ. На підставі результатів досліджень, отриманих в [11], в [12] на основі (1) був запропонований метод (і його алгоритмічна реалізація) виявлення блокової обробки ЦЗ, основними об'єктами аналізу в якому є:

$$SM = \left| \bar{k}_2 - k_2 \right|, \quad SM1 = \left| \bar{k}_4 - k_4 \right|, \quad (4)$$

які характеризують величини збурень кількостей блоків, що розглядаються, різних малих розмірів при різних розташуваннях сітки розбивки матриці ЦЗ. В алгоритмічній реалізації метода k_2 і k_4 (%) - відносні кількості 2×2 - і 4×4 -блоків у кожному ЦЗ, для яких виконувалася умова (3), \bar{k}_2 і \bar{k}_4 (%) – аналогічні параметри, які обчислювалися після зсуву сітки. Алгоритмічна реалізація запропонованого методу є ефективною в умовах різноманітних блокових збурних дій: збурень сингулярних чисел і СНВ блоків (з урахуванням того, що формальним представленням результату будь-якої збурної дії для ЦЗ є сукупність збурень СНЧ і СНВ його матриці (блоків матриці), отриманих за допомогою нормального сингулярного розкладання [2]), при виявленні результатів стиску ЦЗ з втратами, в умовах різних комплексів атак, що містять блокову обробку, при використанні його в якості стеганоаналітичного. Але в розробленому методі передбачається варіант, коли первісна розбивка матриці ЦЗ, що піддається експертизі за його допомогою, включає розбивку при блоковій обробці зображення.

Ця вимога легко виконується, якщо відомим є місце розташування сітки розбивки при блоковій обробці, але якщо досліджуване зображення, наприклад, зазнало обрізки, результатом якої є зміна розташування початку координат для ЦЗ і яка дуже легко і часто реалізується, то відомості про розташування сітки, яка використовувалася при блоковій обробці, будуть втраченими. Це не дає в загальному випадку можливості забезпечити умову первісної розбивки матриці, наведену вище, що звужує область застосування методу, розробленого в [12].

МЕТА РОБОТИ ТА ПОСТАНОВКА ЗАДАЧ

Метою роботи є забезпечення можливості виявлення порушення цілісності ЦЗ внаслідок його блокової обробки в умовах, коли місце розташування сітки, що використовувалася при обробці, є невідомим, шляхом удосконалення методу виявлення блокової обробки ЦЗ, розробленим в [12].

Для досягнення мети в роботі розв'язуються наступні *задачі*:

1. З урахуванням очікуваного зростання обчислювальної складності в результаті удосконалення, серед параметрів, які аналізуються методом, розробленим в [12], вибрати такі, що є більш придатними для досягнення поставленої мети, що дасть можливість зменшити кількість операцій при аналізі зображення при конкретній розбивці його матриці;

2. Удосконалити метод виявлення блокової обробки ЦЗ та розробити його алгоритмічну реалізацію;

3. Провести оцінку ефективності, в тому числі порівняльну, розробленої алгоритмічної реалізації удосконаленого методу виявлення блокової обробки ЦЗ.

Об'єкт дослідження: процес виявлення порушення цілісності цифрового контенту.

Предмет дослідження: метод виявлення блокової обробки цифрового зображення.

Методи дослідження: матричний аналіз [13], теорія збурень [2], методи цифрової обробки зображень [5].

ОСНОВНА ЧАСТИНА

Параметри (4) ведуть себе по-різному у випадку оригінального ЦЗ і такого, яке зазнало блокову обробку: для оригінального ЦЗ [11]

$$(SM \leq 1) \& (SM1 \leq 1) \quad (5)$$

Співвідношення (5) порушується у випадку блокової обробки зображення.

Конкретні розміри блоків, що використовуються при експертизі ЦЗ, а саме 2×2 і 4×4 , обрані таким чином, щоб забезпечити можливість для блоків малого розміру задовольняти умові (3). В [11] показано, що такі блоки належать областям ЦЗ з малими перепадами значень яскравості. З урахуванням того, що чим менше розміри блоку, тим точніше ці блоки при розбивці матриці зображення будуть покривати такі області (в сенсі зменшення різниці між площею області і площею фігури, що складається з блоків розбивки (рис.1)), як параметр, що буде аналізуватися при експертизі цілісності ЦЗ буде використовуватися SM , що характеризує зміну кількості 2×2 -блоків ЦЗ, що задовольняють умові (3), при зміні розташування сітки розбивки.

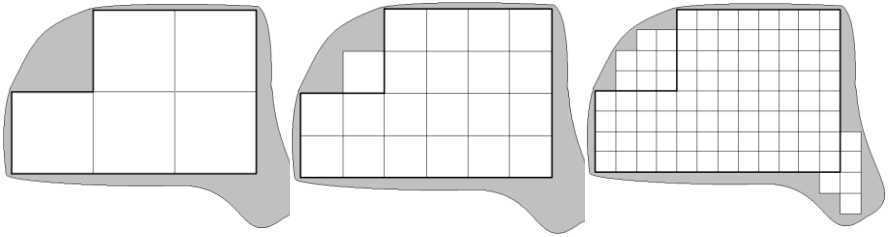


Рис.1 – Область ЦЗ із малими перепадами значень яскравості й блоки розбивки матриці зображення малих розмірів, що її складають

Якісний стрибок в кількості 2×2 -блоків, що задовольняють умові (3), буде відбуватися при зміні взаємного розташування сіток розбивки, що використовуються при блоковій обробці ЦЗ та при його експертизі за допомогою 2×2 -блоків з варіанту, представленим на рис.2(а), до варіанту, що відповідає рис.2(б), що докладно пояснено в [11]. Якщо припустити, що ЦЗ, яке піддається експертизі, попередньо зазнало обрізки, то визначити, якою повинна бути розбивка при експертизі для того, щоб отримати варіант рис.2(а) (варіант рис.2(б)) взаємного розташування сіток, не представляється можливим. Але, якщо сітка розбивки при експертизі буде зсуватися в одному й тому ж обраному напрямі на 1 піксель не більше, ніж $(l-1)$ разів, то на якомусь кроці зміна двох послідовних взаємних розташувань сіток розбивки при блоковій обробці і експертизі зображення обов'язково буде відповідати варіантам, представленим на рис.2. Наслідком цього буде $SM > 1$ на обговорюваному кроці.

З урахуванням вище наведеного основні кроки, що відповідають удосконаленому методу виявлення блокової обробки зображення, який розрахований на випадок невідомого місця розташування сітки розбивки при блоковій обробці, наступні.

Крок 1. Обрати напрям v зсуву сітки розбивки матриці ЦЗ.

Крок 2. Матриця F ЦЗ, що піддається експертизі, розбивається на непересічні блоки малого розміру $l \times l$ пікселів.

Крок 3. Для отриманих блоків визначити відносну кількість k_l таких, що задовольняють (3).

Крок 4. Матриця F ЦЗ, що піддається експертизі, розбивається на непересічні $l \times l$ -блоки, при цьому сітка розбивку зсунута відносно попередньої в напрямку v на 1 піксель.

Крок 5. Для отриманих блоків визначити відносну кількість \bar{k}_l таких, що задовольняють (3).

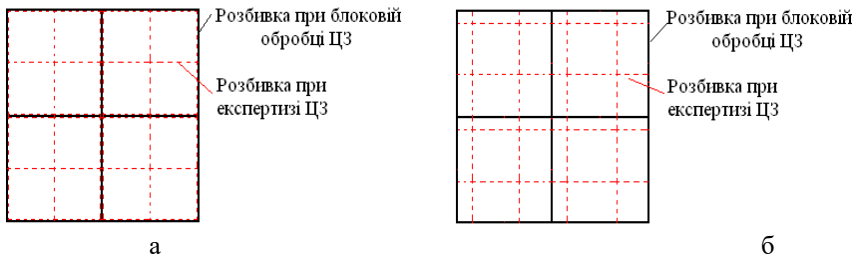


Рис. 2 – Варіанти взаємного розташування сіток розбивки матриці ЦЗ на блоки при блоковій обробці й експертизі ЦЗ: а - сітка розбивки, що використовується при експертизі ЦЗ, містить у собі сітку, використану при обробці ЦЗ; б - сітки розбивки зсунуті одна відносно іншої

Крок 6. Визначити

$$SM = \left| \bar{k}_l - k_l \right|.$$

Крок 7.

Якщо

$$SM > 1$$

то

цілісність ЦЗ порушена; перехід на крок 8,

інакше

якщо

кількість зсувів сітки не перевищує $(l-1)$ разів

то

$$k_l = \bar{k}_l,$$

повернутися на крок 4,

інакше

цілісність ЦЗ не порушена.

Крок 8. Кінець експертизи.

При алгоритмічній реалізації методу використовувалися наступні значення параметрів: $\nu = -45^\circ$, $l = 2$, $m = 8$ (як найчастіше використовуваний на практиці розмір блоку при блокувій обробці ЦЗ (збереження зображення в форматах Jpeg, Jpeg2000, блокові стеганографічні методи та алгоритми [10])).

Зауваження. Завдяки тому, що запропонований удосконалений метод виявлення блокувій обробки ЦЗ сам є блокувим, обчислювальна складність його алгоритмічної реалізації SAu для експертизи ЦЗ з $n \times n$ -матрицею становить $O(n^2)$, як і для алгоритмічної реалізації SA для методу, запропонованого в [12], але обчислювальна складність SAu максимально

може перевищувати складність SA в $t = \frac{1}{2}(l-1)$ разів, де l – розмір квадратних блоків, що використовуються при експертизі, якщо кількість зсувів сітки розбиття при експертизі буде максимальною. Коефіцієнт $\frac{1}{2}$

в виразі для t виникає завдяки тому, що в удосконаленому методі виявлення блокувій обробки ЦЗ аналізується лише параметр SM (4) на відміну від первісного методу [12], що говорить про розбивку матриці ЦЗ лише з одним розміром блоку при експертизі, тоді як в методі-прототипі розбивка відбувається на блоки двох розмірів, і аналізуються обидва параметри (4).

Оцінка ефективності алгоритмічної реалізації SAu методу проводилася шляхом обчислювального експерименту, в якому були задіяні:

- множина M_1 : 200 ЦЗ розміром 400×400 пікселів в форматі Tif [15];
- множина M_2 : 300 ЦЗ розміром 400×400 пікселів в форматі Tif [16];
- множина M_3 : 80 ЦЗ розміром 400×400 пікселів в форматі Tif, отриманих непрофесійними відеокамерами.

При тестуванні алгоритма SAu на першому етапі як блокува обробка розглядався стиск з втратами (збереження в формат Jpeg з коефіцієнтами якості $QF \in Q = \{50,65,75,85,95\}$). Виявлення в якості блокувій обробки стиску з втратами було обрано, оскільки стиск на сьогодні – це сама поширена операція блокувій обробки ЦЗ. Результати збереження з втратами ЦЗ з множин $M_i, i = 1,2,3$, з коефіцієнтами

якості $QF \in Q$ склали множини $M_i^{(j)}, i=1,2,3, j \in Q$, відповідно. ЦЗ в множинах $M_i^{(j)}, i=1,2,3, j \in Q$, піддавалися обрізці, в результаті якої початок координат зсувався від одного до 5 пікселів вздовж осі OX і від 1 до 5 пікселів вздовж OY , після чого піддавалися експертизі за допомогою алгоритмічної реалізації удосконаленого методу. Інформація щодо помилок 1-го роду наведена в табл.1. Середнє значення помилок 1-го роду по всьому експерименту – 9.4%. Помилки 2-го роду склали 10.1%.

Для зручності порівняння ефективності розробленого в роботі алгоритму SA_u із сучасними аналогами за отриманими даними були розраховані значення точності виявлення порушення цілісності [17] (accuracy (ACC)):

$$ACC = (TP + TN) / (TP + FN + TN + FP), \quad (6)$$

де TP (*True Positive*) — число правильно виявлених ЦЗ, цілісність яких була порушена (істинопозитивний результат); TN (*True Negative*) — число правильно виявлених оригінальних ЦЗ (істинонегативний результат); FP (*False Positive*) — число оригінальних ЦЗ, помилково прийнятих за такі, цілісність яких була порушена (хибнопозитивний результат (хибна тривога) або помилка другого роду); FN (*False Negative*) — число ЦЗ, цілісність яких була порушена, помилково визнаних оригінальними (хибнонегативний результат або помилка першого роду). Результати експерименту наведені в табл.1, рис.3.

Таблиця 1 – Результати оцінки ефективності розробленого алгоритму виявлення блокової обробки ЦЗ (стиску з втратами) в умовах невизначеності сітки розбивки матриці зображення при блоковій обробці

Множина	Помилки 1-го роду	ACC
$M_1^{(50)} \cup M_2^{(50)} \cup M_3^{(50)}$	11.1	0.89
$M_1^{(65)} \cup M_2^{(65)} \cup M_3^{(65)}$	9.7	0.90
$M_1^{(75)} \cup M_2^{(75)} \cup M_3^{(75)}$	9.7	0.90
$M_1^{(85)} \cup M_2^{(85)} \cup M_3^{(85)}$	6.6	0.92
$M_1^{(95)} \cup M_2^{(95)} \cup M_3^{(95)}$	9.8	0.90

Як впливає з результатів обчислювального експерименту (рис.3), алгоритмічна реалізація удосконаленого методу виявлення блокової

обробки ЦЗ внаслідок його стиску з втратами лише незначно поступається по ефективності первісному методу, значно перевищує при $QF > 60$ по ефективності метод, запропонований в [18]. Запропонований метод може використовуватися в якості стеганоаналітичного, аналізу ефективності чого присвячений другий етап обчислювального експерименту.

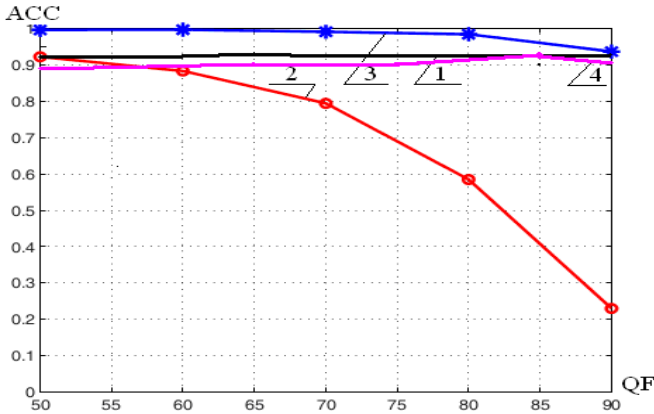


Рис. 3 – Результати виявлення наявності стиску з втратами в ЦЗ, збережених в форматі Jpeg: 1- алгоритмічна реалізація методу блокової обробки, запропонованого в [12]; 2 – алгоритм, запропонований в [18]; 3 – алгоритм, запропонований в [19]; 4 – алгоритм SAu

На сьогоднішній день одним з найпоширеніших стеганографічних методів залишається метод модифікації найменшого значущого біта – LSB-метод [10]. І хоча цей метод не є блоковим, виявлення результатів його роботи принципово можливе за допомогою розробленого удосконаленого методу у випадку, коли в якості контейнера використовується ЦЗ, яке зберігається у форматі з втратами. Найбільшу увагу тут треба звернути на забезпечення можливості детектування вкладень додаткової інформації з малою пропускнуною спроможністю прихованого каналу зв'язку (ПСПК): $ПСПК \leq 0.1 \text{біт/піксель}$, оскільки для такої ПСПК абсолютна більшість існуючих стеганоаналітичних методів виявляються неробочими або неефективними, а частота використання LSB-методу саме з малою ПСПК сьогодні зростає. Результати оцінки ефективності SAu представлені в табл.2.

Результати порівняльного аналізу ефективності розробленого алгоритму SAu із сучасними стеганоаналітичними алгоритмами, спрямованими на виявлення LSB-вкладень, за допомогою показника

ACC (7), представлені в табл.3, де використовуються позначення: SS1 (2006) [20], SS2 (2006) [21], SS3 (2008) [22], SS4 (2009) [23], SS5 (2015) [24], SS6 (2016) [25], SA [12]. Як показують отримані результати, ефективність розробленого алгоритму, як і первісного алгоритму SA, практично не залежить від величини ПСПК, при цьому він залишається ефективним в умовах, на які більшість аналогів взагалі не розрахована: при пропускній спроможності прихованого каналу зв'язку, меншій 0.05 біт/піксель (табл.3). При інший ПСПК значення параметра ACC для розробленого алгоритму SAu порівнянні зі значеннями ACC для розглянутих аналогів.

Таблиця 2 – Помилки першого роду при стеганоаналізі стеганоповідомлень, сформованих LSB-методом, розробленим алгоритмом SAu (%) в умовах обрізки ЦЗ-контейнера

Множина ЦЗ-контейнерів	Пропускна спроможність прихованого каналу зв'язку, побудованого за допомогою LSB-методу (біт/піксель)	
	0.1	0.01
$M_1^{(65)} \cup M_2^{(65)} \cup M_3^{(65)}$	8	8.8
$M_1^{(75)} \cup M_2^{(75)} \cup M_3^{(75)}$	8	9.2
$M_1^{(85)} \cup M_2^{(85)} \cup M_3^{(85)}$	6.4	6.4
Середнє значення	7.5	8.1

Таблиця 3 – Порівняння ефективності, що оцінюється за допомогою ACC , розробленого алгоритму із сучасними стеганоаналітичними алгоритмами виявлення LSB-вкладень в ЦЗ-контейнері в форматі з втратами

ПСПК, (біт/піксель)	SS1 (2006)	SS2 (2006)	SS3 (2008)	SS4 (2009)	SS5 (2015)	SS6 (2016)	SA (2019)	SAu (2019)
0.1	0.9846	0.7727	0.9943	0.9937	0.988	0.970	0.94	0.92
0.05	0.9769	0.6432	0.9283	0.9319	0.968	0.941	0.94	0.92
0.01	0.5692	0.5094	-	-	-	-	0.94	0.91

ВИСНОВКИ

В роботі вирішена важлива науково-практична задача забезпечення можливості виявлення порушення цілісності ЦЗ внаслідок його блокової обробки в умовах, коли місце розташування сітки, що

використовувалася при обробці, є невідомим. В ході роботи удосконалено метод виявлення блокової обробки ЦЗ, запропонований в [12].

В ході тестування алгоритмічної реалізації методу встановлено, що в умовах виявлення наявності стиску ЦЗ з втратами (найпоширеніший вид блокової обробки) її ефективність порівнянна з ефективністю існуючих аналогів, а для метода, запропонованого в [18], значно перевищує.

При оцінці ефективності розробленого методу в якості стеганоаналітичного для виявлення вкладень додаткової інформації LSB-методом в умовах малої пропускної спроможності прихованого каналу зв'язку встановлено, що SA_u є ефективним при малій ПСПК, зокрема при ПСПК=0.01 біт/піксель значення АСС=0.91, що поступається лише алгоритму-прототипу SA, в той час, як більшість існуючих аналогів в таких умовах взагалі не працюють.

ЛІТЕРАТУРНІ ДЖЕРЕЛА

- [1] Redi, J. A. Digital image forensics: a booklet for beginners // *Multimedia Tools and Applications*, 2011. – Vol. 51. – No.1. – P. 133 –162.
- [2] Кобозева, А.А. Анализ информационной безопасности: монография / А.А. Кобозева, В.А. Хорошко. – К.: ГУИКТ, 2009. – 251 с.
- [3] Kobozeva, A.A. General Principles of Integrity Checking of Digital Images and Application for Steganalysis // *Transport and Telecommunication*, 2016. – Vol. 17. – Issue 2. – P. 128 –137.
- [4] Бобок, И.И. Теоретическое развитие общего подхода к проблеме выявления нарушений целостности цифровых контентов, основанного на анализе полного набора формальных параметров // *Информатика та математичні методи в моделюванні*, 2017. – Т.7. – №3. – С. 170 –177.
- [5] Гонсалес, Р. Цифровая обработка изображений; пер. с англ. П.А. Чочиа. – М. : Техносфера, 2006. – 1070 с.
- [6] Bergman, C. Unitary embedding for data hiding with the SVD // *Security, steganography and watermarking of multimedia contents VII, SPIE*, 2005. – Vol.5681. – P.619 – 630.
- [7] Nagarajan, D. Block Processing And Edge Detection For A Dicom Image // *International Journal of Pure and Applied Mathematical Sciences*, 2016. – Vol. 9. – No 1. – P. 9 – 16.
- [8] Мельник, М.А. Стеганоалгоритм, устойчивый к сжатию // *Інформаційна безпека*, 2012. – № 2(8). – С. 99 –106.
- [9] Козіна, М.О. Метод перевірки автентичності інформації, що передається стеганографічним каналом зв'язку // *Вісник Вінницького політехнічного інституту*, 2015. – №1. – С. 99 – 104.
- [10] Стеганография, цифровые водяные знаки и стеганоанализ: [монография]. – М.: Вузовская книга, 2009. – 220 с.

- [11] Бобок, І.І. Дослідження змін властивостей параметрів блоків цифрового зображення при блоковій обробці як основа методу виявлення порушення його цілісності // Правове, нормативне та метрологічне забезпечення системи захисту інформації в Україні, 2019. – №1. – С. 65–74.
- [12] Бобок, І.І. Дослідження змін властивостей параметрів блоків цифрового зображення при блоковій обробці як основа методу виявлення порушення його цілісності // Правове, нормативне та метрологічне забезпечення системи захисту інформації в Україні, 2019. – №2. – С.67–75.
- [13] Гантмахер, Ф.Р. Теория матриц [Текст]: монографія. – 5-е изд. – М. : Физматлит, 2004. – 559 с.
- [14] Костирка О.В. Модифікація стійкого до збурних дій стеганоперетворення просторової області зображення-контейнера // Інформатика та математичні методи в моделюванні, 2016. – №1. – С.85–93.
- [15] Gloe T., Böhme R. (2010). The ‘Dresden Image Database’ for benchmarking digital image forensics. Proceedings of the 25th Symposium on Applied Computing (ACM SAC 2010). Sierre, 2010. –Vol. 2. – P. 1585–1591.
- [16] Hsu, Y.-F. Detecting image splicing using geometry invariants and camera characteristics consistency // Proceedings of the IEEE International Conference on Multimedia and Expo (ICME’06). Toronto, 2006. –P.549–552.
- [17] Geetha, S. Close color pair signature ensemble adaptive threshold based steganalysis for LSB embedding in digital images // Transactions on Data Privacy, 2008. – Vol.1. – Iss.3. – P. 140 – 161.
- [18] Tjoa, S. Block size forensic analysis in digital images // IEEE International Conference on Acoustics, Speech and Signal Processing, 2007. – Vol. 1. – P. I-633 – I-639.
- [19] Luo W., Huang J., Qiu G. (2009) A Novel Method for Block Size Forensics Based on Morphological Operations. In: Kim HJ., Katzenbeisser S., Ho A.T.S. (eds) Digital Watermarking, IWDW 2008. Lecture Notes in Computer Science. – vol 5450. Springer, Berlin, Heidelberg.
- [20] Xiang-dong Chen. Detect LSB Steganography with Bit Plane Randomness Tests // 6-th World Congress on Intelligent Control and Automation, Dalian, 2006. – P. 10306 –10309.
- [21] Zou, D. Steganalysis based on Markov model of thresh-oldeed prediction-error image // 2006 IEEE international conference on multimedia and expo, ICME 2006 – Proceedings, 2006. – P. 1365 –1368.
- [22] Chen C. JPEG Image Steganalysis Utilizing both Intrablock and Interblock Correlations // IEEE International Symposium on Circuits and Systems. Seattle, Washington, USA, 2008. – P. 3029136136 –3032.
- [23] Huang F. Calibration based universal JPEG steganalysis // Science in China series F: Information sciences, 2009. – Vol. 52. – No. 2. – P. 260 –268.
- [24] Xue, B. Steganalysis of LSB replacement for multivariate Gaussian covers // 2015 IEEE China Summit and International Conference on Signal and Information Processing (ChinaSIP), Chengdu, 2015. – P. 836 – 840.

[25] Lin, Q. Local ternary pattern based on path integral for steganalysis // 2016 IEEE International Conference on Image Processing (ICIP), Phoenix, AZ, 2016. – P. 2737 – 2741.

IMPROVEMENT OF METHOD OF DETECTION OF DIGITAL PROCESSING OF DIGITAL IMAGE

Khoroshko V., Kobozeva A., Bobok I.

The integrity of information is one of the criteria for its security. This makes the task of identifying integrity violations of any information content relevant if this content is used for non-entertaining purposes. The work is devoted to the analysis of the integrity of digital images. Currently, image processing is often block-based. In block processing, individual blocks are processed sequentially or in parallel. These blocks are obtained by partitioning the image matrix. Recently, the authors have developed an effective method for identifying block processing. To use the method, information is needed on the location of the grid for partitioning the image matrix during block processing. In practice, this information is most often inaccessible to an expert. This happens, for example, when cropping an image after processing it. The aim of the work is to enable the detection of block image processing when the location of the grid of the matrix splitting is unknown. The goal is achieved by improving the method for identifying block processing, which was developed earlier. The proposed improvement uses a single parameter for analysis. This parameter reflects the change in quantity blocks small size when shifting the mesh, for which the angle between the normalized vector of squares of singular numbers and the singular vector corresponding to the maximum singular number is equal to the angle between the first vector of the standard basis of the corresponding space and the n -optimal vector. The algorithm was tested under lossy image compression conditions, and also as a steganoanalytic method for LSB waveform analysis with low bandwidth of a hidden communication channel. Determined that the developed algorithm is slightly inferior in efficiency to the algorithm taken as a basis, comparable in efficiency with existing analogues.

Keywords. Digital image, integrity violation, image block processing, singular vector, singular number, partition grid.

МАТЕМАТИЧЕСКОЕ ПРЕДСТАВЛЕНИЕ ВЕРОЯТНОСТНОЙ НАДЕЖНОСТИ ЗАЩИТЫ ИНФОРМАЦИИ В ЗАВИСИМОСТИ ОТ НАПРАВЛЕНИЯ ВЗЛОМА

Журиленко Б., Николаева Н.

В работе представлено теоретическое исследование для разработки методологии нового подхода к проектированию, анализу состояния и своевременной модернизации работающей технической защиты информации (ТЗИ) в зависимости от направления происходящего процесса взлома. В работе, теоретически более строго получена функция, определяющая вероятностную надежность проектируемой технической защиты, и ее зависимость от параметров направления процесса взлома. Функция позволяет из параметров процесса взлома вычислить один или другой неизвестный параметр времени или попытки. Например, по известной попытке взлома можно получить возможное время, когда произойдет взлом ТЗИ или, наоборот, по известному времени определить попытку. Поскольку реальный процесс взлома является случайной величиной от попытки и времени взлома, поэтому направление взлома не отображается прямой линией.

В этих условиях функция направления линии взлома будет определяться как среднеквадратичное или среднее значение результатов реальных попыток и его времени. Результаты исследования позволяют в случае проектирования ТЗИ по распределению максимумов вероятности взлома определить максимальные вероятности взлома для любых направлений этого процесса. Получено выражение для распределения вероятности взлома. С помощью этого распределения показана возможность определения вероятности, попытки и времени реального процесса взлома ТЗИ, вероятность взлома которого не будет максимально возможной. В этом случае также можно определять вероятность процесса защиты информации во времени, а также анализировать ее реальное состояние, если злоумышленник поменял направление процесса взлома. Результаты проведенных исследований могут иметь большое значение для разработки новой методологии ТЗИ, учитывающей вложенные финансовые затраты в защиту, эффективность защиты и направление процесса взлома.

Ключевые слова. Техническая защита информации, коэффициент эффективности защиты, распределение максимума вероятности взлома, попытка взлома, время попытки взлома, линия направления взлома.

ВВЕДЕНИЕ

В настоящее время в Украине теоретически исследуются защиты информации, в основном, с использованием системного подхода [1], экспертной оценки анализа с помощью нечетких множеств [2-3] и теории игр [4]. Однако при наличии определенных положительных моментов, связанных с разработкой защиты информации этими методами, все они имеют существенные недостатки. Основной

недостаток этих подходов является качественная оценка защиты и невозможность оценки ее состояния в процессе работы.

Техническая защита информации (ТЗИ) в различных странах осуществляется в соответствии со своими нормативными документами и разрабатываемыми методами. Различные фирмы и предприятия, для которых создается ТЗИ, в первую очередь, интересуется экономическая выгода применения той или иной защиты. Для технических и экономических расчетов, используемых в проектировании технической защиты из всех возможных параметров, понятных для заказчика и разработчика, являются величины рисков полных финансовых потерь, величины рисков потерь вложенного финансирования и вероятности взлома защиты, как в начальной стадии защиты, так и в процессе ее работы. Как и в работе [5], для расчетов финансовых затрат могут использоваться известные начальные финансовые потери без защиты, риски потерь вложенных финансовых затрат на выбранную защиту с данной вероятностью взлома, риски потерь полных финансовых затрат, вероятность взлома и эффективность выбранной защиты.

ПОСТАНОВКА ПРОБЛЕМЫ

В открытой литературе [2-4,6] приводятся методы расчета рисков, финансовых затрат и оценка эффективности защиты информации. Однако, нет конкретных рекомендаций расчетов, которые определялись бы конкретными параметрами такими как: эффективность финансовых расходов на создание ТЗИ, оптимизация финансовых потерь в случае взлома защиты, критериев необходимости дополнительных затрат на восстановление защиты до необходимого технического уровня и, соответственно, оптимизации финансовых потерь. В рассматриваемых работах построенная ТЗИ, в основном, имеет качественную оценку, которая отвечает начальным условиям использования защиты.

С другой стороны для разработчика при проектировании и эксплуатации ТЗИ важно знать возможную вероятность взлома защиты на каждом этапе ее работы в зависимости от направления процесса взлома. Реальное направление взлома можно определить из попыток и времени этих попыток взлома, хотя этот процесс носит случайный характер. Основным требованием для вероятностной надежности должна быть ее зависимость от параметров попыток взлома. Зависимость вероятностной надежности от времени определяется направлением взлома и носит зависимый характер от попытки взлома. Отсутствие попытки взлома не должно изменять вероятностную надежность с изменением времени, и только при наличии попытки взлома вероятностная надежность должна изменяться. В общем случае распределение вероятностной надежности ТЗИ должно определяться попыткой и временем этой попытки взлома. В реальных условиях при

взломе ТЗИ фиксируемыми фактами или параметрами могут быть попытка взлома и ее время. Зная в каждый момент времени вероятность взлома работающей ТЗИ, разработчик может предсказать попытку и время возможного взлома ТЗИ. Направление взлома можно определить всегда, например, по количеству попыток и времени этих попыток взлома [7]. Таким образом, данный подход к построению ТЗИ позволит получить для нее количественную оценку в виде вероятности и оценить состояние в процессе ее работы. Эти результаты дадут возможность разработчику принять решение о замене используемой ТЗИ или ее модернизации, что позволит сэкономить финансовые и материальные ресурсы, вкладываемые в защиту информации.

Из открытых источников неизвестны защиты, которые разрабатывались бы по нормативным документам и которые обеспечивали бы контроль их состояния от количества попыток взлома во времени. С другой стороны, контроль количества попыток и времени взлома позволил бы определить интенсивность и направление взлома. Поскольку направление взлома зависит от двух параметров, то вероятностная надежность также должна зависеть от двух этих параметров и должна быть связана между собой направлением взлома. Существуют публикации Б. Журиленко, в которых сделана попытка разработать методологию построения защиты, позволяющей осуществлять контроль ее состояния в процессе работы, давать рекомендации для модернизации ТЗИ в зависимости от финансовых вложений на защиту, эффективности создаваемой защиты и направлению взлома. Однако в этих работах недостаточно строго была показана вероятностная надежность в зависимости от направления взлома.

ФОРМУЛИРОВАНИЕ ЦЕЛИ ИССЛЕДОВАНИЙ

Актуальность работы заключается в том, чтобы в отличие от нормативных документов разработать новый подход к требованиям ТЗИ, опирающийся на реальные физические процессы взлома информации.

Научная новизна заключается в разработке новой методологии в подходе к проектированию, анализу рабочего состояния работающей ТЗИ с целью экономии финансовых затрат, вкладываемых в защиту.

Целью работы является получение распределения максимума вероятности взлома ТЗИ и распределение вероятности взлома защиты в зависимости от направления взлома, которые определяются двумя параметрами – попыткой и временем этой попытки взлома.

Задача исследования – разработка методологии и способа получения распределений вероятностей взлома ТЗИ с учетом направления взлома.

Объект исследования – является процесс технической защиты информации.

Предмет исследования – распределение вероятностной надежности ТЗИ в зависимости от направления взлома.

Методы исследования – основываются на математическом представлении реального процесса взлома защиты информации.

ИЗЛОЖЕНИЕ ОСНОВНОГО МАТЕРИАЛА ИССЛЕДОВАНИЙ

Теоретическое обоснование распределения вероятностной надежности ТЗИ в зависимости от направления взлома.

Для получения выражения распределения вероятностной надежности ТЗИ от направления взлома примем следующие предположения. Пусть t_0 – некоторый параметр создаваемой ТЗИ, зависящий от направления взлома, и вид которого необходимо определить, t – текущее время, в течение которого осуществляется защита, $p'(t)$ – вероятность защищенности ТЗИ во времени.

Определим риски защищенности ТЗИ во времени, как

$$(t_0 + t) \cdot p'(t) = f(t), \quad (1)$$

где $f(t)$ – произвольная положительная функция, так как левая часть выражения (1) не может быть отрицательной.

Анализируя выражение (1), отметим, что для обеспечения защиты информации функция рисков защищенности $f(t)$ с ростом времени t должна быть хотя бы постоянной. Если функция $f(t)$ со временем будет уменьшаться, то используемая защита является неэффективной и ее необходимо заменить на другую более эффективную защиту. Если $f(t)$ увеличивается со временем, то такая защита в процессе эксплуатации более эффективна, так как риски защищенности увеличиваются со временем.

Таким образом, чтобы иметь более эффективную защиту, выбираем в виде степенного многочлена первого порядка [8] в соответствие с левой частью формулы (1) и требованием независимости вероятностной надежности ТЗИ от времени, когда нет попытки взлома.

$$f(t) = \alpha + \beta \cdot t, \quad (2)$$

Поскольку $f(t)$ произвольная положительная функция, то коэффициенты в выражении (2) должны быть $\alpha \geq 0$, $\beta \geq 0$ при любом значении $t \geq 0$.

Выражение вероятности защищенности из (1) будет иметь вид

$$p'(t) = \frac{f(t)}{(t_0 + t)} \quad (3)$$

Учитывая начальные условия при $t=0$, вероятность защищенности будет полная и $p'(0)=1$. Тогда получим

$$p'(0) = \frac{f(t)}{(t_0)} = 1 \quad ; f(t) = t_0 \quad (4)$$

Следовательно, получим вероятность защищенности ТЗИ в виде

$$p'(t) = \frac{f(t)}{f(t) + t} = \frac{\alpha + \beta \cdot t}{\alpha + \beta \cdot t + t} \quad (5)$$

Определим вероятность взлома во времени

$$p(t) = 1 - p'(t) = \frac{t}{\alpha + \beta \cdot t + t} \quad (6)$$

Считаем независимость вероятности взлома от результатов предыдущих попыток и, если с очередной попытки взлом не произошел, то вероятность взлома используемой защиты остается той же. Такое распределение попыток взлома будет подчиняться геометрическому закону распределения вероятностей [9].

В этом случае вероятность события взлома на m – той попытке может быть записана как

$$P(m, t) = [p'(t)]^{m-1} \cdot p(t) = \left[\frac{\alpha + \beta \cdot t}{\alpha + \beta \cdot t + t} \right]^{m-1} \cdot \left[\frac{t}{\alpha + \beta \cdot t + t} \right] \quad (7)$$

Определим, на какой m – той попытке по времени будет экстремум кривой вероятности взлома $P(m, t)$. Для этого возьмем первую производную выражения (7) по времени и приравняем ее нулю. В результате получим

$$\alpha + \beta \cdot t = (m - 1) \cdot t \quad \text{или} \quad (8)$$

$$f(t) = (m - 1) \cdot t = f(m, t)$$

Поскольку ранее было принято, что $\alpha \geq 0$, $\beta \geq 0$ при любом значении времени $t \geq 0$, то вторая производная по времени будет больше нуля, что соответствует максимуму вероятности взлома (7) при значении функции (8). Сравнивая (8) с (2) получим $\alpha = 0$, $\beta = (m - 1)$. Таким образом, максимумы вероятностей взлома $P(m, t)$ на m – той попытке будут описываться выражением

$$P(m, t) = \left[\frac{f(m, t)}{f(m, t) + t} \right]^{m-1} \cdot \left[\frac{t}{f(m, t) + t} \right] = \quad (9)$$

а поверхность максимумов вероятностей взлома $P(m, t)$ – функцией $f(m, t)$ любых попыток и времени взлома

$$P(m, t) = \left[\frac{f(m, t)}{f(m, t) + t} \right]^{\frac{f(m, t)}{t}} \cdot \left[\frac{t}{f(m, t) + t} \right]. \quad (10)$$

Выражение (9) или (10) должны соответствовать физическому требованию зависимости вероятностной надежности $P(m, t)$ от попыток взлома и независимости от времени, когда нет попыток взлома. Для доказательства этого требования запишем выражение (9) в виде

В результате получим

$$\begin{aligned} \lim_{t \rightarrow \infty} P(m, t) &= \lim_{t \rightarrow \infty} \left\{ \left[\frac{(m-1) \cdot t}{(m-1) \cdot t + t} \right]^{m-1} \cdot \left[\frac{t}{(m-1) \cdot t + t} \right] \right\} = \\ &= \lim_{t \rightarrow \infty} \left[\frac{(m-1)^{(m-1)}}{(m)^{(m)}} \right] = const. \end{aligned} \quad (11)$$

Из выражения (11) видно, что если нет последующей попытки взлома, то независимо от текущего времени (вплоть до бесконечного времени) вероятность взлома остается постоянной в соответствие с предыдущей попыткой. С другой стороны, если попытка возможного взлома стремится к бесконечности, то вероятность взлома будет определяться выражением

$$\lim_{m \rightarrow \infty} P(m, t) = \lim_{m \rightarrow \infty} \left[\frac{(m-1)^{(m-1)}}{(m)^{(m)}} \right] = \lim_{m \rightarrow \infty} \left[\left(\frac{m-1}{m} \right)^{m-1} \cdot \frac{1}{m} \right] = \lim_{m \rightarrow \infty} \left[\frac{1}{e \cdot m} \right] = 0 \quad (12)$$

Таким образом, если попытка взлома происходит на бесконечности, то вероятность взлома будет равна нулю, что соответствует физическому требованию вероятностной надежности и здравому смыслу. Функция $f(m, t)$, присущая данной технической защите, определяет ее защитные свойства и направление взлома. Эта функция также определяет поверхность максимумов вероятностей взлома $P(m, t)$ и зависит от координат точек взлома. Соотношение между координатами m и t точек взлома при постоянных значениях функции $f(m, t)$ представлены на рис.1 линиями 1, 2, 3, 4. С возрастанием номера линии от 1 до 4 значение функции будет возрастать соответственно 1, 10, 20, 40. Линии 5, 6 дают направление взлома, которое определяется двумя точками взлома. В реальных условиях начальная точка взлома может определяться началом координат, то есть $m-l=0$ и $t=0$. Таким образом, пересечение линии постоянства функции (гиперболические линии на рис.1) и линии направления взлома дает значение максимума вероятности взлома в каждой точке пересечения с данной попыткой взлома.

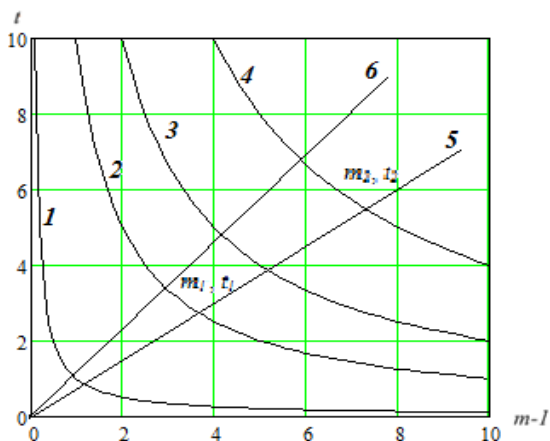


Рис.1– Соотношение между координатами m и t точки взлома при постоянных значениях функции $f(m, t)$ и линии направления взлома. Линия 1 соответствует функции $f(m, t) = 1$, линия 2 - $f(m, t) = 10$, линия 3 - $f(m, t) = 20$, линия 4 - $f(m, t) = 40$, линии – 5, 6 дают направления взлома.

В реальных не начальных условиях каждой определенной попытке взлома соответствуют значения m_1, t_1 и m_2, t_2 . Каждая последующая попытка взлома будет иметь значения $m_2 > m_1, t_2 > t_1$ и, следовательно, согласно выражению (8) значение $f(m, t)$ должно возрасти с увеличением времени или попытки взлома. Направление взлома будет определяться линией между координатами взлома. На рис.1 этот факт представлен прямой линией направления взлома (линия 5) между двумя значениями $f(m_1, t_1)$ и $f(m_2, t_2)$ и координатами m_1, t_1 и m_2, t_2 . Функцию $f(m, t)$ в направлении линии взлома с изменением одной из координат можно представить в виде функции времени

$$f(t) = \left[(m_1 - 1) + \frac{m_2 - m_1}{t_2 - t_1} \cdot (t - t_1) \right] \cdot t, \quad (13)$$

и в виде функции попытки взлома

$$f(m) = \left[t_1 + \frac{t_2 - t_1}{m_2 - m_1} \cdot (m - m_1) \right] \cdot (m - 1) \quad (14)$$

При одной и той же попытке взлома выражения (8), (13) и (14) равны между собой и соответствуют максимуму вероятности взлома в данной точке. Выражение (8), отвечающее за параметры защиты, определяет поверхность максимумов вероятности взлома, которая описывается выражением (10) и, следовательно, описывается двумя координатами m и t точки взлома. Выражение (13) описывает кривую максимумов вероятности взлома от одной координаты времени t точки взлома. Выражение (14) описывает кривую максимумов вероятности взлома от координаты m точки взлома. Таким образом, можем записать

$$f(m, t) = f(t) = f(m) \quad (15)$$

Введем понятие интенсивности или частоты попыток взлома

$$\omega = \frac{m_2 - m_1}{t_2 - t_1} \quad (16)$$

В процессе построения, контроля или модернизации ТЗИ может возникать необходимость по одному из известных параметров m или t определить другой, используя функцию $f(t)$ или $f(m)$ и направление взлома. Это позволит при оценке качества ТЗИ определить возможную попытку и ее время взлома.

Учитывая равенство (15), из выражений (13) и (14) найдем зависимость времени от попытки взлома

$$t(m) = \frac{\sqrt{A^2 + \frac{4}{\omega} \cdot f(m)}}{2} - \frac{A}{2}, \text{ где } A = t_1 + \frac{m_1 - 1}{\omega}, \quad (17)$$

и зависимость попытки взлома от времени

$$m(t) = \frac{\sqrt{B^2 + 4 \cdot \omega \cdot f(t)}}{2} - \frac{B}{2} + 1, \quad (18)$$

где $B = \omega \cdot t_1 - (m_1 - 1)$.

Следует учитывать, что в (18) при первой попытке взлома $m(t) = 1$, то есть соответствует предстоящей реальной попытке взлома, когда реальное начальное время еще равно нулю.

Функция $f(m, t)$ определяет поверхность максимумов вероятности взлома, но не учитывает эффективность защищенности, то есть дает значение максимума вероятности взлома при коэффициенте эффективности защиты (КЭЗ) равном $\gamma = 1$, что соответствует взлому на бесконечной попытке. В реальных условиях взлом происходит на конечной попытке при КЭЗ меньше единицы. В работах [5,7,10] показано, что коэффициент эффективности защиты может быть вычислен через значения вероятностей двух любых известных попыток взлома и представлен в виде

$$\gamma = \frac{\ln P_1 - \ln P_2}{\ln [P(m_1 t_1)] - [P(m_2 t_2)]}, \quad (19)$$

где P_1, P_2 – реальные известные вероятности в первой и второй точках взлома соответственно, $P(m_1, t_1), P(m_2, t_2)$ – расчетные вероятности в первой и второй точках взлома соответственно.

Учитывая КЭЗ в соответствие с [10], выражение (19)

$$P(m, t) = \left\{ \left[\frac{f(m, t)}{f(m, t) + t} \right]^{\frac{f(m, t)}{t}} \cdot \left[\frac{t}{f(m, t) + t} \right] \right\}^\gamma. \quad (20)$$

При проектировании технической защиты информации параметры взлома закладываются разработчиком и соответствуют исходным данным. Для разработчика защиты важно знать

вероятностную надежность ТЗИ в проектируемом направлении взлома и в направлении реального процесса взлома. Чтобы построить проектируемую поверхность с выбранными разработчиком параметрами защиты, в выражениях (9) или (10) необходимо выразить степень через параметры конкретной проектируемой попытки взлома, например, $m = m_c$, $t = t_c$. Тогда (9) будет иметь вид

$$P(m, t) = \left\{ \left[\frac{f(m, t)}{f(m, t) + t} \right]^{m_c - 1} \cdot \left[\frac{t}{f(m, t) + t} \right]^\gamma \right\} \quad (21)$$

а (10)

$$P(m, t) = \left\{ \left[\frac{f(m, t)}{f(m, t) + t} \right]^{\frac{f(m_c, t_c)}{t_c}} \cdot \left[\frac{t}{f(m, t) + t} \right]^\gamma \right\} = \left\{ \left[\frac{f(m, t)}{f(m, t) + t} \right]^{\frac{f(m_c)}{t_c}} \cdot \left[\frac{t}{f(m, t) + t} \right]^\gamma \right\} \quad (22)$$

В выражении (14) заменяем m на m_c попытку взлома и получаем $f(m_c)$. Время t_c время взлома берется из исходных параметров. Если есть проектируемое направление взлома и задано время взлома, то вместо функции $f(m_c)$ используем $f(t_c)$, заменяя t на t_c в выражении (22). На рис.2 построена проектируемая поверхность максимумов вероятности взлома ТЗИ по формуле (20). По этой формуле каждая точка строится по максимуму вероятности взлома ТЗИ с эффективностью защиты $\gamma = 0,7$, выбранной для иллюстрации математического представления вероятностной надежности ТЗИ в направлении взлома. На рисунке представлена поверхность с выбранным направлением взлома по линии 5, а линия 6 соответствует другому, например, реальному направлению взлома. Точки пересечения поверхности с линиями дают координаты максимумов вероятности взлома в данном направлении. Для линии 5 это будут, выбранные разработчиком защиты, исходные координаты $m_m = 9$, $t_m = 6$ с максимумом вероятности взлома в данной точке, а для линии 6 – реального направления взлома в точке $m_m = 12$, $t_m = 11$. На рис.3 представлена поверхность распределения вероятности взлома с максимумом в точке с выбранным направлением взлома по линии 5 ($m_c = 9$, $t_c = 6$). Линия 6 соответствует другому реальному направлению взлома, но на поверхности распределения вероятности взлома, которая рассчитана для проектируемого направления взлома по линии 5. Из рис. 3 видно, что при изменении направления взлома (линия 6) надежность ТЗИ будет меняться и при ее проектировании необходимо это учитывать.

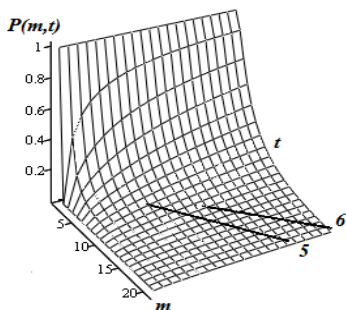


Рис.2 - Поверхность распределения максимумов вероятности взлома

Поверхность распределения максимумов вероятности взлома (рис.2), построенная по формуле (20) с эффективностью защиты $\gamma = 0,7$; 5, 6 – линии взлома, направление которых соответствуют линиям рис.1. Точки пересечения поверхности с линиями дают координаты максимума вероятности взлома в данном направлении.

Поверхность распределения вероятности взлома (рис.3) в точке с выбранным направлением взлома по линии 5 и максимумом в точке $m_c = 9, t_c = 6$. Линия 5 соответствует выбранному направлению, а линия 6 реальному направлению взлома

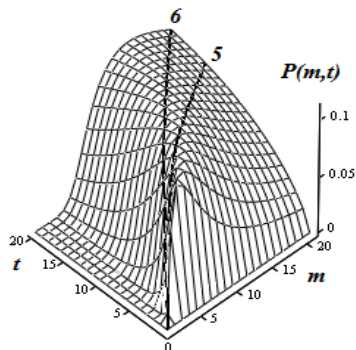


Рис.3 - Поверхность распределения вероятности взлома

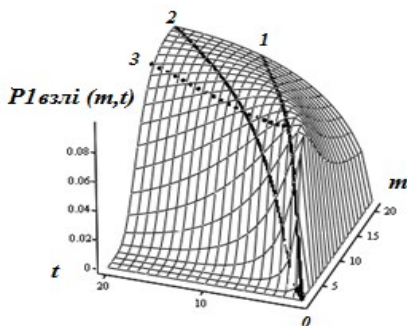
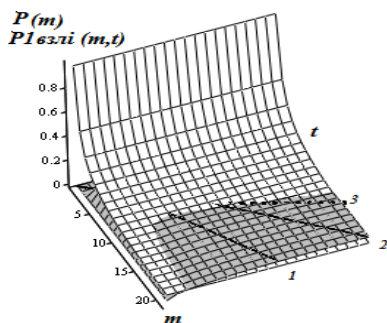


Рис.4 - Распределение вероятности взлома: а – с проектируемым направлением взлома по линии 1(серая поверхность); б – реальная поверхность взлома (белая поверхность), рассчитываемая по формуле $P(m)=1/m$

На рис.4а представлена поверхность с максимумом вероятности взлома в точке с другим проектируемым направлением взлома по линии

1, максимум которого в точке $m_c = 10$, $t_c = 5$. Линия 2 соответствует другому реальному направлению взлома, но на поверхности, определяемой направлением взлома по линии 1. Линия 3 дает направление реального взлома, если злоумышленник изменил реальный процесс нападения. На поверхности по координатам m , t видны максимумы значений вероятностей взлома. Точка пересечения обоих максимумов и направления линии дает точку максимума вероятности взлома в данном направлении. Такая точка на рис.4а представлена пересечением поверхности с линией 1 проектируемого направления процесса взлома с координатами $m_m = 10$, $t_m = 5$, а для линии 2 с $m_m = 12$, $t_m = 11$. Из рис.4б видно, что реальный процесс взлома с поверхностью вероятности взлома описываемой выражением $P(m)=1/m$ (белая поверхность) и поверхностью спроектированной защиты (серая поверхность) будет проходить с вероятностью определяемой линией пересечения белой и серой поверхностей. В данном случае, согласно рис.4б, процесс взлома с расчетным максимальным значением вероятности будет только для направления проектируемой ТЗИ (линия 1), а для остальных направлений значение вероятности взлома будет меньше (линии 2 и 3). Если направление реального процесса взлома близко к проектируемому направлению защиты, то взлом ТЗИ может произойти при значениях близких к проектируемой попытке взлома $m_{взл}$, особенно при небольших увеличениях по времени между попытками взлома. При значительных увеличениях по времени между попытками взлома, взлом защиты может не произойти совсем. Аналогичная ситуация, когда взлом не произойдет, возможна, если попытки взлома будут следовать очень часто друг за другом.

ВЫВОДЫ

Из рисков защищенности ТЗИ получена функция $f(m, t)$, зависящая от направления процесса взлома, которая присуща данной технической защите и определяет вероятностную надежность технической защиты в направлении взлома.

Из функции направления процесса взлома перспективным является выражение, позволяющее по одному из параметров m или t , определять другой. Это важно при проектировании, анализе состояния и модернизации ТЗИ, потому что позволит по одному из известных параметров по направлению взлома найти другой. Например, по известной попытке взлома можно оценить возможное время, когда произойдет взлом защиты.

В данной работе получено распределение вероятности взлома ТЗИ для направления проектируемого процесса взлома, зависящего от параметров попытки, времени этой попытки взлома, финансовых вложений в защиту и коэффициента эффективности защищенности. При

проектировании защиты направление взлома выбирается в виде прямой линии, которая строится по требуемым исходным данным. Следует заметить, что реальный процесс взлома является случайной величиной, как по попыткам, так и времени взлома, и не отображается прямой линией. Для анализа состояния работающей технической защиты направление взлома может определяться по методологии, как это было предложено в работе [10].

По полученным в работе выражениям построена поверхность максимумов вероятности взлома (рис.2), по которой в точках пересечения поверхности и линии определяется наиболее вероятное значение взлома и координаты точки взлома. Построены поверхности распределения вероятности взлома (рис.3, рис.4а) по линиям проектируемых направлений взлома (линии 5, линия 1). Результаты работы позволяют оценить состояние остаточной вероятности надежности работающей ТЗИ реального процесса взлома в выбранных злоумышленником направлениях нападения.

Проведенные исследования в перспективе позволят создать новую методологию проектирования, модернизации и анализа состояния работающего комплекса технической защиты информации, с учетом вложенного в защиту финансирования, эффективности разработанной защиты и выбранного разработчиком направления взлома.

СПИСОК ЛИТЕРАТУРЫ

- [1] Домарев В.В. Безопасность информационных технологий. Системный подход, К.: ООО «ТИД «ДС», 2004. – 992 с.
- [2] Корченко А.Г. Построение систем защиты на нечетких множествах. Теория и практические решения, К.: «МК-Пресс», 2006. – 320 с.
- [3] Архіпов О.Є. Оцінювання якості роботи експертів за даними багатооб'єктної експертизи // Захист інформації: науково-технічний журнал. - К.: НАУ, 2011. –№4 (53). – С. 45 – 54.
- [4] Грищук Р.В. Теоретичні основи моделювання процесів нападу на інформацію методами теорії диференціальних ігор та диференціальних перетворень: Монографія, Житомир : Рута, 2010. – 280 с.
- [5] Журиленко Б.Е.Оценивание финансовых затрат на построение системы защиты информации // Науково-технічний журнал «Захист інформації», 2018. – №4(20). – С. 231–239. DOI: 10.18372/2410-7840.20.13424
- [6] Кравченко В.І. Використання теорії нечітких множин для визначення втрат на захист інформації // Науково-технічний журнал «Захист інформації», 2011. – №1. – С.85 – 90.
- [7] Журиленко Б.Е. Моделирование процесса взлома и анализа рабочего состояния технической защиты информации // Безпека інформації, 2016. – №1(22). – С. 26 – 31.

- [8] Андре Анго Математика для электро- и радиоинженеров, М.: Из-во «Наука», 1964. – 772 с.
- [9] Румшинский Л.З. Элементы теории вероятностей , М.: Изд-во «Наука», Главн. Ред. Физ.-мат. Лит., 1970. – 256 с.
- [10] Журиленко Б.Е.Методология построения и анализа состояния комплекса технической защиты информации с вероятностной надежностью и учетом временных попыток взлома // Научно-технический журнал «Захист інформації», 2015. – №3(17). – С.196 – 204

MATHEMATICAL REPRESENTATION OF PROBABILISTIC RELIABILITY OF INFORMATION PROTECTION DEPENDING ON HACKING DIRECTION

Zhurilenko B., Nikolaeva N.

In this paper, a theoretical study is presented for further development of methodology of a new approach to designing, analyzing the state and timely modernization of working technical information protection (TIP), depending on the direction of the ongoing hacking process. In the work, there was obtained a theoretically more rigorous function that determines the probabilistic reliability of the designed technical protection, and its dependence on the direction parameters of the hacking process. The function allows one or another unknown parameter, either time or attempt, to be calculated from the parameters of the hacking process. For example, using a known hacking attempt, you can get the possible time when a TIP hacking will occur, or, conversely, using a known time to determine the attempt. Since the real process of hacking is a random variable from the attempts and time of hacking, the direction of the hack is not displayed by a straight line. Under these conditions, the direction function of the cracking line will be determined as the mean square or average value of the results of real attempts and its time. The results of the study allow, in the case of designing a TIP to distribute the maximum probabilities of breaking, to determine the maximum probabilities of breaking for any direction of this process. An expression is obtained for the probability distribution of hacking. Using this distribution, the possibility of determining the probability, attempt and time of a real process is shown for hacking TIP with breaking probability different from maximum possible one. In this case, it is also possible to determine the probability of the process of protecting information over time, as well as analyze its real state if the attacker changed the direction of the hacking process. The results of the research can be of great importance for the development of a new TIP methodology that takes into account the financial investment in defense, the effectiveness of the defense and the direction of the fracture process.

Keywords/ Technical protection of information, protection efficiency coefficient, distribution of the maximum probability of hacking, attempted hacking, time of attempted hacking, line of direction for breaking.

ІНТЕЛЕКТУАЛЬНА СИСТЕМА ВИЯВЛЕННЯ АТАК В ЛОКАЛЬНИХ МЕРЕЖАХ

Шпінарева І.

Інструменти виявлення атак мають дуже широкий спектр рішень, від застосування правил, отриманих досвідом, до використання методів машинного навчання, включаючи кілька методів біологічного натхнення. Задачі забезпечення безпеки корпоративних інформаційних систем традиційно вирішується шляхом побудови системи інформаційної безпеки, де одним з найважливіших аспектів безпеки є безпека локальної комп'ютерної мережі.

Станом на сьогоднішній день, кількість нових атак на корпоративні комп'ютерні мережі зростає кожного дня шляхом вторгнення у них за допомогою підключення до інтернету. У зв'язку з чим існуючі системи не завжди здатні виявити наявність атаки і вимагають застосування нових методів і підходів до виявлення атак. В роботі спроектована система виявлення аномальної поведінки мережі за допомогою двох нейронних мереж – перша мережа перевіряє наявність атаки, друга мережа класифікує атаку у разі її наявності.

Ключові слова. Інформаційна безпека, виявлення вторгнення, штучна нейронна мережа, KDD, машинне навчання TCP/IP.

ВСТУП

З самого початку передачі інформації за допомогою комп'ютерних мережевих ресурсів, були розглянуті питання про загрози безпеки інформаційної системи за допомогою різних методів, найбільш часто використовуваних у системах виявлення атак[1].

Аналітичні дані компаній [2], що спеціалізуються в сфері захисту інформації, таких як McAfee, Symantec, Trustware і Kaspersky Labs, показують, що протягом 2018 року стабільно зростала кількість інцидентів, пов'язаних зі спробами порушення безпеки інформаційних систем. Крім цього, спостерігається відносно стабільний ріст кількості нових зразків атакуючих дій. Згідно з міжнародним дослідженням ЕУ в даній області [3], кількість атак різних типів зросла у середньому на дев'ять відсотків у період 2017-2018 рр. Ці дані показують, що існуючі системи виявлення атак не можуть виявляти нові атаки і нові різновиди атак, та вимагають розробку нових методів для покращення результатів виявлення. Сучасні системи виявлення атак повинні виконувати розподілений збір і аналіз інформації, а також інтелектуальний її аналіз.

Станом на сьогоднішній день, існують два класи методів виявлення атак – метод зловживання та метод аномалій. Виявлення атак методами зловживання виконується за допомогою правил, отриманих на основі досвіду або з використанням алгоритмів машинного навчання

відповідно до наборів даних, сигнатур, отриманих з мережевих сценаріїв зі звичайним трафіком, і за допомогою ін'єкційних атак до одного і того ж сценарію. Цей клас методів використовують сучасні системи виявлення атак доповняючи його аналізом мережі у якій вони застосовуються.

Метою роботи є розробка та реалізація системи виявлення вторгнень яка може задовольнити наступні вимоги: система повинна виявляти погрозу застосовуючи методи аномалій; мати низьку ймовірність помилково-позитивних та помилково-негативних результатів; вміти виявляти нові різновидності однієї атаки.

ОГЛЯД МЕТОДІВ АНОМАЛІЙ

А. Аналіз методів аномалій

Методи аномалій засновані на використанні інформації про "нормальну" поведінку системи та її порівнянні з параметрами спостережуваної поведінки. Вони орієнтовані на побудову моделі штатного, або нормального, функціонування системи або користувача. До таких методів відносять:

1) *поведінкові методи*: Вейвлет-аналіз, статистичний аналіз, аналіз ентропії, спектральний аналіз, фронтальний аналіз та кластерний аналіз.

Переваги цих методів полягають у можливості визначити розподілені атаки, в тому числі і в часі, визначити взаємозв'язок між різними подіями, а також кореляція подій дозволяє визначити значущі події серед досліджуваних.

До недоліків цих методів відносять проблему чутливості методів, яка залежить від заданої величини відхилень, точності моделі штучної мережі та ці методи дуже сильно залежать від функції подібності [4].

2) *методи машинного навчання*: Дерево рішень, Баєсова мережа, MAP-сплайни, алгоритми кластеризації та алгоритми регресії.

До переваги цих методів відносять легкість реалізації, низькі вимоги до підготовки даних, малий час обробки великої кількості даних і використовують модель білого ящика, що допомагає перевірити модель з використанням статистичних тестів.

Недоліками цих методів є проблема навчання, що вимагає ретельного підбору коефіцієнтів та проблема побудови оптимальних моделей.

3) *методи обчислювального інтелекту*: Нейронні мережі, генетичні алгоритми, нечітка логіка, імунні системи, метод опорних векторів та ройові алгоритми.

Перевагами цих алгоритмів є можливість бистої обробки великої розмірності даних, також у паралельному режимі, легкі для реалізації, добре працюють навіть на поганому наборі даних та здатні вирішувати нелінійні задачі [5].

Недоліками методів є потрібність як у позитивних, так і негативних прикладах, вони вимагають багато пам'яті і процесорного часу, потребують виборі оптимальної функції ядра або активації.

Враховуючи ці особливості, було вирішено використовувати нейронні мережі типу – багатошаровий перцептрон Румельхарта, та навчати їх алгоритмом зворотного поширення помилки.

БАЗА ДАНИХ ДЛЯ АНАЛІЗУ АТАК

NSL-KDD – це набір даних, запропонований Тавалае та ін.[6]. Набір даних NSL-KDD є скороченою версією оригінального набору даних KDD 99. Він складається з тих самих функцій, що і KDD 99, 41 змінна та одного атрибуту класу. Атрибут класу має 21 значення, що потрапляють під чотири типи нападу: Probe attack, User to Root (U2R), Remote to Local (R2L) та атаки Denial of Service (DoS). Цей набір даних має атрибут дворядкового класу.

Таблиця І. Класи і приклади типів атак

Клас атак	Тип атак
DoS	Back, Land, Neptune, Pod, Smurf, Teardrop, Apache2, Udpstorm, Processtable, Worm
Probe	Satan, Ipsweep, Nmap, Portsweep, Mscan, Saint
R2L	Guess_Password, Ftp_write, Imap, Phf, Multihop, Warezmaster, Warezclient, Spy, Xlock, Xsnoop, Snmptguess, Snmptgetattack, Httptunnel, Sendmail, Named
U2R	Buffer_overflow, Loadmodule, Rootkit, Perl, Sqlattack, Xterm, Ps

Нижче представлені короткі описи атак:

– DOS. Відмова в обслуговуванні є атакою, яка виснажує ресурси жертви, тим самим роблячи її нездатною обробляти нормальні запити.

– Probing. Мета атаки полягає в віддаленому спостереженні і зборі інформації про жертву.

– U2R. Даний тип атак, використовує доступ до існуючого облікового запису користувача в комп'ютерній системі жертви, отриманий, наприклад, методами соціальної інженерії, щоб, використовуючи вразливості, отримати доступ до облікового запису суперкористувача.

– R2L. Неавторизований доступ з віддаленого комп'ютера дозволяє зловмисникові увійти в віддалену машину і отримати локальний доступ до комп'ютера.

Кожна з представлених атак, по суті своїй, є цілим класом, кожен з яких включає в себе безліч конкретних прикладів (табл. 1).

Крім того, в NSL-KDD є достатня кількість навчальних та тестових прикладів, що робить можливим проведення експериментів.

А. Опис бази даних виявлення вторгнень

Записи з NSL-KDD, що використовуються для навчання та тестування, складаються з двох основних компонентів: перший з них являє собою набір функцій що описують подію, а другий це клас, який інформує тип події, тобто якщо це нормальна або ненормальна поведінка, всі ці зразки організовані у файлі CSV.

В. Атаки без контенту

У деяких атаках надсилається інформація вище 4-го рівня моделі OSI, щоб скористатися вразливістю програми, тому довжина корисного навантаження четвертого рівня відрізняється від нуля. Таки типи нападів зазвичай називають "атаками на основі контенту" [7].

З іншого боку, є деякі атаки, яким не потрібно надсилати інформацію на рівень сеансу або вище, це означає, що корисний набір можливого пакету четвертого рівня порожній, ці напади будуть називатися атаками без контенту.

Оскільки аналізується трафік локальної комп'ютерної мережі, отримати інформацію про зміст атаки, а точніше, зміст переданого пакета по мережі немає можливості, в слідстві чого, потрібно проаналізувати існуючий набір атак і вибрати характеристики, які не пов'язані із вмістом атаки. З 41 доступних змінних, 13 на основі контенту, це означає, що такі дані не є необхідними для виявлення атак, не пов'язаних з контентом.

Найважливішою особливістю при виявленні атак без контенту є прапор статусу з'єднання, тобто 4-а особливість в наборі даних. Найважливіші можливі стани прапору статусу наведені в табл. 2.

Історичні події можуть допомогти в процесі виявлення, бо деякі атаки зазвичай приймають різні кроки, щоб бути успішними. Аналізуючи кроки, можна дізнатись, що в мережевому трафіку трапляється щось не так, наприклад, в атаці з портів, можна знайти запити, які надходять до не запропонованої служби, викликаючи з'єднання з режимом REJ, це не обов'язково означає, що присутня атака на порти portsweep, а можливо, що хтось намагається отримати інформацію про послуги, що пропонуються в мережі.

Таблиця 2. Флаг стану з'єднання

Стан	Значення
SF	Нормальне SYN/FIN завершення
REJ	З'єднання відхилено, початковий SYN викликав відповідь RST
S0	Стан 0: відправив SYN, відповіді нема
S1	Стан 1: з'єднання встановлено (обмінялись SYN), більше ніякої активності
S2	Стан 2: з'єднання встановлено, ініціатор закрив свою сторону
S3	Стан 3: з'єднання встановлено, відповідач закрив свою сторону
RSTO RSTR OTH	Ініціатор оновлює з'єднання Відповідач оновлює з'єднання Друге, цей стан не підтримується.

Через характер нападів на комп'ютерну мережу були обрани усі характеристики, зокрема характеристик 10-22, ці характеристики відносяться до атак з контентом. Описи кожної характеристики можна знайти в [8].

МОДУЛІ СИСТЕМИ

Створена система здатна збирати інформацію з мережевого інтерфейсу. Архітектура системи представлена на рис.1.

Модуль сенсору складається з застосування утиліти Tshark [9]. На момент початку роботи, система починає процес Tshark, та кожні 30 секунд зберігає трафік мережі в обраній папці. Збір даних здійснюється за протоколами TCP, UDP і ICMP. При запуску цього модулю, запускається робота модуля аналізатора пакетів, який інтерпретує збережені пакети у форматі XML, після чого збирає інформацію про пакети та передає аналізатору з'єднань. Аналізатор з'єднань отримує пакети, та в залежності від прапорів, встановлених в пакеті, вирішує чи встановити нове з'єднання, оновити інформацію вже існуючого або треба видалити з'єднання. Модуль аналізатору з'єднань відстежує стан з'єднання по описаним прапорам в табл. 2, при установленні яких, передає інформацію про з'єднання модулю збору статистики, а також, слідкує за часом з'єднання, та виконує роботу двох секундного вікна. Якщо з'єднання було оновлено понад дві секунди тому, дане з'єднання перевіряється на атаку у разі необхідності, після чого видаляється зі списку з'єднань. Збірник статистики при отриманні з'єднання збирає статистичні дані за останні дві секунди та утворює вектор з 28-ми змінними для аналізу на наявність атаки.

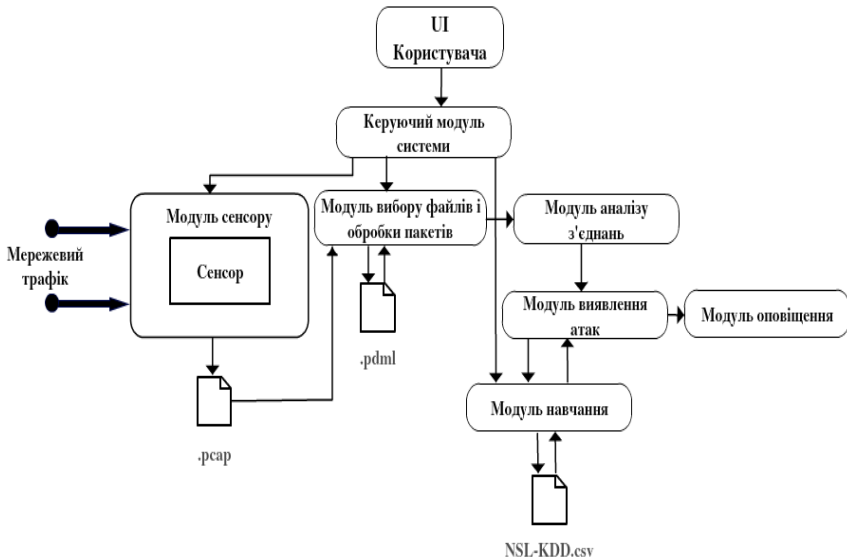


Рис. 1 – Архітектура системи виявлення вторгнень

У якості вихідних даних модуля очікується вектор в форматі NSL-KDD, що складається з наступних полів:

1) Основні характеристики мережесвих з'єднань

Duration – Тривалість з'єднання

Protocol_type – Протокол, що використовується при з'єднанні

Service – Цільовий сервіс, що використовується

Flag – Статус з'єднання: нормальне, помилка

Src_bytes – Кількість байтів, переданих від джерела до приймача за одне з'єднання

Dst_bytes – Кількість байтів, переданих від приймача до джерела за одне з'єднання

Land – Якщо джерело та приймач має однакові номери портів, то параметр набуває значення 1, якщо не однакові – 0

Wrong_fragment – Загальна кількість пошкоджених фрагментів у конкретному з'єднанні

Urgent – Кількість термінових пакетів у конкретному з'єднанні. Терміновий пакет – це пакет, в якому активований біт терміновості URG

2) Характеристики пов'язані з часом мережевого з'єднання

Count – Кількість під'єднань до цільового хоста протягом часового інтервалу в 2с.

Srv_count – Кількість під'єднань до поточної служби (номеру порта) за останні 2с.

Error_rate – Відсоток з'єднань з помилкою типу SYN для даного хоста джерела

Srv_error_rate – Відсоток з'єднань з помилкою типу SYN для даної служби джерела

Error_rate – Відсоток з'єднань з помилкою типу REJ для даного хоста джерел

Srv_error_rate – Відсоток з'єднань з помилкою типу REJ для даної служби джерела

Same_srv_rate – Відсоток з'єднань зі службою

Diff_srv_rate – Відсоток з'єднань з різними службами

Srv_diff_host_rate – Відсоток з'єднань з різними хостами

3)Характеристики пов'язані з хостом мережевого з'єднання

Dst_host_count – Кількість з'єднань з хостом

Dst_host_srv_count – Кількість з'єднань зі службою

Dst_host_same_srv_rate – Відсоток з'єднань з даною службою на даному хості

Dst_host_diff_srv_rate – Відсоток з'єднань з різними службами на даному хості

Dst_host_same_src_port_rate – Відсоток з'єднання з даним хостом при поточному номері порта джерела

Dst_host_srv_diff_host_rate – Відсоток з'єднань зі службою різних хостів

Dst_host_error_rate – Відсоток з'єднань з помилкою типу SYN для даного хоста приймача

Dst_host_srv_error_rate – Відсоток з'єднань з помилкою типу SYN для даної служби приймача

Dst_host_rej_rate – Відсоток з'єднань з помилкою типу REJ для даного хоста приймача

Dst_host_srv_rej_rate – Відсоток з'єднань з помилкою типу REJ для даної служби приймача

Виявлення «аномальної» поведінки системи здійснюється модулем аналізу атаки. Модуль аналізу атаки містить у собі каскад з двох нейронних мереж. Ідея каскаду НМ полягає в розподілі головної задачі системи виявлення атак на кілька малих задач, які вирішуються різними нейронними мережами. Головною задачею системи являється виявлення атак і їх класифікація. Перша мережа аналізує наявність атаки, друга мережа класифікує тип атаки. Але, оскільки перша задача вирішує питання наявності атаки, а нейронні мережі мають недолік помилкового спрацювання, то друга нейронна мережа крім класифікації відомих атак буде перевіряти результати першої нейронної мережі на наявність атаки або на розпізнавання невідомих атак. Цей варіант також дає подвійну перевірку від помилково-позитивних результатів системи. Модуль нейронної мережі являє собою конструктор

для створення багатошарової нейронної мережі – багатошаровий перцептрон. Нейронні мережі мають три шари. Перший – шар вхідних параметрів, який має розмір 28, другий – схований шар, який має розмір 84 та третій – вихідний шар, який має розмір 2 або 5, відповідно до типу мережі. Система розпізнає такі класи атак: DOS (Denial of service), Probing, R2L (Remote to local attack), U2R (User to root). Система працює у двох режимах, перший режим – режим налаштування, який навчає кожну нейронну мережу, другий режим – аналіз трафіку у реальному часі.

ТЕСТУВАННЯ

А. Метрика для оцінки якості

Для порівняння продуктивності і ефективності методів виявлення вторгнень в мережі використовуються такі метрики [10]: ймовірності істинно-позитивних (True Positive Rate, TPR) та помилково-позитивних результатів (False Positive Rate, FPR). FPR є ймовірністю отримати оповіщення у разі перевірки нормальної поведінки. До того, ймовірність помилково-негативного рішення (False Negative Rate, FNR) означає, що система не видає оповіщення у разі наявності зловмисної поведінки. Рівняння (1) та (2) відображують FPR та FNR.

$$FPR = \frac{\text{number of false positive}}{\text{number of negative}} \quad (1)$$

$$FNR = \frac{\text{number of false negative}}{\text{number of positive}} \quad (2)$$

Звідси, ймовірність TPR та істинно-негативних результатів (True Negative Rate, TNR) дорівнюють:

$$TPR = 1 - FNR \quad \text{та} \quad TNR = 1 - FPR \quad (3)$$

Також, до цих показників можна додати ймовірність правильної класифікації (True Rate, TR) і ймовірність помилкової класифікації (False Rate, FR).

$$TR = \frac{\text{number of true positive} + \text{number of true negative}}{\text{number of positives} + \text{number of negatives}} \quad (4)$$

$$FR = 1 - TR.$$

Рівняння (5) відображує міру чутливості (sensitivity), вона визначається як частка нормальної поведінки:

$$\text{Sensitivity} = \frac{\text{number of true positives}}{\text{number of true positives} + \text{number of false negatives}} \quad (5)$$

Але, ця міра недостатньо змістовна, оскільки вона може бути тривіально досягнута шляхом класифікації всієї поведінки як зловмисної. Ще одна метрика, міра специфіки (specificity). Це частка справжніх негативів всієї розглянутої негативної поведінки:

$$\text{Specificity} = \frac{\text{number of true negatives}}{\text{number of true negatives} + \text{number of false positives}} \quad (6)$$

При класифікації всього трафіку як нормального, міра специфіки досягається повністю. F-міра є показником, що поєднає міри чутливості та специфіки:

$$F - \text{measure} = \frac{2 \times \text{sensitivity} \times \text{specificity}}{\text{sensitivity} + \text{specificity}} \quad (7)$$

A. Тестування різних нейронних мереж

Тестування і вибір розміру, функції активації та показника навчання проводились у ОС Windows® 10-64 з процесором Intel Core i7 2600k (3.4 GHz), 12,0 Гб ОЗУ. Реалізація нейронних мереж була створена за допомогою Microsoft CNTK[11], та система була реалізована на мові C#. Навчання НМ здійснюється методом зворотного поширення помилки. Точність розбіжності очікуваного і отриманого результатів визначається автоматично завдяки бібліотеці CNTK, яка виконує ітеративний прохід засновуючись на отриманій кількості ітерацій 300 і мінібатчами по 64 вектора. Система виявлення атак була навчена на виборці NSL-KDD на 133,504 векторах та протестована на 15040. Сама вибірка векторів мала 95% варіації, та кожен вектор унікальний, тобто навчальна та вибірка для тестування не мають однакових векторів.

Каскад складається з двох нейронних мереж. Перша нейронна мережа, повинна встановлювати, чи є наявність атаки, друга повинна встановлювати клас атаки. Тобто, перша нейронна мережа на виході у якості результату надає вектор з двох значень – є атака чи ні, друга нейронна мережа у якості відповіді надає вектор з п'ятьма значеннями – Probe, R2L, U2R, DoS та «не є атакою». Перші значення відповідають класам атак, а остання застосовується для подвійної перевірки, якщо перша нейронна мережа помилково вирішила що зв'язок небезпечний.

Для ефективного знаходження мережевих атак необхідно вибрати структури нейронних мереж, використаних у каскаді. Для вибору структури нейронних мереж були протестовані п'ять функцій активації з різними коефіцієнтами навчання та кількості шарів. Таблиця III відповідає трьом шаровим нейронним мережам з коефіцієнтом навчання 0.02. Перший шар відповідає 28 змінним які передає система, другий шар містить кількість нейронів кратну трьом(84), або п'яти(140), та останній шар містить два, якщо це перша нейронна мережа та п'ять, якщо друга. Кількість нейронів у другому шарі було обраний емпіричним шляхом у

ході тестування. Кількість шарів нейронних мереж дорівнює трьом, при спробі збільшення їх кількості, показники точності значно падали.

Перша нейронна мережа показала найкращий результат за показниками Tr, Fr, TPR, FNR, Recall та F-measure з функцією активації ReLU та середнім шаром розміром 84 (табл.3). Цю нейронну мережу було обрано для починання каскаду.

Таблиця 3. Оцінка якості першої нейронної мережи (28/84/2) з різними функціями активації та $\eta=0,02$

	Tr	Fr	TPR	TNR	FPR	FNR	Recall	Precision	F-measure
Tanh	95.47	4.53	97.44	93.39	2.56	6.61	93.39	97.44	95.37
ReLU	96.30	3.70	97.34	95.21	2.66	4.79	95.21	97.34	96.26
Sigmoid	95.35	4.65	97.68	92.87	2.32	7.13	92.87	97.68	95.22
ELU	95.70	4.30	96.93	94.40	3.07	5.60	94.40	96.93	95.65
SELU	95.04	4.96	96.97	92.99	3.03	7.01	92.99	96.97	94.94

Друга нейронна мережа показала більш розподілений результат. Нейронна мережа з функцією активації ReLU та з середнім шаром розміром 140 видає кращий результат за показниками TPR(96,38%), FPR(3,6%), Recall (95%) та F-measure(96,32%), а нейронна мережа з середнім шаром розміром 84 видає кращий результат за показниками Tr та Fr. Оскільки головними для нас є показники Tr та Fr, у якості другої нейронної мережі була обрана мережа з функцією активації ReLU, середнім шаром розміром 84 та коефіцієнтом навчання $\eta=0.02$ (табл.4).

Таблиця 4. Оцінка якості другої нейронної мережи (28/84/5) з різними функціями активації та $\eta=0,02$

	Tr	Fr	TPR	TNR	FPR	FNR	Recall	Precision	F-measure
Tanh	95.94	4.06	98.35	93.36	1.65	6.64	93.36	98.35	95.79
ReLU	96.02	3.98	98.10	93.81	1.90	6.19	93.81	98.10	95.90
Sigmoid	95.60	4.40	98.66	92.32	1.34	7.68	92.32	98.66	95.39
ELU	95.85	4.15	98.51	93.01	1.49	6.99	93.01	98.51	95.68
SELU	95.54	4.46	97.75	93.18	2.25	6.82	93.18	97.75	95.41

Також кожна нейронна мережа була протестована з коефіцієнтом навчання 0.04 або 0.08 та з більшою кількістю шарів. При використанні трьох шарів та коефіцієнту навчання $\eta=0.04$, показники Tr та Fr мають гірший результат у середньому на 0.3 відсотка, а при збільшенні коефіцієнту навчання цей показник значно зростає, проте, при збільшенні кількості шарів та використанні коефіцієнту навчання 0.02, показники Tr та Fr відрізняються вже на 10,3 відсотки у середньому.

Таблиця 5 демонструє порівняння між обраними нейронними мережами першого і другого типу окремо, та створений з них каскад, ReLU_1 – нейронна мережа першого типу, відповідно ReLU_2 –

нейронна мережа другого типу. Як можна бачити з таблиці, каскад з цих мереж показав кращий результат за показниками Tr, Fr, TNR, FNR та Precision. Це показує, що створений каскад в загалі краще працює, а також має нижчу ймовірність прийняти безпечний трафік за небезпечний. Покращити показники TPR та FPR можна при підвищенні показників першої нейронної мережі.

Таблиця 5. Порівняння якості каскаду

	Tr	Fr	TPR	TNR	FPR	FNR	Recall	Precision	F-measure
ReLU 1	96.38	3.62	97.65	94.96	2.35	5.03	95.03	97.65	96.32
ReLU 2	96.31	3.69	97.49	94.46	2.51	5.53	95.05	97.49	96.25
Cascade	96.91	3.09	98.69	96.78	1.31	3.21	96.78	98.7	97.73

ВИСНОВКИ

Перспективними напрямками при проектуванні систем виявлення атак в даний час бачиться використання методів машинного навчання і гібридизація підходів, яка дозволила б поєднувати в собі переваги сигнатурних і евристичних методів, а також використання технологій великих даних і проактивного моніторингу безпеки. Одним з основних вимог, що пред'являються до цих рішень, є забезпечення адаптивної і високо масштабованої аналітичної обробки подій, що забезпечує інтелектуальне управління великими обсягами даних про безпеку в реальному або близькому до реального масштабі часу

Метою роботи було створити систему аналізу інтернет трафіку для виявлення атаки на локальну комп'ютерну мережу. В роботі для виявлення інформативних ознак атак були розглянуті нейронні мережі прямого поширення з різними функціями активації та кількості шарів, на підставі яких був створен каскад нейронних мереж у якості нового засобу аналізу даних трафіку. Запропонований каскад показав кращий результат – 96.91 відсотків, ніж використання кожної нейронної мережі окремо, та була створена система аналізу мережевого трафіку для виявлення атак.

Система була протестована на локальній мережі що містить 14 комп'ютерів. На локальну мережу була імітована DoS атака у розмірі 26,102 пакетів, 93.41% з яких було коректно ідентифіковано як атаку.

ЛІТЕРАТУРНІ ДЖЕРЕЛА

- [1] Nikishova A., Churilina A. Distributed intrusion into information system of enterprise detection Izvestiya UFU, Technical science, 2013. – №12 (149).
- [2] Information Security Threats, www.anti-malware.ru/threats/information-security-threat.
- [3] Garuba, M.; Chunmei Liu; Fraites, D.; "Intrusion Techniques: Comparative Study of Network Intrusion Detection Systems," Information Technology: New Generations, 2008. ITNG 2008. Fifth International

- Conference on. – no. 7-9. – P.592 – 598, DOI: 10.1109/ITNG.2008.231.
- [4] A. A. Branitsky, I. V. Kotenko, “Analysis and classification of methods for detecting network attacks”, Tr. SPIIRAN, 45 (2016). – P. 207 – 244.
- [5] I.M. Shpinareva «The comparative characteristic of the intrusion detection systems on basis of neural networks»/ Book of abstracts of the International Scientific Conference “Computer Algebra and Information Technology” / 20-26 August 2012, Odessa.- P. 86 – 89.
- [6] Tavallae, M.; Bagheri, E.; Wei Lu; and Ghorbani, A.A detailed analysis of the KDD CUP99 data set. Proceedings of the 2009 IEEE Symposium on Computational Intelligence in Security and Defense Applications (CISDA 2009). – P.1– 6.
- [7] Pasupulati, A.; Coit, j.; Levitt, K; Wu, S.F.; Li, S.H.;Kuo, J.C.; Fan, KP.; "Buttercup: on network-based detection of polymorphic buffer overflow vulnerabilities," Network Operations and Management Symposium, 2004. NOMS 2004. IEEE/IFIP. – Vol.1. – P.235 – 248.
- [8] Dhanabal, L. and Dr. S. P. Shantharajah. “A Study on NSL-KDD Dataset for Intrusion Detection System Based on Classification Algorithms.” 2015. Tshark [Электронный ресурс] – Режим доступа: <http://www.wireshark.org/docs/man-pages/tshark.html>.
- [9] Holz T. 13 security measurements and metrics for networks, Dependability Metrics, 2008. – P. 157 – 165. DOI: 10.1007/978-3-540-68947-8_13
- [10] CNTK [Электронный ресурс] – Режим доступа: <https://www.microsoft.com/en-us/cognitive-toolkit>
- [11] Lukatsky A.V. Obnaruzhenie atak [Attack Detection]. SPb.: BHV-Petersburg, 2003. – 608 p.

INTELLECTUAL SYSTEM FOR DETECTION ATTACK IN LOCAL NETWORKS

Spinareva I.

Attack detection tools have a very wide range of solutions, from applying the rules of experience to using machine learning methods, including several biological inspiration methods. Corporate information systems security has traditionally been addressed by building an information security system where one of the most important aspects of security is the security of the local area network. As of today, the number of new attacks on corporate computer networks is increasing every day by invading them through an internet connection. Therefore, existing systems may not always be capable of detecting an attack and require the use of new methods and approaches to detecting attacks. The system designed to detect anomalous behavior of the network using two neural networks - the first network checks for an attack, the second network classifies the attack if it exists.

Keywords. Information Security, Intrusion Detection, Artificial Neural Network, KDD, TCP / IP Machine Learning. Correction

СТРУКТУРА ЗАЩИТЫ ИНФОРМАЦИИ В МЕЖДУНАРОДНОЙ ГИДРОГРАФИЧЕСКОЙ ОРГАНИЗАЦИИ

Михайлов С., Шевцов Ю.

Публикация «S-63 IHO Data Protection Scheme» («Схема защиты данных Международной гидрографической организации»), далее называемая «схемой», описывает рекомендуемый стандарт защиты информации электронной навигационной карты (ЭНК). Он определяет конструкции безопасности и рабочие процедуры, которые должны соблюдаться для обеспечения правильной работы схемы защиты данных и предоставления спецификаций, позволяющих участникам создавать системы, совместимые с S-63, и распространять данные безопасным и коммерчески жизнеспособным способом.

Стандарт «Защита информации электронных картографических навигационных систем» был принят в качестве официального стандарта МГО государствами-членами МГО в декабре 2002 года (IHO CL 66, 2002). Он определяет роли и обязанности по защите данных электронной навигационной карты (ЭНК), созданных Национальными гидрографическими службами, и распространяется среди клиентов с системами ECS/ ECDIS (Система отображения электронных карт и информации).

Ключевые слова. Стандарт защиты информации, электронная навигационная карта, гидрографическая служба.

ОБЩЕЕ ОПИСАНИЕ

В этом документе описывается метод защиты информации электронных навигационных карт (ЭНК) и поддержания целостности службы ЭНК с несколькими службами данных, обслуживающими большую клиентскую базу. Защита данных имеет три раздела: 1. Защита от пиратства: для предотвращения несанкционированного использования данных путем шифрования информации электронной навигационной карты (ЭНК); 2. Селективный доступ: для ограничения доступа к информации электронных навигационных карт (ЭНК) только тех клиентов, которые поставлены на учет; 3. Аутентификация: обеспечить уверенность в том, что данные электронной навигационной карты (ЭНК) поступают из официальных источников.

Защита от пиратства и выборочный доступ достигаются за счет шифрования информации электронных навигационных карт (ЭНК) и предоставления разрешений ячеек для их расшифровки. Серверы данных будут шифровать данные электронных навигационных карт (ЭНК), предоставленные странами-производителями, прежде чем отправлять их клиенту данных. Зашифрованная электронная навигационная карта (ЭНК) затем дешифруется ECS / ECDIS до того, как будет перематрирована и импортирована в системную электронную

навигационную карту (SENC). Аутентификация обеспечивается посредством цифровых подписей в данных. В схеме конкретно не рассматривается способ защиты информации электронной навигационной карты (ЭНК) или системной электронной навигационной карты (SENC), когда она находится в приложении конечного пользователя. Это ответственность производителя оригинального оборудования. Схема позволяет для массовой рассылки зашифрованных электронных навигационных карт (ЭНК) на жестком носителе (например, CD-ROM или DVD) и может быть использована для всех клиентов с действующей лицензией, содержащей набор разрешений. Избирательный доступ к отдельным ячейкам поддерживается за счет предоставления пользователям с лицензированным пакетом разрешающих документов, содержащий зашифрованные ключи. Эта лицензия создана с использованием уникального идентификатора оборудования системы и является уникальным для каждого клиента данных. Следовательно, лицензии не могут быть обменены между индивидуальными клиентами информации. В схеме используется алгоритм сжатия для уменьшения размера набора данных. Незашифрованные данные электронных навигационных карт содержат много повторяющихся шаблонов информации, например, координатная информация. Поэтому сжатие всегда применяется до того, как информация электронных навигационных карт зашифрована и не сжата после дешифрования в клиентской системе данных (как правило, электронная навигационная карта (ECS) или система отображения электронных карт и информации (ECDIS)). Существует несколько типов пользователей схемы: администратор схемы (AC), который всегда один; сервер данных (СД), которых может быть много; клиент данных (КД), которых может быть много; производитель оригинального оборудования, которых может быть много. Более подробное объяснение этих терминов приведено ниже.

Администратор схемы.

Администратор несет полную ответственность за поддержание и координацию схемы. Роль администратора схемы реализуется Международным гидрографическим бюро (ИГБ) в качестве секретариата МГО от имени государств-членов МГО. Администратор схемы отвечает за контроль членства в схеме и обеспечение того, чтобы все участники работали в соответствии с определенными процедурами. Администратор схемы сохраняет цифровой сертификат верхнего уровня, используемый для работы со «Схемой защиты данных», и является единственным органом, который может подтвердить личность других участников схемы. Администратор схемы также является хранителем всей документации, относящейся к схеме защиты данных.

Серверы данных.

Серверы данных отвечают за шифрование и подпись данных электронной навигационной карты в соответствии с процедурами и процессами, определенными в схеме. Серверы данных выдают лицензии электронной навигационной карты т.е. разрешения, чтобы клиенты данных с действительными разрешениями пользователей могли расшифровать данные электронной навигационной карты. Серверы данных будут использовать информацию уникального идентификационного ключа (M_KEY) и уникального идентификатора (HW_ID), предоставленную администратору схемы, для выдачи зашифрованных ключей ячейки электронной навигационной карты для каждой конкретной установки. Несмотря на то, что ключи ячейки, используемые для шифрования каждой ячейки, идентичны, они будут зашифрованы с использованием уникального идентификатора (HW_ID) и поэтому не могут быть переданы между другими системами отображения электронных карт и информации (ECDIS) от того же производителя. Примерами серверов данных являются гидрографические офисы, реселлеры с добавленной стоимостью и региональным координационным центром электронной навигационной карты.

Клиенты данных.

Клиенты данных являются конечными пользователями информации электронной навигационной карты и получают защищенную информацию от серверов данных. Программное обеспечение клиента данных отвечает за аутентификацию цифровых подписей электронной навигационной карты, дешифрование информации электронной навигационной карты в соответствии с процедурами, определенными в схеме. Эта схема не мешает агентам или дистрибьюторам предоставлять услуги передачи данных своим клиентам.

Оригинальный производитель оборудования.

Производители оборудования и программного обеспечения должны создавать программное приложение и самостоятельно проверять и проверять его в соответствии с условиями, предусмотренными администратором схемы. Стандарт S-63 содержит тестовые данные для проверки и проверки приложений. Администратор схемы предоставит успешным OEM-производителям собственный уникальный ключ производителя и идентификацию (M_KEY и M_ID). Изготовитель должен обеспечить безопасный механизм в своих программных системах для уникальной идентификации каждой установки конечного пользователя. Схема требует, чтобы каждая установка имела уникальный идентификатор оборудования (HW_ID). Программное приложение сможет расшифровывать ключи ячейки с помощью уникального

идентификатора оборудования (HW_ID), хранящегося в аппаратных или программных средствах, подключенных или запрограммированных в приложении, чтобы впоследствии расшифровать и открыть данные электронной навигационной карты. Затем значение циклической проверки избыточности(CRC), содержащееся в электронной навигационной карте (ENC), может быть проверено для установления целостности базовых данных.

Взаимодействие участников схемы.

Администратор схемы, который может быть только один, аутентифицирует личность других участников схемы. Все серверы данных и системные производители должны обратиться к администратору схемы, чтобы стать участниками схемы и при приеме им назначается уникальной для них конфиденциальной информацией. Клиенты данных являются клиентами серверов данных и OEM-производителей, где серверы данных предоставляют услуги передачи данных и OEM-производители оборудование для дешифрования и отображения этих сервисов.

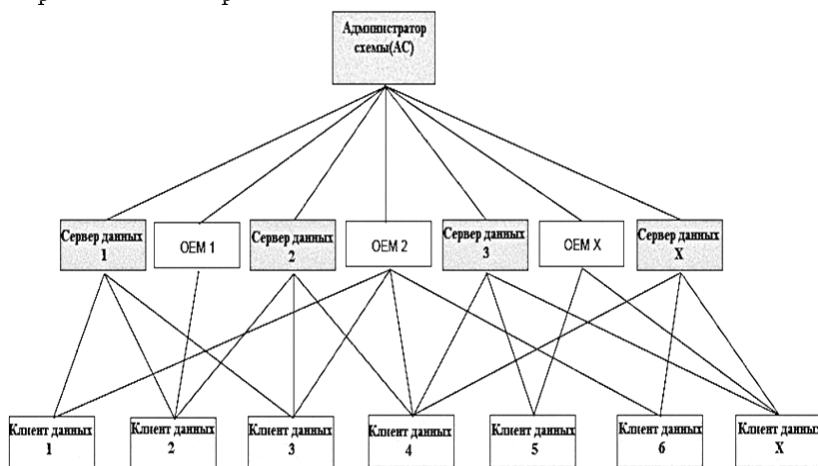


Рис.1– Совместимость с предыдущими версиями

Версия 1.1 стандарта была подготовлена с учетом опыта, накопленного серверами данных и производителями ECS / ECDIS во время работы схемы согласно версии 1.0. Эта версия пытается более четко определить стандарт, удалив дублирование и возможную неопределенность. Она также содержит дополнительные механизмы, которые позволяют производителям сделать их системы более интуитивными для пользователей ECS / ECDIS. Перечень изменений в

стандарте. 1. Удаление ненужного дублирования. 2. Спецификация того, как и при каких условиях должны использоваться определенные файлы.

3. Удаление зависимости получения доступа разрешения от редакции ячейки. 4. Дополнительная информация, позволяющая клиентам данных управлять данными ENC более эффективно и эффективно. 5. Идентификация стратегии загрузки для обеспечения более эффективной загрузки зашифрованных ЭНК.

Структура документа

Основная часть документа обычно может быть разбита на четыре части. В первой части подробно описываются компоненты, которые являются фундаментальными для схемы и описывают их назначение и конструкцию. Второй определяет, как все компоненты объединяются в набор обмена электронной навигационной карты(ENC) S-63. Третий описывает роли и обязанности каждого типа пользователей, участвующих в схеме. Наконец, есть раздел, который определяет различные сообщения об ошибках и предупреждения, которые должны отображаться на клиенте данных при возникновении определенных условий.

ОСНОВНОЙ ДОКУМЕНТ.

1.СЖАТИЕ ДАННЫХ

1.1 Обзор

Файл электронной навигационной карты(ENC) из-за своей структуры содержит повторяющиеся шаблоны информации. Примерами этого являются последовательная нумерация идентификатора объекта (FOID) или небольшие изменения в координатной информации в файле электронной навигационной карты (ENC). Поэтому данные электронной навигационной карты (ENC) хорошо реагируют на сжатие с уменьшением в размере от 30% до 60%, что значительно снижает стоимость передачи данных электронной навигационной карты (ENC) в конечный пункт назначения. Сжаты только файлы электронной навигационной карты ENC (база и обновление). Файлы ENC всегда сжимаются до того, как они зашифрованы, поскольку эффективность любого алгоритма сжатия основана на существовании структурированных данных.

1.2 Алгоритм сжатия

В схеме безопасности используется алгоритм ZIP1 для сжатия и распаковки данных электронной навигационной карты(ENC). Он идентичен алгоритму, используемому во многих коммерческих приложениях, например. WinZip, PKZIP. Потенциальные серверы данных и OEM-производители должны знать, что в прошлом произошли ошибки, когда серверы данных сжимают данные и интерпретируются популярными реализациями алгоритма ZIP как «текстовые» данные.

Если данные не сжаты с неправильными параметрами, это может привести к повреждению файла электронной навигационной карты(ENC), ведущего к ошибкам проверки целостности. Серверу данных и OEM-производителям рекомендуется тщательно выполнять (архивирование) сжатие / несжатие в своих системах.

1.3 Сжатые файлы

Схема безопасности сжимает только базовую ячейку электронной навигационной карты(ENC) и файлы обновлений. Никакие другие файлы в наборе обмена S-57 не будут сжаты.

2.ШИФРОВАНИЕ И ОБРОБОТКА ДАННЫХ

2.1 Какие данные зашифрованы?

В рамках Схемы используется только один алгоритм шифрования. Только данные в файлах электронной навигационной карты (ENC) база или обновленная ячейка внутри набора обмена S-57, то есть текстовых или графических файлов, остаются незашифрованными. Схема шифрует полное содержимое базы данных электронной навигационной карты (ENC) или обновляет файлы данных. Другая информация в зашифрованной схеме включает в себя системный идентификатор (HW_ID) OEM-системы, который зашифрован и предоставлен клиенту данных в форме пользовательского ввода. Клавиши соты, используемые для шифрования файлов данных электронной навигационной карты (ENC), сами зашифровываются сервером данных и предоставляются клиентам данных в качестве разрешений сотовой сети.

2.2. Алгоритм шифрования

Каждый отдельный файл ячейки электронной навигационной карты(ENC) шифруется с использованием уникального ключа ячейки. Один и тот же ключ ячейки используется для шифрования всех обновлений, выпущенных для ячейки. Тем не менее, эта схема позволяет увеличивать и изменять ключи ячеек по усмотрению сервера данных. Ключи соты доставляются клиентам данных в виде разрешений на ячейки. Информация электронной навигационной карты ENC(базовые ячейки и обновления) шифруется с использованием 40-битного ключа.

Пользовательское разрешение и содержимое разрешений для ячеек шифруются с использованием 48-битного ключа. Схема шифрует всю информацию, используя алгоритм Blowfish. Алгоритм непатентован и доступен в общественном достоянии (www.counterpane.com). Blowfish - это алгоритм блочного шифрования, который работает с 64-разрядными (8 байт) данными. Это требуется, чтобы источники данных были дополнены, если они не кратны 8 байтам. Схема защиты использует алгоритм заполнения «DESinCBCMode», когда любые источники данных должны быть дополнены. Это соответствует режиму ЭКК (электронная кодовая книга) DES.

3. ЛИЦЕНЗИРОВАНИЕ ДАННЫХ

3.1 Введение

Лицензирование – это метод, которым серверы данных пользуются, чтобы предоставить выборку клиентов, данных современным элементам электронной навигационной карты (ENC) в течение данного периода времени. Чтобы взаимно действовать со схемой фактически должны быть средства, где установки клиента данных могут открыть кодируемые элементы электронной навигационной карты (ENC). Чтобы открыть данные, установка клиента данных должна иметь доступ к кнопочным переключателям с самовозвратом, которые используются, чтобы кодировать элементы электронной навигационной карты (ENC).

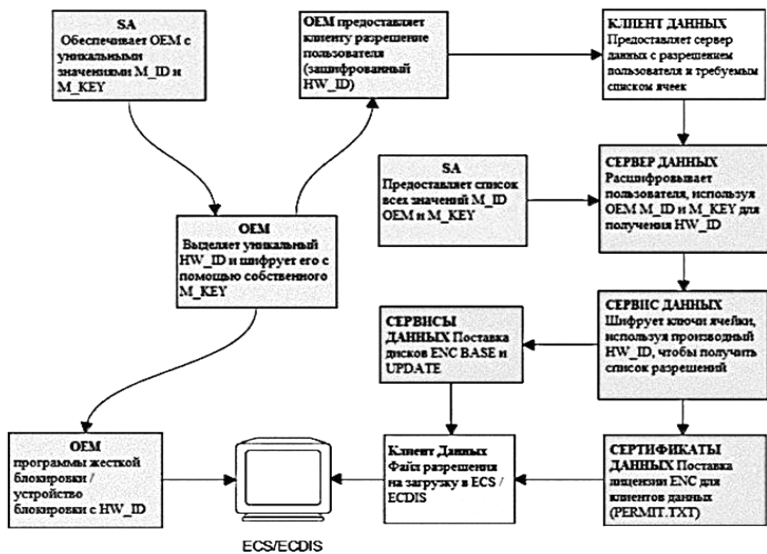


Рис.2 – Схема лицензирования ENC высокого уровня

Эти кнопочные переключатели с самовозвратом поставляются клиенту, чтобы кодировать, в файле разрешения. Чтобы сделать каждый набор, клеточные разрешения эксклюзивны, клеточные кнопочные переключатели с самовозвратом должны кодироваться, набор и комбинация, этот уникальный для каждого клиента данных. Оригинальные производители оборудования назначают уникальное имя, уникальный идентификатор (Hw_id) каждой из их установок и обеспечивают кодируемую копию в форме пользовательского разрешения, к каждому клиенту данных. Уникальный идентификатор (Hw_id), сохраненный в пользовательском разрешении кодирует оригинальных производителей оборудования, кодируют уникальный

идентификатор (Hw_id) с их собственным уникальным кнопочным переключателем изготовителя (M_key) с самовозвратом таким образом, что уникальный идентификатор (Hw_id) не может копировать другой изготовитель. Серверы данных имеют доступ к оригинальным производителям оборудования уникальным кнопочным переключателям производителя (M_keys) и могут, поэтому расшифровать уникальный идентификатор (Hw_id), сохраненный в пользовательском разрешении.

Серверы данных кодируют свои кнопочные переключатели с самовозвратом с изготовителями уникального идентификатора (Hw_id), производя набор клеточных разрешений.

3.2 Пользовательское разрешение.

Пользовательское разрешение создает оригинальное производительское оборудования и поставляется клиентам данных как часть их оборудования таким образом, что они могут получить необходимый доступ к кодируемым электронным навигационным картам (Encs) от серверов данных. Следующая секция определяет композицию и формат пользовательского разрешения. Все клиенты данных с установками, способными к пользованию данными, защищенными со схемой S-63, должны иметь уникальную аппаратную идентификацию, уникальный идентификатор (Hw_id), что вместе определяет потребителя.

Такой уникальный программный идентификатор (Hw_id) часто реализуется как механизм защиты от несанкционированного копирования или другими средствами, гарантирующими уникальную идентификацию для каждого оборудования. Уникальный идентификатор (Hw_id) неизвестен клиенту данных, но оригинальное производительское оборудование обеспечит пользовательское разрешение, которое кодируемый вариант уникального идентификатора (Hw_id) и уникальный для данного клиента данных. Пользовательское разрешение создает взятие заданного уникального идентификатора (Hw_id) и кодируя это с кнопочным переключателем изготовителя (M_key) с самовозвратом. Циклическая проверка избыточности (CRC)32 алгоритм есть бегут на кодируемом уникальном идентификаторе (Hw_id) и результат присоединил к этому. Наконец изготовитель прилагает их заданный идентификатор изготовителя (M_id) до конца получающейся строки. Уникальный идентификационный ключ (M_key) и уникальный идентификатор (M_id) поставляет администратор схемы и они уникальны для каждого изготовителя. Клиент данных получает доступ к S-63 кодируемые электронные навигационные карты (Encs), сообщая это пользовательское разрешение серверу данных, кто может затем выдать специфику клеточных разрешений. Поскольку пользовательское разрешение содержит уникальный идентификатор (M_id), это может быть использовано серверами данных, чтобы идентифицировать уникальный идентификационный ключ (M_key), чтобы затем

использовать его для расшифровки данных. Список изготовителя уникального идентификационного ключа (M_key) и уникальный идентификатор (M_id) выпущен и обновлен администратором схемы ко всем серверам данных, подписывающимся на схему. Этот список будет обновлен периодически как новые оригинальные производительские оборудования будут присоединяться к схеме.

3.2.1 Определение разрешения пользователя.

Разрешение пользователя состоит из 28 символов и должно быть записано в виде текста американского стандартного кода для обмена информацией (ASCII) со следующим обязательным форматом и длиной поля:

Зашифрованный HW_ID	Проверка суммы (CRC)	Идентификатор производителя M_ID
16 символов	8 символов	4 символа

Любой алфавитный символ может быть записан в верхнем регистре.

Пример: Пользовательское разрешение

:73871727080876A07E450C043031

Первые 16 символов - Зашифрованный HW_ID

Следующие 8 символов-Проверка суммы (CRC)

Последние 4 символа-Идентификатор производителя M_ID

3.2.2 Формат HW_ID.

HW_ID - это пятнадцатичное шестнадцатеричное число, определенное оригинальным производителем оборудования. Такой уникальный идентификатор (HW_ID) может быть реализован как ключ или другими средствами, обеспечивающими уникальную идентификацию каждой установки. Уникальный идентификатор (HW_ID) должен храниться в системе безопасным образом. Оригинальный производитель оборудования(ОЕМ)должен назначить уникальный идентификатор (HW_ID)для каждой установки. Рекомендуется, чтобы уникальный идентификатор(HW_ID)не были последовательными. Он будет храниться в зашифрованном виде пользовательского разрешения. Уникальный идентификатор (HW_ID) зашифрован с использованием алгоритма шифрования (Blowfish) с уникальным идентификационным ключом (M_KEY) в качестве ключа, приводящего к шестнадцатеричному числу (8 байтов). Зашифрованный идентификатор (HW_ID) затем представляется в форме американского стандартного кода для обмена информацией (ASCII) в пользовательском разрешении как 16 символов. Пример HW_ID: A79AB. Пример зашифрованного HW_ID: 73871727080876A0

3.2.3. Формат суммы (CRC).

Сумма проверки - это шестнадцатеричное число из восьми символов. Оно генерируется за счет зашифрованного уникального

идентификатора (HW_ID) и преобразования его в шестнадцатеричный вид. Затем он хешируется с использованием алгоритма CRC32 и байта, преобразованного в шестнадцатеричную строку с 8-ми символами. Контрольная сумма не зашифрована и позволяет проверять целостность разрешения пользователя. Контрольная сумма в приведенном выше примере: 7E450C04

3.2.4 Формат M_ID.

M_ID - представляет собой двухсимвольный буквенно-цифровой код, выраженный как представление американского стандартного кода для обмена информацией (ASCII), предоставляемое администратором схемы. Администратор схемы предоставит всем лицензированным производителям собственную уникальную комбинацию ключей и идентификаторов производителя (M_KEY и M_ID). Изготовитель должен защитить эту информацию. Администратор схемы предоставит всем лицензированным серверам данных полный список всех заводских кодов, когда новые производители подписываются на эту схему. Эта информация используется сервером данных для определения того, какой ключ (M_KEY) используется для дешифрования уникального идентификатора (HW_ID) в пользовательском разрешении во время создания разрешений на ячейку клиента данных. Уникальный идентификатор (M_ID) в приведенном выше примере: 01 или 3031 (ASCII).

3.2.4 Формат M_KEY.

M_KEY представляет собой пятизначное шестнадцатеричное число, предоставленное администратором схемы (SA). Оригинальный производитель оборудования использует этот ключ для шифрования назначенного идентификатора (HW_ID) при создании разрешений пользователя. Оригинальный производитель оборудования должен хранить его надежно. Этот ключ используется сервером данных для дешифрования назначенных идентификаторов (HW_ID). Пример уникального идентификационного ключа (M_KEY) - 123AB или 3132334142 (ASCII)

3.3 Разрешение ячейки.

Чтобы расшифровать ячейку электронной навигационной карты, клиент данных должен иметь доступ к ключу шифрования, используемого для его шифрования. Поскольку ключи шифрования известны только серверу данных, должно быть средство доставки этой информации клиентам данных защищенным образом. Эта информация предоставляется сервером данных, например, региональный координационный центр электронной навигационной карты (RENC) клиенту данных в зашифрованном виде, известном как разрешение ячейки. Для доставки разрешения ячейки предоставляется один файл и называется PERMIT.TXT. Этот файл может содержать несколько

разрешений ячеек на основе покрытия электронной навигационной карты (ENC), требуемого клиентом данных. Файл PERMIT.TXT будет доставлен либо на жестком носителе, либо с использованием онлайн-сервисов в соответствии с рабочими процедурами сервера данных. Эти процедуры будут доступны клиентам данных при покупке лицензии. Каждая запись разрешения ячейки также содержит дополнительные поля, которые предоставляются, чтобы помочь оригинальным производителям оборудования- системам управлять лицензией клиента данных и разрешать файлы с нескольких серверов данных. Клиенты данных могут получить лицензию на доступ к электронной навигационной карте (ENC), предоставив серверу данных свой уникальный пользовательский доступ. Затем серверы данных могут извлекать уникальный идентификатор (HW_ID) из пользовательского разрешения, используя уникальный идентификационный ключ (M_KEY) клиента данных и создавать разрешения для конкретных клиентов на основе этого значения. Поскольку разрешения на ячейки выдаются для определенного уникального идентификатора (HW_ID), они не могут быть переданы между установками (системами клиента данных). Этот способ связывания разрешения с установкой поддерживает создание зашифрованных компакт-дисков, которые распространены среди всех клиентов данных, подписываемых на услугу. Система клиента данных расшифровывает разрешение ячейки, используя назначенный уникальный идентификатор (HW_ID), надежно хранимый с помощью аппаратных или программных средств. Затем расшифрованные ключи ячеек могут использоваться системой для дешифрования ячейки электронной навигационной карты. Т.к. несколько серверов данных создают файлы разрешений для электронной навигационной карты в своей службе, ответственность за управление файлами разрешений от нескольких серверов данных лежит на системе клиента данных.

3.3.1. Файл разрешений (PERMIT.TXT).

Разрешающая способность ячеек всегда предоставляется в файле PERMIT.TXT, имя файла всегда указывается в верхнем регистре (UPPERCASE), равно как и любые алфавитные символы, содержащиеся в файле. Файл полностью закодирован в стандартном коде для (ASCII):

Раздел	Описание
:Заголовок	Сюда относится дата создания файла и версия формата.
: Разрешения ENC	ENC (официальные) с сервера данных перечислены в этом разделе.
: Разрешения ECS	ECS (неофициальные) с сервера данных могут быть перечислены в этом разделе.

Сервер данных предоставляет информацию о том, как файлы разрешений будут доступны, будь то на жестких носителях или онлайн-сервисах.

3.3.2 Форматы файлов разрешений – заголовков.

В следующей таблице определяется содержимое, и формат каждого заголовка раздела в файле разрешения.

Раздел	Имя поля	Значение
Дата и время	:Дата	Имя поля, дата и время разделяются символом пробела (SP <h20>). Дата будет указана как YYYYMMDD и время как HH:MM с использованием 24-часовых часов. Пример:: ДАТА 20050809 11:11
Версия мета-разрешения	:Версия	Целое число в диапазоне от 1 до 99. Оно будет увеличено на 1 для каждой новой версии спецификации формата файла разрешения. S-63 Edition 1.1 определяет значение как «2», то есть: ВЕРСИЯ 2
Тип разрешения: сотовой ячейки	:ENC	Поле содержит определение разрешений, доступных в лицензии на распространение ENC с сервера данных. В верхнем регистре поле обозначается следующим знаком: ENC
Тип разрешения: сотовой ячейки	:ENS	Поле содержит определение мета-разрешений, доступных в лицензии на распространение ECS от DataServer. Поле обозначается следующим ярлыком в верхнем регистре: ECS

Пример:

: DATE 20080809 11:11

: ВЕРСИЯ 2

: ENC [Список разрешенных разрешений ячеек для официальных ENC]

: ECS [Список разрешений для разрешенных ячеек для других векторных продуктов]

3.3.3. Разрешенные поля записи.

Запись разрешения сотовой ячейки состоит из разделенных запятыми полей:

Поле	Значение
Сотовая ячейка	Определено в разделах 4.3.4 и 4.3.5
Индикатор уровня обслуживания	0 для разрешения на подписку 1 для разрешения на покупку
Номер издания [необязательно]	Номер выпуска DSID-EDTN ячейки ENC (только для серверов данных)
Идентификатор сервера данных	Это двухзначный буквенно-цифровой номер, выданный SA
Комментарий	Свободное текстовое поле для комментариев по разрешению ячейки и т.д.

Примечание:

Поле «Номер издания» больше не является обязательным требованием в S-63, версия 1.1. OEM-производители, реализующие

версию 1.1, больше не должны создавать зависимости в своих системах, которая проверяет взаимосвязь между номером издания ENC и ключом ячейки, используемым для его шифрования. Клиенты данных должны проверять, есть ли в строке разрешения действительный ключ ячейки. Серверы продолжают поддерживать файлы версии PERMIT.TXT версии 1.0 до тех пор, пока не будет определено, что он больше не требуется.

3.3.4 Определение разрешения ячейки.

В таблице указаны поля, содержащиеся в разрешении сотовой ячейки, с определением цели каждого из них.

Поле	Цель
Имя ячейки	имя ячейки позволяет системам базы данных связывать правильный ключ шифрования с соответствующим зашифрованным файлом ячейки ENC.
Дата истечения срока действия	это срок, когда истекает срок действия лицензии базы данных. Системы должны предотвращать установку новых элементов ENC, новых выпусков или обновлений, созданных после этой даты.
Зашифрованный ключ ячейки 1 (ECK1)	ECK1 содержит ключ дешифрования для текущей версии ячейки ENC.
Зашифрованный ключ ячейки 2 (ECK2)	ECK2 содержит ключ дешифрования, который будет использоваться, когда следующий ключ ячейки будет повторен. Будущий ключ содержится в разрешении соты, чтобы позволить серверам данных периодически изменять Ключ соты без одновременного выдачи новых разрешений ячеек всем Клиентам данных.

3.3.5 Формат разрешений ячеек

Разрешение ячеек должно быть записано в виде текста ASCII со следующим обязательным форматом и длиной поля:

Поле	Символы	Формат
Имя ячейки	8	Буквенно-цифровую строку, следующую за соглашением, определенным в S-57 Раздел 3.1 Приложение В, раздел 5.6 для имен ячеек, исключая расширение имени файла. Пример: NO4D0613
Срок годности	8	Числовая строка, которая содержит дату истечения срока действия лицензии для каждой ENC в формате YYUUMMDD. Пример: 20000830 (30 августа 2000 г.)
ECK1 и ECK2	16	Ключи ячейки - это 5 байтовых случайных чисел - их шестнадцатеричные представления шифруются с помощью Blowfish, а затем выражаются в шестнадцатеричном виде в разрешении. Примечание. Алгоритм шифрования blowfish приведет к тому, что зашифрованные данные будут дополнены длиной до 8 байтов. Это означает, что зашифрованные сотовые ключи на самом деле имеют длину 8 байтов, хотя они незашифрованы, они всего 5 байтов (10 шестнадцатеричных символов). Пример: ECK1: BEB9BFE3C7C6CE68 ECK2: B16411FD09F6982

Контрольная сумма разрешения ENC	16	Содержит зашифрованную контрольную сумму для Разрешения сотовой ячейки. Он зашифрован с использованием алгоритма Blowfish с конкретным HW_ID клиента данных и является 8-байтным числом. Эта контрольная сумма зашифровывается в отличие от незашифрованной контрольной суммы разрешения пользователя. например Контрольная сумма ENC в приведенном ниже примере: 795C77B204F54D48
----------------------------------	----	--

ВЫВОДЫ.

Описываемая структура взаимодействия участников МГО позволит существенно повысить уровень информационной безопасности конфиденциальных данных участников и обеспечить защиту информации в электронных картографических навигационных системах.

СПИСОК ЛИТЕРАТУРЫ

- [1] Издание S57 3.1: Стандарт передачи ИО для цифровых гидрографических данных. Международное гидрографическое бюро. – 114 с.
- [2] Стандарт цифровой подписи (DSS), Публикация федеральных стандартов обработки информации (FIPSPub) 186. – 130 с.
- [3] Информационные технологии. Взаимосвязь открытых систем. Каталог. Аутентификация. X.509 версия 3 – Международный союз электросвязи. – 64 с.

INFORMATION PROTECTION STRUCTURE IN THE INTERNATIONAL HYDROGRAPHIC ORGANIZATION

Mikhailov S., Shevtsov Yu.

The publication “S-63 IHO Data Protection Scheme”, later referred to as “the scheme”, describes the recommended standard for the protection of ENC information. It defines security constructs and operating procedures that must be followed to ensure that the data protection scheme is operated correctly and to provide specifications that allow participants to build S-63 compliant systems and distribute data in a secure and commercially viable manner.

The Standard was adopted as the official IHO standard, by the IHO member states in December 2002 (IHO CL 66, 2002). It defines the roles and responsibilities for protecting ENC data produced by National Hydrographic Offices and distributed to customers with ECS / ECDIS systems.

Keywords. Information Security Standard, electronic navigation chart, hydrographic service

ГИБРИДНЫЕ МОДЕЛИ ПРИНЯТИЯ РЕШЕНИЙ С ИСПОЛЬЗОВАНИЕМ ИММУННОГО ПОДХОДА

Кораблев Н., Фомичев А., Соловьев Д., Чуприна А.

Современные системы поддержки принятия решений (СППР) характеризуются обработкой больших объемов информации в условиях неопределенности, что требует применения эффективных методов и моделей, использующих различные интеллектуальные технологии параллельной обработки информации. В статье предлагаются гибридные модели принятия решений (МПР), представленные как на основе нейронной сети (НС), так и в виде правил нечеткого вывода, обучение и адаптация которых осуществляются на высокопроизводительных системах с использованием иммунных моделей в зависимости от изменения характеристик окружающей среды и свойств объекта принятия решений (ОПР). Процесс обучения и адаптации НС состоит в настройке и коррекции параметров сети, а также числа нейронов в скрытых слоях и связей между ними. Процесс обучения и адаптации моделей нечеткого вывода (МНВ) состоит в настройке формы и параметров функций принадлежности, а также параметров и структуры модели (числа правил вывода).

Ключевые слова. Принятия решений, гибридная модель, обучение, адаптация, нейронная сеть, нечеткий вывод, искусственные иммунные системы.

ВВЕДЕНИЕ

В настоящее время актуальной задачей является разработка методов, моделей и информационных технологий принятия решений для управления сложными объектами. В условиях неопределенности при принятии решений, а также изменения свойств окружающей среды и состояния ОПР, учесть которые априори невозможно, необходимо использовать интеллектуальные подходы к созданию МПР. Существуют различные подходы к созданию таких моделей, которые используют аппарат теории нечетких множеств, искусственные нейронные сети, генетические алгоритмы, искусственные иммунные системы (ИИС), байесовские сети, экспертные системы и др. [1-9]. Перспективным является разработка гибридных МПР, использующих преимущества каждой из технологий искусственного интеллекта.

Среди МПР особое место занимают модели на основе НС и модели нечеткого вывода (МНВ), которые широко применяются в условиях неопределенности, в случае нелинейности объекта принятия

решений (ОПР), а также в системах, основанных на использовании экспертной информации [10-12]. Основные трудности при использовании таких моделей для решения задачи принятия решений связаны с априорным определением их компонентов. При наличии набора неструктурированных данных определение этих компонентов представляет собой сложную проблему. Кроме того, одной из задач при формировании МПР является устранение их избыточности. Поэтому специфика задачи принятия решений приводит к необходимости устранения этих недостатков.

Как правило, СППР функционируют в условиях априорной и текущей неопределенности. Поэтому необходимо использовать МПР, которые будут эволюционировать в условиях изменения внешней среды и свойств ОПР, т.е. адаптироваться под эти изменения [1]. Параметры адаптивных моделей должны настраиваться путем их оптимизации в соответствии с некоторым критерием, формируемым по данным из обучающей выборки. Решение подобной задачи оптимизации представляет собой трудную задачу по ряду причин. Во-первых, это большая размерность вектора параметров модели. Во-вторых, большая размерность пространства поиска тесно связана с проблемой многоэкстремальности целевой функции. Кроме того, характер целевой функции полностью зависит от данных обучающей выборки, которые меняются с каждым новым моделируемым ОПР.

Для обучения и адаптации МПР может использоваться эволюционный подход. В [10] рассмотрены эволюционные методы формирования нечетких баз правил и их редукции на основе генетического алгоритма. Однако ГА не всегда эффективны при решении задачи мультимодальной оптимизации. В [11] исследованы нейросетевые и гибридные алгоритмы обучения для повышения эффективности МПР в сложных средах. Развитие эволюционных вычислений ведется по пути создания новых методов, использующих возможности динамического изменения параметров алгоритмов оптимизации и различные варианты распараллеливания процессов обработки нечеткой информации. Таким требованиям удовлетворяют ИИС [13]. В [14] рассмотрено применение ИИС в нечетких системах классификации, в частности для генерации базы нечетких правил. В [15] предложена гибридная МПР на основе нейронной сети, обучение и эволюция которой осуществляются с использованием иммунных моделей клонального отбора и иммунной сети. В [16] предложена МПР на основе нечеткой логики, эволюция которой осуществляется на основе иммунного подхода.

В данной работе рассматриваются гибридные МПР, представленные в виде НС и МНВ, обучение и адаптация которых осуществляется на основе использования ИИС. Это позволяет в режиме реального времени корректировать как структуру, так и параметры МПР

в зависимости от изменения характеристик окружающей среды и свойств ОПР.

ПОСТАНОВКА ЗАДАЧИ

Пусть X – множество наборов признаков (факторов) описания ОПР, Y – множество состояний (классов) ОПР, а функционирование МПР описывается как отображение пространства факторов X на пространство состояний Y в виде соотношения:

$$F : X \rightarrow Y, X \subset R^m, Y \subset R^n. \quad (1)$$

МПР должна иметь аналитическую зависимость соотношения (1), согласно которой по известным значениям вектора входных факторов $X^m = \{x_1, x_2, \dots, x_m\} \subset X$ можно оценить значения выходного вектора состояний (классов) $Y^n = \{y_1, y_2, \dots, y_n\} \subset Y$.

Для обучения и адаптации МПР строится целевая функция:

$$\Phi(P) = \sqrt{\frac{1}{n \cdot m} \sum_{i=1}^n \sum_{j=1}^m [F_i(x_j, P) - y_i]^2} = \min_P, \quad (2)$$

минимизирующая среднеквадратическое отклонение между фактическими значениями выходной переменной y_i и полученной на основе МПР оценкой $F_i(x_j, P)$. Здесь x_j – входные переменные, P – параметры модели, состав которых определяется исходя из выбранного способа описания зависимости (1) и алгоритма обучения.

Ставится задача разработки алгоритмов обучения и адаптации МПР на основе ИИС. С этой целью по данным из обучающей выборки необходимо оптимизировать параметры и структуру МПР в соответствии с критерием (2).

НЕЙРОСЕТЕВАЯ МОДЕЛЬ ПРИНЯТИЯ РЕШЕНИЙ С ИММУННЫМ ОБУЧЕНИЕМ

1. Описание МПР в виде нейронной сети

Для описания связи между входными и выходными переменными в (1) используется МПР в виде многослойной НС. Используется трехслойный персептрон, имеющий входной слой, содержащий число нейронов, равное числу входных признаков, один промежуточный слой и выходной слой, содержащий число нейронов, равное числу классов принимаемых решений. В промежуточном слое НС в качестве функций активации используются сигмоидальные функции:

$$z_m = f(u_m) = \frac{1}{1 + e^{-\lambda_m u_m}}, \quad u_m = \sum_{n=1}^N w_{n,m} x_n + w_{o,m}, \quad (3)$$

где z_m , $m = \overline{1, M}$ – выходной сигнал m -го нейрона промежуточного

слоя, состоящего из M нейронов, имеющих N входов; $x_n, n = \overline{1, N}$ – n -я компонента входного вектора признаков; $w_{n,m}$ – весовой коэффициент n -го входного признака x_n , поступающего на вход m -го нейрона промежуточного слоя; $w_{o,m}$ – значение смещения; λ_m – коэффициент, определяющий крутизну функции активации $f(u_m)$. Нейроны выходного слоя имеют пороговую функцию активации и используются для принятия решения о принадлежности анализируемой СППР к определенному классу решений:

$$y_k = \varphi\left(\sum_{m=1}^M v_{m,k} \cdot z_m + v_{o,k}\right) = \begin{cases} 1, & \text{если } y_k > 0, \\ 0, & \text{если } y_k \leq 0, \end{cases} \quad k = \overline{1, K} \quad (4)$$

где $v_{m,k}$ – весовые коэффициенты; $v_{o,k}$ – смещение; K – число выходов НС.

2. Иммунное обучение и адаптация нейросетевой МПР

Существующие методы обучения НС являются трудоемкими. Основной их чертой является отсутствие возможности модификации структуры и параметров НС. Для реализации этого предлагается выполнять обучение НС на основе ИИС. Задача обучения НС рассматривается как требование минимизации целевой функции (2), в которой в векторе параметров P НС используются параметры $w_{n,m}, w_{o,m}, v_{m,k}, v_{o,k}$ и $\lambda_m, n = \overline{1, N}, m = \overline{1, M}, k = \overline{1, K}$.

Основной идеей обучения НС с использованием ИИС является представление решаемой задачи в виде антигена, а возможные ее решения – в виде антител [13]. Сформируем популяцию антигенов $Ag_i = \langle x_1^i, x_2^i, \dots, x_n^i, y_1^i, y_2^i, \dots, y_k^i \rangle, i = \overline{1, S}$, где $x_1^i, x_2^i, \dots, x_n^i$ – входные переменные, а $y_1^i, y_2^i, \dots, y_k^i$ – выходные переменные модели для i -го примера обучающей выборки. В виде популяции антигенов выступает обучающая выборка относительно входных $x_n, n = \overline{1, N}$ и выходных $y_k, k = \overline{1, K}$ переменных. В качестве антител используются векторы настраиваемых параметров. В одном антителе кодируются все настраиваемые параметры НС: $w_{n,m}, w_{o,m}, v_{m,k}, v_{o,k}$ и $\lambda_m, n = \overline{1, N}, m = \overline{1, M}, k = \overline{1, K}$.

Для решения задачи используется модель кодирования настраиваемых параметров в виде адаптивного структурированного мультиантитела (рис. 1) [14], состоящего из двух частей, каждая из

которых может обрабатываться независимо друг от друга.

$w_{1,1}, \dots, w_{1,M}, \dots, w_{N,1}, \dots, w_{N,M}; w_{0,1}, \dots, w_{0,M}; \lambda_1, \dots, \lambda_M$	$v_{1,1}$...	$v_{M,K}$	$v_{0,M}$
ab_0	ab_1	...	ab_M	ab_{M+1}
Часть 1	Часть 2			

Рис. 1 – Структура мультиантитела

Популяция мультиантител представлена в виде $mAb = \{mAb_1, mAb_2, \dots, mAb_N\}$, где $mAb_i = \{ab_0, ab_1, ab_2, \dots, ab_K\}$, $i = \overline{1, S}$ – i -е адаптивное мультиантитело, длина которого L изменяется в процессе выполнения иммунного алгоритма; S – размер популяции мультиантител. Каждое мультиантитело популяции характеризуется полным множеством настраиваемых параметров НС. В части 1 мультиантитела закодированы весовые коэффициенты $w_{n,m}$, значения смещений $w_{o,m}$ и коэффициенты функции активации λ_m . В части 2 закодированы коэффициенты $v_{m,k}$ и смещения $v_{o,k}$ выходного слоя НС. Вторая часть является адаптивной, т.к. в процессе обучения, помимо оптимизации коэффициентов, изменяется и их общее количество, т.е. выполняется структурный синтез. Это приводит к изменению количества нейронов в скрытом слое НС и связей между ними. Структурированный способ формирования мультиантитела позволяет повысить эффективность иммунного алгоритма за счет отдельного применения иммунных операторов к каждой из частей. Вычисление аффинности выполняется для мультиантитела в целом, оперируя обоими составляющими.

В качестве вычислительной модели ИИС используются принципы клонального отбора и сетевого взаимодействия [13]. При клональном отборе антитело распознает антиген, клонируется, и полученные клоны подвергаются мутации. Если в результате мутации аффинность клона улучшается, то соответствующее антитело заменяется своим клоном, т.е. осуществляется клональный отбор. При использовании модели иммунной сети антитела взаимодействуют не только с антигенами, но и с другими антителами. Это обеспечивает эффект взаимной стимуляции и супрессии, что приводит к сокращению числа нейронов в промежуточном слое НС.

Алгоритм обучения, выполняемый в режиме off-line, представляет собой итерационную процедуру последовательной идентификации наблюдений из обучающей выборки. Антигены представляют примеры обучающей выборки. Каждое антитело кодирует одно из возможных решений, а количество антител в популяции соответствует количеству

примеров в обучающей выборке. Алгоритм обучения НС представляет собой следующую последовательность шагов:

1. Инициализация начальной популяции мультиантител mAb – выполняется случайным образом.

2. Вычисление аффинности Aff_{mAb-Ag} для каждого мультиантитела. Необходимо подставить параметры, закодированные в мультиантителе, в НС. На входы НС подаются входные признаки x_n и вычисляются значения выходных переменных y_k . Аффинность мультиантитела mAb с антигеном Ag вычисляется в виде:

$$Aff_{mAb-Ag} = (1 + d_{mAb-Ag})^{-1}, \quad (5)$$

где d_{mAb-Ag} – расстояние Хэмминга между полученными значениями выходов НС y_k , $k = \overline{1, K}$ и желаемыми y_k^r для всех S антигенов популяции Ag :

$$d_{mAb-Ag} = \sum_{k=1}^K y_k, \quad \text{где } y_k = \begin{cases} 1, & \text{если } y_k \neq y_k^r, \\ 0, & \text{если } y_k = y_k^r. \end{cases} \quad (6)$$

3. Клонирование мультиантител пропорционально их аффинности и формирование популяции клонов Cl . Параметрами оператора клонирования являются количество антител для клонирования g и кратность клонирования мультиантитела N_{Cl} . Используется фиксированное значение параметра g . Кратность клонирования мультиантитела N_{Cl} регулируется в процессе работы иммунного алгоритма в зависимости от аффинности мультиантитела по соотношению:

$$N_{Cl} = \alpha * N_{Cl_min} + (1 - \alpha) * N_{Cl_max}, \quad (7)$$

где $\alpha = \frac{Aff_{best} - Aff_{mAb-Ag}}{Aff_{best}}$; N_{Cl_min} и N_{Cl_max} – минимальная и максимальная кратность клонирования мультиантитела; Aff_{best} – лучшее значение аффинности в текущем поколении.

4. Мутация клонов обратно пропорционально аффинности мультиантитела и формирование популяции мутированных клонов MC . Мутация выбранных параметров мультиантитела mAb выполняется путем добавления гауссовского шума:

$$mAb_{i+1} = mAb_i + N(0, \sigma_i). \quad (8)$$

Для изменения среднеквадратического отклонения случайной величины σ_i используется соотношение:

$$\sigma_{i+1} = \sigma_i \frac{Aff_{best} - Aff_{mAb-Ag}}{Aff_{best} - Aff_{worst}}, \quad (9)$$

где Aff_{worst} – худшее значение аффинности в текущем поколении.

5. Вычисление аффинности популяции мутированных клонов MC в соответствии с (5). Если в результате мутации аффинность улучшилась, замена клонами соответствующих мультиантител в популяции mAb .

6. Вычисление аффинности антител внутри части 2 мультиантитела. Супрессия антител, аффинность которых больше заданного порога δ_{net} . Вычисление аффинности выполняется в соответствии с выражением:

$$mAb_{i+1} = mAb_i + N(0, \sigma_i), \quad (10)$$

где $d_{ab_i-ab_j}$ – расстояние между i -м и j -м антителами части 2 мультиантитела:

$$d_{ab_i-ab_j} = \|ab_i - ab_j\| = \sqrt{\sum_{m=0}^M (v_{m,i} - v_{m,j})^2}, \quad i, j = \overline{1, K}. \quad (11)$$

Выполнение супрессии путем удаления антител ab_i с аффинностью, большей заданного порога δ_{net} , позволяет уменьшить количество нейронов и связей между ними в скрытом слое и устранить их избыточность.

7. Проверка критерия останова. В качестве критерия останова используется либо достижение заданного порога аффинности, либо достижение заданного количества поколений работы алгоритма.

Результатом работы алгоритма будет мультиантитело с лучшей по популяции аффинностью, определяющее структуру НС и содержащее настроенные ее параметры. Шаги 1-5 алгоритма соответствуют принципу клонального отбора. Алгоритм на данных этапах работает с обеими частями мультиантитела. Шаг 6 соответствует принципу сетевого взаимодействия. Если ранее мультиантитело обрабатывалось как обычное антитело, то на данном шаге работа выполняется только с частью 2 мультиантитела, состоящей из отдельных антител.

В процессе функционирования СППР возможно изменение характеристик окружающей среды и свойств ОПР. В этом случае необходимо адаптировать МПР под эти изменения, т.е. осуществлять ее эволюцию. Предлагается для этого использовать процедуру обучения НС на основе ИИС, но в качестве обучающей выборки на вход будут поступать реальные значения входных признаков МПР, функционирующей в режиме on-line. Используется равномерное распределение нейросетевой обработки данных, основанное на

динамическом перераспределении наборов нейронов НС между вычислителями в зависимости от объема обрабатываемых данных, что позволяет существенно сократить время обучения и функционирования НС, а также уменьшить на порядок вычислительную сложность [17]. Предлагаемый подход позволит корректировать в реальном времени структуру и параметры МПР.

3. Экспериментальные исследования нейросетевой МПР

В качестве тестовой рассматривалась задача проектирования и анализа локальной компьютерной сети (ЛКС), которая является примером трудно формализуемой задачи, для решения которой необходим интегрированный подход. При проектировании ЛКС необходимо определить выходные параметры, которым должна удовлетворять сеть, и начальные условия (входные параметры), которые задаются перед процессом проектирования.

Анализ ЛКС позволил выделить входные признаки, наиболее существенные из которых следующие: x_1 – стоимость сети; x_2 – количество и расположение пользователей; x_3 – простота установки и изменения конфигурации сети; x_4 – пропускная способность сети; x_5 – надежность сети; x_6 – безопасность сети; x_7 – возможность расширения сети.

Исходя из требований проектирования ЛКС, имеем следующие основные выходные параметры: 1) сетевая операционная система (ОС); 2) топология сети; 3) сетевая технология.

Варианты выбора ОС: 1) Novell NetWare; 2) Microsoft Windows (NT, Vista, 7, 8, 10); 3) UNIX системы (Solaris, FreeBSD); 4) GNU/Linux системы; 5) IOS; 6) ZyNOS компании ZyXEL.

Варианты выбора топологии сети: 1) звезда; 2) шина; 3) кольцо; 4) дерево; 5) полносвязная; 6) ячеистая; 7) смешанная.

Варианты выбора сетевой технологии: 1) Fast Ethernet; 2) Token Ring; 3) FDDI.

Реализация данной задачи методом простого перебора вариантов неприменима, поскольку комбинаторная емкость перебора составляет несколько порядков. Более того, большая часть комбинаций вариантов никогда не будет реализована. Поэтому реализация была выполнена с применением эволюционных НС с использованием ИИС. Задача условно была разбита на три параллельные задачи, каждая из которых решает проблему по одному из выходных параметров. При этом каждый из выходных параметров может определяться значениями не всех входных, а только нескольких из них.

Для каждой из трех задач была создана своя НС, содержащая

определенное число входов и выходов. Так, для выбора ОС – это 7 и 6, для выбора топологии сети – 6 и 7, для выбора сетевой технологии – 6 и 3 соответственно. Количество скрытых слоев – один, содержащий 15 и более нейронов. Структура спроектированной НС для определения ОС приведена на рис. 2.

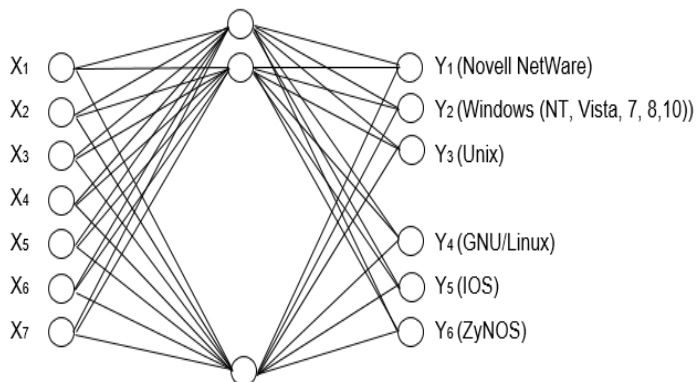


Рис. 2 – Структура нейронной сети для определения ОС ЛКС

Для моделирования полученных результатов использована компьютерная сеть Fast Ethernet с физической топологией «звезда» со скоростью передачи 100 Мбит, которая состоит из 10 четырехядерных вычислителей Intel Core 2 Quad CPU Q8200 @2.33GHz с видеокартой GeForce GTX 460. Распараллеливание выполнено с помощью технологий OpenMP и MS MPI. Для анализа процедуры параллельного и распределенного обучения НС на многопроцессорной системе проведен ряд экспериментов. Анализ показал, что технология MPI позволяет ускорить распределенное обучение НС в 11 раз для топологии передач данных «звезда». Однако при значительном увеличении количества процессоров в системе растет время на передачу данных между ними. Поэтому, при достижении определенного прироста производительности, дальнейшее увеличение количества процессоров в системе дает обратный эффект – производительность начинает падать. Стандарт OpenMP позволяет ускорить параллельное обучение НС до 7 раз. Ускоренная обработка данных позволяет функционировать интеллектуальной СППР в реальном времени.

Таким образом, проведенные исследования показали, что многослойные модульные нейронные сети персептронного типа с иммунным обучением целесообразно использовать для прогнозирования возможных вариантов решений при выборе параметров ЛКС, при этом наблюдается высокая устойчивость получаемых результатов.

НЕЧЕТКАЯ МОДЕЛЬ ПРИНЯТИЯ РЕШЕНИЙ С ИММУННЫМ ОБУЧЕНИЕМ

1. Описание МПР в виде правил нечеткого вывода

Для описания зависимости (1) будем использовать модель нечеткого вывода Такаги-Сугено первого порядка с правилами вида:

$$R_i : \text{if } x_1 \text{ is } A_{i1} \text{ and } \dots \text{ and } x_m \text{ is } A_{im}, \text{ then} \\ y_i = k_{i0} + k_{i1}x_1 + \dots + k_{im}x_m, \text{ is class}_i, i = \overline{1, n}, \quad (12)$$

где $A_{ij}, i = \overline{1, n}, j = \overline{1, m}$, – нечеткие множества и соответствующие им функции принадлежности (ФП), построенные в пространстве значений входных переменных x_j , k_{ij} – коэффициенты аргументов функции, значение выходной переменной y_i определяет класс системы. В качестве ФП используются гауссовы функции:

$$\mu(x) = \exp \left[- \left(\frac{x - c}{\sigma} \right)^2 \right], \quad (13)$$

где c – центр нечеткого множества; σ отвечает за крутизну функции. Для обучения и адаптации нечеткой МПР (11) на основе ИИС необходимо оптимизировать параметры ФП, а также параметры и структуру этой модели в соответствии с критерием (2).

2. Иммунное обучение и адаптация нечеткой МПР

Представим адаптивную систему нечеткого вывода в виде кортежа:

$$AFIS = \langle \{X_i, y_i\}, RB, DB, I, G(RB), L(DB), F \rangle, \quad (14)$$

где $\{X_i, y_i\}, i = \overline{1, n}$ обучающая выборка; RB – база правил; DB – база данных; I – механизм нечеткого логического вывода; $G(RB)$ – генерация и оптимизация базы правил; $L(DB)$ – генерация и оптимизация базы данных; F – функция, оценивающая эффективность нечеткой модели.

Основная трудность использования данной системы заключается в необходимости структурной и параметрической адаптации – поиске оптимального набора параметров термов лингвистических переменных и коэффициентов нечетких правил, а также структуры множества правил нечеткого вывода [14]. Алгоритм формирования нечеткой базы знаний представляет собой итерационную процедуру последовательной идентификации наблюдений из обучающей выборки с помощью ИИС. Антигены представляют собой примеры обучающей выборки. Каждое антитело кодирует одно правило базы знаний. Количество антител в популяции (правил в базе знаний) соответствует количеству примеров в обучающей выборке. Правила кодируются вещественными числами,

представляющими собой индексы нечетких множеств входных переменных и коэффициенты нечетких правил:

$$Ab_i = \langle IFS_{i_1}, IFS_{i_2}, \dots, IFS_{i_m}, k_{i_0}, k_{i_1}, \dots, k_{i_m} \rangle, \quad i = \overline{1, n},$$

где $IFS_{i_1}, IFS_{i_2}, \dots, IFS_{i_m}$ – индексы нечетких множеств для m входных переменных; $k_{i_0}, k_{i_1}, \dots, k_{i_m}$ – коэффициенты нечетких правил (11).

Аффинность антител популяции вычисляется по результатам нечеткого вывода в соответствии с выражением:

$$Af = (1 + \Phi(P))^{-1}, \quad (15)$$

где $\Phi(P)$ – эвклидово расстояние (2). В результате работы иммунного алгоритма лучшие по популяции антитела формируют популяцию клеток памяти, которая представляет множество нечетких правил вывода вида (11):

Полное множество правил (12) для реальных задач может быть достаточно велико. Для сокращения мощности множества правил (12) необходимо выделить некоторое минимальное подмножество, исключив избыточные и противоречивые правила. Данную задачу в предлагаемом алгоритме выполняет иммунный оператор супрессии [13], который обеспечивает сжатие популяции путем связывания антител в зависимости от значения их аффинности относительно других антител популяции. Антитела с похожей аффинностью удаляются из популяции в соответствии с установленным порогом сжатия сети, что приводит к уменьшению числа правил в базе знаний адаптивной системы нечеткого вывода.

Алгоритм структурной адаптации можно представить в виде следующей последовательности шагов:

1. Генерация начальной популяции антител Ab .
2. Цикл для каждого антигена Ag_i :

2.1. В МНВ подставляются параметры ФП и правило вывода с коэффициентами, взятыми из антитела. Вычисляются аффинности антител Ab к антигену Ag_i . Выбор n лучших антител для клонирования N раз.

2.2. Клонирование выбранных антител, формирование популяции клонов C .

2.3. Мутация популяции клонов C и формирование популяции клонированных и мутированных антител C^* .

2.4. Вычисление аффинности популяции C^* к антигену Ag_i . Редактирование популяции антител Ab_i и заполнение популяции клеток памяти M . Супрессия антител Ab внутри популяции.

- 2.5. Замена d худших антител новыми в популяции антител Ab .

3. Проверка критерия останова. Если не достигнут – переход к шагу 2, в противном случае – переход к шагу 4.

4. Конец.

Результат – популяция клеток памяти M , которая представляет собой набор правил вывода с вычисленными коэффициентами нечетких правил.

Параметрическая адаптация включает в себя процедуру оптимизации значений параметров ФП, целью которой является поиск такой модели, с таким набором параметров ФП, для которой качество принятия решений будет наилучшим. В роли антигена выступает обучающая выборка; в роли антител – параметры ФП вида (13) для входных переменных. Антитело представляет собой строку фиксированной длины:

$$Ab = \langle c_{11}, \dots, c_{1m}, \dots, c_{n1}, \dots, c_{nm}, \sigma_{11}, \dots, \sigma_{1m}, \dots, \sigma_{n1}, \dots, \sigma_{nm} \rangle,$$

где c_{ij}, σ_{ij} , $i = \overline{1, n}$; $j = \overline{1, m}$ – параметры гауссовых ФП вида (12) для n входных переменных, каждая из которых имеет m термов. Каждая из ФП задана на универсуме $X = [x_{min}, x_{max}]$, в качестве которого выбран замкнутый интервал действительных чисел. Исходный вектор параметров ФП генерируется случайным образом. Предлагается вещественное кодирование антител, что позволяет повысить точность найденных решений.

На вход адаптивной МНВ подается множество значений входных переменных из обучающей выборки. На основании параметров ФП, взятых из антител, и сгенерированной на этапе структурной адаптации базы правил, выполняется нечеткий вывод. По результатам нечеткого вывода вычисляется аффинность каждого антитела со всеми антигенами популяции в соответствии с (15). Поиск минимума функции (2) ведется в допустимом множестве изменения входных переменных x_i : $x_i^{min} \leq x_i \leq x_i^{max}$, $i = \overline{1, m}$. Антитела с лучшей аффинностью подвергаются клонированию и мутации. Антитела с худшим значением аффинности удаляются из популяции и заменяются новыми. Работа иммунного алгоритма прекращается при достижении критерия останова. База данных адаптивной системы нечеткого вывода формируется по лучшему антителу популяции.

Алгоритм параметрической адаптации можно представить в виде такой же последовательности шагов, как и алгоритм структурной адаптации, только в нем не используется иммунный оператор супрессии антител внутри популяции. Результат – антитело с лучшей аффинностью, содержащее настроенные параметры ФП.

Обобщенный алгоритм обучения адаптивной МНВ представляет собой итерационную процедуру, в которой организовано взаимодействие

процедур параметрической и структурной адаптации. На подготовительном этапе для каждого ОНР формируется обучающая выборка $(x_1^{(k)}, x_2^{(k)}, \dots, x_m^{(k)}, y^{(k)})$, $k = \overline{1, K}$, где $x_1^{(k)}, x_2^{(k)}, \dots, x_m^{(k)}, y^{(k)}$ – значения входных переменных x_1, x_2, \dots, x_m и выходной переменной y в k -ом примере; K – общее количество примеров в обучающей выборке. На основании известных минимальных и максимальных значений каждой входной переменной область определения этих переменных разбивается на отрезки, на каждом из которых задается гауссова ФП. На этапе структурной адаптации формируется N наборов нечетких правил, в каждом поколении иммунного алгоритма происходит модификация популяции в соответствии с выбранной схемой применения иммунных операторов, вычисление приспособленности каждого антитела, ранжирование антител по степени аффинности в популяции и формирование МНВ, на основе которой выполняется параметрическая адаптация. Иммунный алгоритм параметрической адаптации итерационно повторяется заданное число раз. В конце из полученной популяции выбирается наиболее приспособленное антитело, которое является результатом адаптации МНВ.

3. Экспериментальные исследования нечеткой МНР

Для оценки эффективности предложенных адаптивных МНВ были проведены экспериментальные исследования для двух тестовых функций (ТФ):

$$F_1(x_1, x_2) = x_1^2 + x_2^2, \quad (16)$$

$$F_2(x_1, x_2) = 3(1 - x_1)^2 \exp(-x_1^2 - (x_2 + 1)^2) - 10\left(\frac{x_1}{5} - x_1^3 - x_2^5\right) \exp(-x_1^2 - x_2^2) - \frac{1}{3} \exp(-(x_1 + 1)^2 - x_2^2), \quad (17)$$

которые исследовались в диапазоне входных переменных $x_1, x_2 \in [-3, 3]$. ТФ 1 является унимодальной функцией, а ТФ 2 – мультимодальной. Обучающая выборка сформирована в соответствии с диапазоном входных переменных и составляет 121 пример. По данным из обучающей выборки каждая ТФ восстанавливается с помощью адаптивной МНВ. Начальные значения параметров ФП устанавливаются таким образом, чтобы ФП были равномерно распределены по области определения функций. Эксперименты проводились с использованием разного количества лингвистических термов для каждой из двух входных переменных.

Были выбраны параметры иммунных алгоритмов структурной и параметрической адаптации МНВ. Размер популяции антител для структурной адаптации МНВ соответствовал количеству примеров в

обучающей выборке.

Количество антител для клонирования n влияет на скорость сходимости алгоритма. Использовался оператор пропорционального клонирования [13], который клонирует n антител пропорционально их аффинности.

Вероятность мутации влияет на скорость сходимости алгоритма. Большая вероятность мутации поддерживает разнообразие антител в иммунной системе на достаточно высоком уровне и позволяет находить глобальный экстремум. Небольшая вероятность мутации позволяет исследовать локальную область вокруг антитела, делая небольшие шаги в поиске антител с лучшей аффинностью, ведущих к оптимальному значению. Использовался оператор обратно-пропорциональной мутации [13], согласно которому вероятность мутации обратно пропорциональна аффинности антитела.

При редактировании популяции антитела с худшей по популяции аффинностью заменялись новыми случайно сгенерированными антителами, что обеспечивает поддержку разнообразия в популяции.

Оператор супрессии удаляет похожие антитела для того, чтобы избежать избыточности и поддерживать разнообразие в популяции [13]. Оператор супрессии обеспечивает сжатие популяции путем связывания антител в зависимости от значения их аффинностей относительно других антител популяции с установленным порогом сжатия сети, равным 0,001. Критерием останова является достижение заданного количества поколений работы алгоритма.

В табл.1 приведены данные по среднеквадратическому отклонению по результатам 10 запусков алгоритма для тестовых функций (16) и (17).

Таблица 1 – Результаты экспериментов

Количество лингвистических термов	Тестовая функция	
	1	2
3	0.2	0.69
5	0.051	0.082
7	0.012	0.05

Из табл. 1 видно, что использование большего количества ФП для каждой переменной позволяет повысить точность восстановления функции. Восстановленные поверхности тестовых функций (6) и (17) и полученные ФП представлены на рис. 3 и рис. 4.

При адаптации МНВ изменяются не только коэффициенты нечетких правил, но и количество правил нечеткого вывода за счет выполнения супрессии. Количество правил сократилось с 25 до 19 для ТФ (15), и до 23 для ТФ (16).

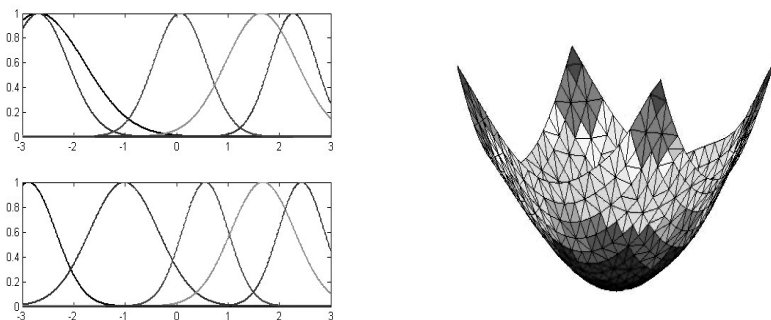


Рис. 3 – Настроенные ФП и восстановленная поверхность для ТФ (15)

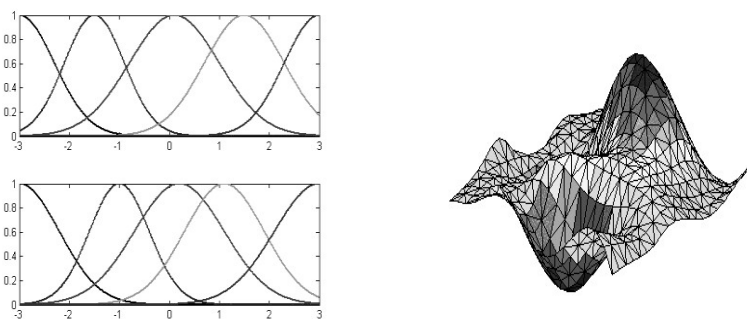


Рис. 4 – Настроенные ФП и восстановленная поверхность для ТФ (16)

ВЫВОДЫ

В работе рассмотрено решение актуальной задачи повышения качества функционирования интеллектуальных СППР на основе использования гибридных МПР с иммунным обучением. Эффективность решения этой задачи определяется выбором МПР в виде многослойной НС и МНВ Такаги-Сугено первого порядка, а также разработкой методов их параллельного обучения и эволюции на основе иммунного подхода. Эволюция МПР рассматривается как задача их адаптации к условиям изменения внешней среды и свойств ОПР, состоящая из процедур коррекции структуры (числа нейронов в скрытых слоях и связей между ними – для НС; числа правил вывода – для МНВ) и параметров моделей, использующих иммунные модели клонального отбора и иммунной сети.

Осуществлена программная реализация и проведены экспериментальные исследования предложенного подхода с использованием технологий параллельного программирования OpenMP и MS MPI на примере проектирования локальной компьютерной сети и

восстановления тестовых функций, которые подтвердили эффективность предложенных эволюционных МПР, использующих нейросетевой, нечеткий и иммунный подходы.

СПИСОК ЛИТЕРАТУРЫ

- [1] Снитюк В.Е. Эволюционные технологии принятия решений в условиях неопределенности. – К.: «МП Леся», 2015. – 347 с.
- [2] Zohuri B. Neural Network Driven Artificial Intelligence: Decision Making Based On Fuzzy Logic, 2017. – 379 p.
- [3] Ocak H. A medical decision support system based on support vector machines and the genetic algorithm for the evaluation of fetal well-being //Journal of medical systems, 2013. – Т. 37. – №. 2. – P. 99 – 113.
- [4] Мелихова О.А. Методы поддержки принятия решений на основе нейронных сетей // Актуальные проблемы гуманитарных и естественных наук, 2015. – № 9 – 1. – С. 52 – 59.
- [5] Бова В.В. Применение искусственных нейронных сетей для коллективного решения интеллектуальных задач / // Известия Южного федерального университета. Технические науки, 2012. – № 8. – С. 131–136.
- [6] Бухаров О.Е., Боголюбов Д.П. Распараллеленная самообучающаяся система поддержки принятия решений на генетических алгоритмах и нейронных сетях // Системный администратор, 2014. – Вып. № 9 (142). – С. 57 – 69.
- [7] Столярчук А. С. Проектирование модели принятия решения на базе нечёткой логики / Ученые заметки ТОГУ, 2015. – том 6. – № 4. – С. 126 – 130.
- [8] Samigulina G. A. Development of decision support systems based on intellectual technology of artificial immune systems //Automation and Remote Control, 2012. – Т. 73. – №. 2. – С. 397 – 403.
- [9] Самигулина Г. А. Разработка интеллектуальных экспертных систем прогнозирования и управления на основе искусственных иммунных систем // Проблемы информатики. –Новосибирск, 2010. –№.1. –С. 15–22.
- [10] Комарцова Л.Г. Эволюционные методы формирования нечетких баз правил // Материалы международной конференции OSTIS-2011. – С. 181 –184.
- [11] Комарцова Л.Г. Исследование нейросетевых алгоритмов обучения в интеллектуальных эволюционных системах // Материалы IV-й международной конференции OSTIS-2014. – С. 233 – 238.
- [12] Бухаров О. Е., Боголюбов Д. П., Мизикин А. А. Разработка оболочки системы поддержки принятия решений с использованием эволюционных алгоритмов. // «Промышленные АСУ и Контроллеры». – М.: НАУЧТЕХЛИТИЗДАТ, 2013. –№7. – С. 37 – 45.
- [13] Dasgupta D. Recent Advanced in Artificial Immune Systems: Models and

Applications // Applied Soft Computing. Elsevier, 2011. – P. 1574–1587.

[14] Korablev N. Immune Approach for Neuro-Fuzzy Systems Learning Using Multiantibody Model / N. Korablev, I. Sorokina // ICARIS 2011, Springer Lecture Notes in Computer Science, 2011. – Vol. 6825. – P. 395–405.

[15] Korablyov M. Hybrid evolutionary decision-making model based on neural network and immune approaches / M. Korablyov, N. Axak, D. Soloviov // Proceedings of the 13-th International Scientific and Technical Conference on Computer Sciences and Information Technologies (CSIT), 2018. – V.1. – P. 378–381.

[16] Korablev N.M. Intelligent Evolution Decision Support System based on Fuzzy Logic and Immune Approach / N.M. Korablev, D.N. Soloviov, R.R. Malyukov // System technologies. №6 (107). – Dnipro, 2016. – P. 98–106.

[17] Axak N. Development of multi-agent system of neural network diagnostics and remote monitoring of patient // Eastern-European Journal of Enterprise Technologies, 2016. – 4/9 (82) – P.4–11.

HYBRID DECISION-MAKING MODELS USING AN IMMUNE APPROACH

Korablev N., Fomichev A., Soloviev D., Chuprina A.

Modern decision support system (DSS), characterized by processing large amounts of information in an uncertain environment, which requires the use of effective methods and models using a variety of intelligent parallel processing technology. The article offers hybrid decision-making models (DMM), presented both based on the neural network (NN) and in the form of fuzzy inference rules, the training and adaptation of which are carried out on high-performance systems using immune models depending on changes in environmental characteristics and properties decision making object (DMO). The process of learning and adapting NN consists in tuning and correcting network parameters, as well as the number of neurons in hidden layers and the connections between them. The process of training and adapting fuzzy inference models (FIM) consists in adjusting the form and parameters of membership functions, as well as the parameters and structure of the model (the number of inference rules).

Keywords. Decision-making, hybrid model, training, adaptation, neural network, fuzzy conclusion, artificial immune systems

COMPUTER MODELING OF THE PROCESSES OF DEVELOPMENT OF INFORMATION TECHNOLOGY IN DYNAMIC PROCESSES OF THE FORMATION OF CLASSES OF THE GENERALIZED ARTIN'S HYPOTHESIS

Vostrov G., Opiata R.

Abstract. The relationship between the processes of formation of classes of primes in the generalized Artin's hypothesis and the theory of information, and as a consequence, information technology, is investigated.

It is proved that probabilistic methods of the theory of information and information technologies are the basis for constructing computer models of classes of primes in accordance with the generalized Artin's hypothesis. Methods for calculating the Artin's constants are developed and the convergence of the estimates of the constants in probability to limit values is established. The foundations of a number-theoretic analysis of Artin's constants and related classes are created.

Keywords. Generalized Artin's classes, Artin's constants, class probabilities, stability of estimates of the Artin's constants? convergence in probability.

INTRODUCTION

The solution of many problems in such applied areas as electronics, electrical engineering, modeling of complex, both deterministic and stochastic, nonlinear dynamic systems, information technology and many other applied areas of human activity, depends on solving a significant number of mathematical problems that have not yet been solved. Artin's hypothesis of primitive roots is one of these fundamental mathematical problems. For almost a century, it remains unresolved. The solution to Artin's problem is important for research in such applied areas as the creation of effective methods of protecting information using cryptographic methods, the development of pseudo-random number generators, the modeling of dynamic processes in stock markets, and the construction of advanced algorithms for testing software products of high complexity. One of the options for cryptographic information protection is the discrete logarithm method. The development of the Monte Carlo method and its application in the theory of modeling complex systems depends on the creation of effective generators of pseudo random numbers given by the laws of probability distribution. The construction of such generators is especially important in the methods of testing software applications. Modeling processes in modern stock markets is not possible without high-quality random number generators with a given distribution law [1-3]. Another urgent applied problem is the modeling of self-organizing nonlinear dynamic systems, which are commonly called synergetics, taking into account the deep modeling of the phenomenon of self-organization in

complex systems consisting of transition sequences from one phase state to another using random number generators with a given probability distribution law [3]. The numerical sequences of iterative models of cyclic fixed points of dynamical systems are determined by the properties of the primes with which they are represented. In this case, the question always arises: what distribution laws obey prime numbers. Riemann in 1869 proposed the function:

$$\xi(s) = \sum_{n=1}^{\infty} \frac{1}{n^s} = \prod_{p \in P} \left(1 + \frac{1}{p^s} + \frac{1}{p^{2s}} + \dots \right)$$

where s is a complex variable, P is the set of all primes [3,4]. The function is called the Riemann's zeta function. Concerning this Riemann's function, a hypothesis was formed according to which all non-trivial zeros of this function are on line $1/2 + iy$, where $i = \sqrt{-1}$ and $y \in R$. It follows that all primes lie on this line since y takes values from a set that includes all primes P . Moreover, for any prime number p $\xi(1/2 + ip) = 0$. In essence, this was the first attempt to find the law of the distribution of primes. So far, the hypothesis has not been proved. Note that any function $f(s)$ that is defined in the complex space C , and therefore $s \in C$, is usually called analytic.

The study of the Riemann's analytic zeta function has been done by many mathematicians. In particular, it was proved that $\xi(-2) = 1 + 2^2 + 3^2 + \dots + k^2 + \dots = 0$ is the value of s is a trivial zero. This paradoxical fact for applicants does not contradict the theory of analytic functions, but from the point of view of the distribution of primes it does not give an adequate description of the distribution of primes on the number line R or Q or Z . Starting with the works of Fermat, Euler, Dirichlet, Gauss, Chebyshev [3,4], systematic attempts were made to establish their distribution law in two-dimensional real space [4]. In 1896, independently, Adamard and Valle Poussin proved that equality is true:

$$\pi(x) = \int_2^x \frac{dt}{\ln t} + O\left(x \cdot e^{-c\sqrt{\ln x}}\right) \quad (1)$$

where $\pi(x)$ is the number of primes $p \leq x$, and the first term in the form of a logarithmic smooth function determines the logarithmic law of the distribution of primes in approximate form, and the second term is the remainder term that describe the inconsistency of the step function $\pi(x)$ when it is approximated by a logarithmic function. In many studies, an analysis of the residual term is given, it is proved that if it is considered as a function of x , it has a fractal nature [3,8].

It is known that the distance between primes increases, and the relation:

$$p_{n+1} - p_n > \varepsilon \frac{\ln(n) \cdot \ln \ln(n) \cdot \ln \ln \ln(n)}{(\ln \ln \ln(n))^2} \quad (2)$$

where n is the prime number in the ordered whole set of primes [5,6], a constant that is generally difficult to calculate. The proof of this relation does not mean at all that absolutely complete information on the distribution of primes in the system of the logarithmic distribution law has been obtained.

It should be noted that the Riemann's hypothesis, numerous studies of its trivial and non-trivial zeros, the proof of the logarithmic law of the distribution of primes were a source of new information that is fundamental to the modern world, on the one hand, and on the other, these results led to the creation of new information technologies. One of the areas of deepening information technology was the formulation in 1927 by the French mathematician Artin's of a hypothesis about the primitive roots of primes $p \in P$, and accordingly the primitive roots of residue groups $(Z/pZ)^*$ modulo prime p . Consider the definition of the primitive root of a prime number p . The numbers $a \neq 1$, $a \neq k^2$ is the primitive (antiderivative) root of the number p , if the following relations are true:

$$\begin{cases} a^{p-1} \equiv 1 \pmod{p} \\ a^{\frac{p-1}{n}} \not\equiv 1 \pmod{p}, \quad n > 1 \end{cases} \quad (3)$$

The n is the divisor of $p-1 = \prod_{i=1}^k p_i^{\alpha_i}$.

Given the definition of a primitive root, Artin's hypothesis is:

$$\pi(x, a) = c(a) \cdot \pi(x) \quad (4)$$

where $\pi(x, a)$ is the number of primes p less than or equal to x , for which $a \neq \pm 1$ and $a \neq k$ are according to (1) their primitive roots, $c(a)$ is the Artin's constant. More precisely, this hypothesis should be presented as follows:

$$\begin{cases} \pi(x, a) = c(a, x) \cdot \pi(x), \\ \lim_{x \rightarrow \infty} c(a, x) = c(a) \end{cases} \quad (5)$$

But then $c(a, x) = \frac{\pi(x)}{\pi(x, a, x)}$ and in probability converges to $c(a)$,

and therefore has a probability theory interpretation: $c(a)$ is the probability of choosing from the set P a prime number p such that a is its primitive root. Note that the first relation in (1) is always satisfied if a and p are coprime numbers according to Fermat's theory [3].

It should be noted that Artin's proposed his ratings for $c(a)$ at $a = 2$. But as proved by Hooley [5], these estimates are not true. He also proved the validity of the relation:

$$\pi(x, 2) = \frac{c(2) \cdot x}{\ln x} + O\left(x \frac{\ln \ln(x)}{(\ln(x))^2}\right) \quad (6)$$

at the same time $c(2) = \prod_{p \in P} \left(1 - \frac{1}{p(p-1)}\right)$ and an assessment of the value

of $c(2) = 0,373955813\dots$ As will be shown later, this estimate is true only with the accuracy of the first two decimal places. This can be easily explained by a very simple consideration. The expression for $c(2)$ proposed in (4) depends on all primes, but this is not true because $a = 2$ is not a primitive root for all primes $p \in P$. The values of the constants $c(a)$ for $a > 2$ and $a \neq k^2$ were not presented in more than one scientific article.

It should be noted that any number $a > 1$ and coprime to p is the basis for considering the recursive function $f(x) \equiv a \cdot x \pmod{p}$, which leads to a recursive iterative sequence.

$$f(x_0 = 1) = 1, f(x_{n+1}) = x_{n+1} \equiv ax_n \pmod{p} \quad (7)$$

According to Fermat's theorem [4], if a is not a primitive root for p , then the process of recursive computations will continue for such m that equality $f(x_n = m) \equiv x_{m-1} \cdot a \pmod{p} = 1$ is achieved i.e.

$$a^m \equiv 1 \pmod{p} \text{ и } m < p - 1 \quad (8)$$

From Fermat's theorem and the properties of the group $(Z/pZ)^*$ residues modulo p , it follows that in this case a is the primitive (antiderivative) root of some subgroup of the group $(Z/pZ)^*$. Moreover, m is the order of this subgroup, which is usually denoted by $card_a(p)$, the

number of adjacency classes for this subgroup is denoted by $ind_a(p)$. According to the cyclic group theorem $(Z/pZ)^*$, the equality:

$$p - 1 = card_a(p) \cdot ind_a(p) \tag{9}$$

The validity of equality (6) follows also from Fermat's theorem. In conclusion, it should be noted that from the relations (5) and (6) the validity of the equation follows:

$$c \equiv a^x \pmod{p} \tag{10}$$

No more complicated analysis of this equation leads to the consideration of four options for its solution: Given a, x, p - calculate c ; Given c, a, p - calculate x ; Given c, x, p - calculate a ; Given c, a, x - calculate p ;

The first equation is solved using recursion (7), (8). The second equation is the formulation of the discrete logarithm problem. The variable p can also be a pseudo prime number equal to p^k or equal to 2^p , as noted in [3], this problem in the general case may be algorithmically unsolvable.

In the third and fourth cases, the solution of such equations is a problem of extreme complexity. There are no publications on this subject.

From the above analysis it follows that equation (5) allows us to study the Artin's hypothesis from a more general point of view, when any natural number $a > 1$ and it can be used as a classifier of the set of all primes in the magnitude of $ind_a(p)$, which is the object of further research. As will be established, Artin's hypothesis of primitive roots will be a frequent case of its more general formulation.

MODELING THE PROCESSES OF GENERATING DYNAMIC INFORMATION ABOUT THE STRUCTURE OF CLASSES OF PRIMES ON A GIVEN BASIS

Now we return to the logarithmic law of the distribution of primes [1-4] in order to pay attention to the fact that the above equality does not provide comprehensive information about the structure of spaces between primes. Obviously, with increasing prime numbers, the gaps between them increase quite significantly. From relation (2) it is completely impossible to conclude by what laws the distance between the gaps changes, how often the dips appear statistically and most importantly how the structure of the decomposition of $p-1$ numbers into simple factors changes. It is especially important to have information about the distribution of smooth primes [3]. This information is especially important when solving the discrete logarithm problem and applying algorithms for solving it in the modern coding theory

and in modern cryptographic methods of information protection. As shown in [7], it is practically impossible to find smooth large prime numbers. It follows that it is of considerable interest to establish the laws of distribution of primes not only with respect to their primitive roots, but to the roots of subgroups of the residue group modulo a prime number. Artin's hypothesis does not imply such detailed studies. Moreover, it is necessary to find the laws of the relationship between the laws of the distribution of primes in sets corresponding to various primitive roots and roots of subgroups. Such tasks were not considered at all.

The second circumstance is that simultaneously with this fact, the dynamics of change in $O\left(x \cdot e^{-\frac{c}{2}\sqrt{\ln x}}\right)$ [8] is investigated. In [3, 4], the

entropy estimate of this estimate was obtained and it was proved that it has a fractal character. These facts are the basis for the formation of proposals on the need to study other models of the distribution of prime numbers. Another problem related to the distribution of primes appeared in 1927, when the famous mathematician Artin's formed a hypothesis about the distribution of primes for which the natural number $a > 1$ is given is its primitive root [1,5].

In addition, it is generally accepted, even at the present time, that it makes sense to study it more fundamentally. The first attempt was made by D. Zagier [8], but not completed. The results obtained by the author confirm the very complex fractal behavior of this component. It follows that it is necessary to significantly improve the study of the depth of classification of primes, taking into account all models for the formation of classes of primes for any given basis $a > 1$. Further more detailed studies of this component confirm that although the logarithmic distribution law is fulfilled, nevertheless, complete information on the dynamic properties of primes and their relationships with their primitive roots remains poorly studied. In the future we will consider any values of the base and large units.

According to Artin's hypothesis [5], the set of such primes has the distribution law $\pi(x, a)$ as an expression:

$$\pi(x, a) = c(a) \cdot \pi(x)$$

where $\pi(x)$ is the distribution of prime numbers, and $c(a)$ is a constant dependent on a . Until now, despite numerous studies, this hypothesis has not been resolved. However, it is not known if this is true for any a values. If the hypothesis is correct, then the question remains how to estimate the constant $c(a)$ for each concrete a and which properties of the number a influence its value. Answers to these questions are still missing. In works [3, 5] a detailed analysis of all the results of research in the field of solving the Artin's

hypothesis is given. It should be noted that the proof of Artin's hypothesis is important both from a theoretical point of view in number theory, and from an applied rehenium point, because it's positive solution is important in cryptography, coding theory, and the theory of dynamical systems. In [6], a generalized Artin's hypothesis was formed for any $a > 1$, i.e. and at the same time a may not be a primitive root. According to Artin's generalized theory, the following equality is true:

$$\pi(x, a, i) = c(a, i) \cdot \pi(x) \tag{11}$$

where $a > 1$, i – is the index of the subgroup of the group $(Z/pZ)^*$ of primes in the classification of prime numbers generated by the numbers a , $c(a, i)$ is a constant. According to the classification built in [6]:

$$P(a, i) = \{p \in P \mid (p-1)/card_a(p) = i\} \tag{12}$$

where $card_a(p)$ is the length of the dynamic recursion $x_{n+1} \equiv ax_n \pmod{p}$ at $x_0 = 1$, P is the set of all primes.

It is not difficult to show that for any $a > 1$ the equality:

$$\sum_{i=1}^{\infty} c(a, i) = 1 \tag{13}$$

This means that primes are evenly distributed in classes $P(a, i)$ for any a . By uniformity is meant that within each class of primes $P(a, i)$ a logarithmic law of the distribution of primes is preserved. The constant $c(a, i)$ determines the measure of puncturing prime numbers, based on the value a . If $i = 1$ then a is the primitive root of all primes $P(a, 1)$. For an arbitrary natural number x , the equality

$$\pi(x, a, i) = c(a, i, x) \cdot \pi(x)$$

Moreover, if $x \rightarrow \infty$, then $c(a, i, x)$ tends to the limit value $c(a, i)$. If we put $i = 1$ then $c(a, 1)$ will be Artin's constant for primitive roots. In this case $a \neq \pm 1$, and $a \neq k^2$ for none $k \in N$. This is true according to Fermat's theorem [3, 4]. Wherein, a is the primitive root of the group of residues $(Z/pZ)^*$ for any $p \in P$ such that $P(a, 1) = \{p \in P \mid (p-1)/card_a(p) = 1\}$. It is important to investigate the classes of primes $P(a, i)$ for $i > 1$ since in this case the positive integer a will be the primitive root for the subgroups of the group $(Z/pZ)^*$ with the index defined by the relations:

$$P(a,i) = \{p \mid (p-1)/\text{card}_a(p) = \text{ind}_a(p)\}$$

where $\text{ind}_a(p) = i$ is the index of the subgroup of $(Z/pZ)^*$. The classes of primes $P(a,i)$ have not yet been studied and the distribution of primes in these classes is not known. In [1], an assumption was made that $P(a,i)$ at $i > 1$ is proportional to $P(a,1)$ with a factor of $1/i^2$. Since $i > 1$ is considered, in this case it is important to know the distribution of prime numbers for the value $a = k^2$. This is an important generalization of Artin's hypothesis. At the same time, the probability of:

$$P(p \in P(a,i) \& p \in P) = |P(a,i)|/|P| = c(a,i)$$

Membership agrees exactly with the provisions of the theory of probability, and therefore, estimating $c(a,i)$ on the basis of successive statistical tests and the law of large numbers is parity.

The determination of $c(a,i)$ for any a,i using analytical methods is unlikely in the near term. However, the formation and development of experimental mathematics [1, 2] opens up another way to solve this problem by using computer simulation of nonlinear dynamic processes for the formation of classes of prime numbers.

The process of modeling the distribution of primes in classes $P(a,1), P(1,2), \dots, P(a,k), \dots$ was reduced to choosing a set of consecutive primes from a set of a sufficiently large sample of these classes. The number of primes analyzed at each interval of natural numbers was chosen to be 500,000. This choice was largely due to the fact that it was previously established that reducing this value leads to more significant fluctuations in estimates, although convergence to the limit over the entire set of any intervals, even if they are not placed consistently, has the same character.

The process of statistical testing of $p \in P$ primes for checking their belonging to class $P(a,i)$ was reduced to calculating for the selected number p the recursive procedure $x_0 = 1, x_{n+1} = ax_n \pmod{p}$ until the pairs $ax_i \equiv 1 \pmod{p}$ were reached at some step i . Then $\text{card}_a(p) = i$ and according to Fermat's theory and the cyclic group theorem the number $p-1$ is divisible by i and then $\text{ind}_a(p) = (p-1)/\text{card}_a(p) = i$, and therefore $p \in P(a,i)$ and if $i = 1$, then a is the primitive root of the cyclic group $(Z/pZ)^*$, and otherwise it is the primitive root of some subgroup. At $i > 1$,

we obtain the primitive roots of the subgroups of the $(\mathbb{Z}/p\mathbb{Z})^*$ residue group with the index $i > 1$. The study of the distribution law of prime numbers p on their belonging to $P(a, i)$ had the character of consistent statistical tests on the set of natural numbers containing the first 500,000 primes. At the first stage, primes p were chosen from the set $\{p_1, p_2, \dots, p_{500000}\}$. With this $x = p_{500000}$. For each $n \in \{2, \dots, x\}$, we had to solve two problems: check n for simplicity, and if $n = p \in P$, then $p - 1$ was decomposed into simple factors, i.e. systematically solved two non-simple problems of checking numbers for simplicity and decomposition into simple factors. An effective algorithm for solving them was created based on probabilistic methods in the theory of elliptic curves. As a result of analyzing $a \in \{2, \dots, k\}$, $P(a, 1), \dots, P(a, l)$ sets were obtained for some $l < x$ and absolutely exact values of their powers were calculated, i.e. $|P(a, 1)|, \dots, |P(a, l)|$, and then estimates of:

$$c(a, 1, x) = |P(a, 1, x)| / \pi(x), \dots, c(a, l, x) = |P(a, l, x)| / \pi(x)$$

while $c(a, 1, x) \rightarrow c(a, 1), \dots, c(a, l, x) \rightarrow c(a, l)$ with $x \rightarrow \infty$ were obtained.

At the next stage, work was also carried out for prime numbers from the $\{p_{500000}, \dots, p_{1000000}\}$ interval and the values of the $c(a, 1), \dots, c(a, l)$ constants were calculated using the same scheme. At the same time l increases. The $\{p_1, \dots, p_{500000}\}$ and $\{p_{500000}, \dots, p_{1000000}\}$ sequences were combined, and the estimates of the generalized Artin's constants were again calculated and the process of their refinement was studied on the basis of the theory of large numbers in probability theory. This procedure continued until $x = p = 179424673$ and this is a ten million prime numbers. It was found that $c(a, 1), \dots, c(a, k)$ in probability converges to some values, the exact values of which are irrational and possibly transcendental numbers. In the process of estimating the $c(a, i)$ constants, two important theorems were proved:

Theorem 1. For any $a \in \{2, 3, \dots, k, \dots\}$ that is not a square, i.e. $a \neq k^2$. The number of non-empty classes of primes tends to infinity at $x \rightarrow \infty$.

Theorem 2. For any $a \in \{2, 3, \dots, k, \dots\}$ that is not a square, i.e. $a \neq k^2$ The number of prime numbers in $P(a, i)$ tends to infinity at $x \rightarrow \infty$.

These theorems are the basis of the convergence of a sequence of statistical tests to marginal values. Since for any $x \in N$ it is obvious that:

$$\sum_{i=1}^k |P(a, i, x)| = \pi(x)$$

$$P(a, i, x) \cap P(a, j, x) = \emptyset$$

at $i \neq j$, it follows from this that:

$$\sum_{i=1}^k c(a, i) = 1$$

This is true for all values of $x \rightarrow \infty$. The review [5] provides an estimate of $c(2, 1)$, which is identified by $c(2, 1)$ in our sense, but $c(2, 1)$ differs from the estimate of $c(2, 1)$ starting from the fifth decimal place and this is a theoretical error of the survey works. For different $a \in \{2, 3, 5, 6, 7, 8, 10, 11, \dots\}$, the behavior of the $c(a, i)$ constants is complex group-theoretic and number-theoretic. The study of their dynamic properties is beyond the scope of this work. It should be noted that the results of computer simulation of the processes of distribution of primes are calculated with an accuracy of the eleventh decimal place for estimates of $c(2, 1), c(3, 1), c(5, 1), c(6, 1), \dots$ values. This cannot be asserted for classes by the $i \geq 2$ index. To achieve the same accuracy with $i \geq 2$, it is necessary to significantly increase the number of prime numbers. With an increase in the i class index $P(a, i)$ more than three requirements and the volume of the analyzed primes increases in accordance with the unexplored laws.

Probability-theoretic interpretation of the constant:

$$c(a) = \frac{\pi(x, a)}{\pi(x)} \text{ at } x \rightarrow \infty$$

Consider the probability space (Ω, F, P) based on:

$$\Omega = \{\omega_1, \dots, \omega_n, \dots\} = \{p_1, \dots, p_n, \dots\} = P$$

Obviously at $x \rightarrow \infty$ the numbers are $\pi(x) \rightarrow \infty$, $\pi(x, a) \rightarrow \infty$, but:

$$\pi(x, a) = |P(a, 1, x)|, \pi(x) = |P(x)|, c(a, 1, x) = \frac{|P(a, 1, x)|}{|P(x)|}$$

and at $x \rightarrow \infty$ it is obvious that:

$$|P(a,1,x)|/|P(x)| \rightarrow c(a,1)$$

is where $x \in P$, $P \rightarrow \infty$,

$$P(a,i,x) = \{p \mid p \leq x \ \& \ (p-1)/\text{card}_a(p) = i\}$$

is at $x \rightarrow \infty$ $P(a,i,x) \rightarrow P(a,i)$. Thus:

$$c(a) = \lim_{x \rightarrow \infty} \pi(x,a)/\pi(x)$$

It follows from Artin's hypothesis that with $c(a,1)$ there is precisely the probability of a random event $P(a,1)$ consisting of a choice of $\Omega = \{p_1, \dots, p_n, \dots\}$ of a prime number p for which a is an original root of the cyclic group $(Z/pZ)^*$. To estimate this probability, the law of large numbers and the method of successive statistical tests were used. The essence of the method is that the first test group was reduced and calculated for $\{p_1, p_2, \dots, p_{500000}\}$ for each $a \in \{2, 3, \dots, 16\}$ evaluation of the values of $c(a,i,x)$ at $x = p_{500000}$ for all possible values of $i = \{1, 2, \dots, k, \dots\}$, that is, $\tilde{c}_1(a,1,x), \dots, \tilde{c}_1(a,k,x), \dots$ was calculated on the next iteration, the same tests were performed for the second iteration on the set $\{p_{500001}, \dots, p_{1000000}\}$. $\tilde{c}_1(a,1,x), \dots, \tilde{c}_k(a,1,x), \dots$ Estimates were obtained at the same time $\tilde{c}_1(a,1,x), \dots, \tilde{c}_k(a,k,x), \dots$, provided that the first and second samples were combined and computed values and were determined by $|\tilde{c}(a,i,x) - c(a,i,x)| \leq \varepsilon$ for all x . The main focus was on $c(a,1,x)$. As a result of some iterations, it was found that for all a the estimates obtained:

$$P(x) = \{p \mid p \leq x\}$$

$$P(a,i,x) = \{p \mid p \leq x \ \& \ (p-1)/\text{card}_a(p) = i\}$$

The order of the cyclic group of the subgroup $(Z/pZ)^*$. If $l = p - 1$, then a is an original root, and if $l < p - 1$ is the original form of the $c(a)$ Artin's measure, $c(a,i)$ is a measure of classes by $P(a,i)$ in P . At that $c(a,i) = |P(a,i)|/|P|$ and at the same time:

$$\sum_{i=1}^{\infty} c(a,i) = 1 \text{ for all } a > 1$$

This applies only to classes with indexes $i = 1$. For $i \geq 2$ it is necessary to increase the number of statistical tests. This is naturally due to the fact that the classes $P(a, i, x)$ for $i \geq 2$ from numerical theorems contain less than prime numbers. In [1] it is stated that this decrease should be of the order of $1/i^2$, but this is an erroneous assertion. The degree of decline essentially depends on the properties of a and requires a separate study. Case $a \in \{4, 9, 16\}$ requires separate investigations, because these numbers cannot be primitive roots of that number p , in accordance with the Fermat theorem [3] cannot be generating elements of groups $(\mathbb{Z}/p\mathbb{Z})^*$. However, they are generating elements of the subgroups of the group $(\mathbb{Z}/p\mathbb{Z})^*$ with even indices. All classes with odd indices are empty sets. Table 1 shows the constants for $c(a, 1)$ for all a except $\{4, 9, 16\}$. Analysis of the table. The table contains over a thousand columns. The analysis of these data is numerically theoretical and group-specific and goes beyond the scope.

The simulation process of the dynamics of the formation of prime numbers was constructed on the following assumptions. Suppose that an ordered set of prime numbers $P = \{p_1, p_2, \dots, p_k, \dots\}$ is given, whose elements are ordered in ascending order. All this set was split into a subset of 500,000 primes. The number of 500,000 is due to the limitations of MS Excel, as a statistical analysis tool, on a number of characteristics of the process of generating prime numbers. Only one restriction is important. We always select 500,000 consecutive primes of the set P . In the current version of Excel, this number can be increased to one million. If you use a powerful computer, you can choose a larger number instead of a million.

The implemented version of the study of dynamic processes for the formation of primes includes the following indicators: the number of a simple number in the p in the ordered set of P , the value of a simple number of p , the value of the recursion length of the numbers $card_a(p)$ at the same value of a for all prime numbers P , the index $ind_a(p)$ of the index of the class:

$$ind_a(p) = (p - 1) / card_a(p)$$

The value of the residues modulo any natural module $n > 1$, for all classes and any other analytic properties of primes or factors of the decomposition of the number of $p - 1$ into simple factors. For each simple multiplier p_i in the:

$$p-1 = \prod_{i=1}^n p_i^{\alpha_i}$$

Decomposition, one parameter of the dynamic process of generating primes is presented, with separate indicators that can be analyzed for any other indicators, the values for them are deduced by the modulus of the natural number $n > 1$. The only exception is $ind_a(p)$. The number of controlled indicators analyzed in the Excel environment can be expanded.

DYNAMICAL PROPERTIES OF THE PROCESSES OF THE FORMATION OF CLASSES PRIMES IN THE GENERALIZED ARTIN'S HYPOTHESIS

According to the idea of experimental mathematics on the first iteration, we proceed from hypothetically known data. But it is also the basis for obtaining experimental information on the basis of which the analytical methods of the theory of numbers yield an expanded representation of the hypothesis in the form H_i . It is possible that at the same time the hypothesis can be corrected or even rejected as not true. From the point of view of information technology in mathematics, the hypothesis H_i is used to develop from the point of view of deepening the experimental mathematics of the model of in-depth studies at the level I_1 .

The iterations process is continued until an analytically based solution of the generated hypothesis is obtained. Since the Artin's generalized hypothesis is considered in the paper, we present the results of the estimation of the constant $c(a, i)$ for the case $a = 4$ and $i = 2$. The number $a = 4$ is a perfect square, and therefore it cannot be a primitive root. In terms of Artin's generalized hypothesis, this is as interesting and important as in the case when a is an original root.

Based on the data presented in [6], we obtained estimates for $c(a, i)$ for $a = 2, 3, \dots, 9, 10$ and $i = 1, 2, \dots, 9, \dots$. It is shown that their values are stable for class $P(4, 2)$ i.e. class with $ind_4(p) = 2$ to within a fourth decimal place. They are presented in the table 1.

An analysis of the data in the tables shows that for these numbers Artin's hypothesis is true on the set of primes $|P| = 10^9$.

The estimates for the $c(a, i)$ constants given in table 2 have the unique $i = 1$ property, which is that for $a \in \{2, 3, 5, 6, 7, 10, 11, 12, 13, 14, 15\}$ they coincide with the accuracy of the third decimal place. An analysis of

these classes will make it possible to establish that these classes have the same number of common primes. For any pair of $P(a_i,1) \cap P(a_j,1)$ at $i \neq j$, the values of:

$$|P(a_i,1,x) \cap P(a_j,1,x)|/|P(x)| = 0.1473$$

at $x = 179424673$ take the same value, with the exception of the pair:

$$|P(3,1,x) \cap P(12,1,x)|/|P(x)| = 2 \cdot 0.1473$$

This means that all these sets have the same number of common identical primes. An analysis of this fact shows that the formation of classes for:

$$a \in \{2,3,5,6,7,10,11,12,13,14,15,16,17,29,53\}$$

In the generalized Artin's hypothesis is subject to the same mathematical laws. Cases $a \in \{4,9\}$ do not relate to this fact because, according to Fermat's theorem, these values are not primitive roots. Special attention should be paid to the case of a pair $(3,12)$ for which classes:

$$P(3,1,x) \cap P(12,1,x)$$

with probability 0.8 regarding these classes, consist of the same prime numbers. The study of these facts will require the creation of new methods of mathematical analysis of the formation of classes in the classical and, therefore, generalized Artin's hypotheses. The solution to this problem is beyond the scope of this work.

The data in table 2 and table 3 allow us to make an important conclusion that there are many primitive roots for which the generalized Artin's constant $c(a,1)$ is equal to the same value 0.3739.... In addition, from the same table it follows that all pairs of $(P(a,i), P(a,j))$ sets have sets of common primes of the same power. Deepening research in this direction will parallelly create the theoretical basis for solving the classical Artin's hypothesis. The generalized Artin's hypothesis for all classes $P(a,1), \dots, P(a,i), \dots$ will require additional studies based on probabilistic computer simulation on the set of prime numbers of data beyond the limits of the first hundred million.

The results of experimental mathematics in table 1 of the first iteration confirm that Artin's hypothesis is correct. The estimates of the constants are obtained with the accuracy of the third decimal place. For $a \in \{2,3,5,6,7,8,10\}$ the:

$$\sum_{i=1}^{\infty} c(a, i) = 1$$

and for $a \in \{4, 9\}$ all $c(a, 2i + 1) = 0$ and:

$$\sum_{i=1}^{\infty} c(a, 2i) = 1$$

This is due to the fact that for all $a = k^2$ this is true because they are primitive roots of $(Z/pZ)^*$ groups, but primitive roots of their subgroups with even indices [3]. The results obtained are the basis for constructing an analytical proof of Artin's hypothesis and its generalization. The $c(a, 1)$ ratings given in the table for the set of primitive roots $\{2, 3, \dots, 16\}$ are obtained for the first time based on the results of computer simulation. The literature is known estimation $c(2, 1)$, which, starting from the fourth decimal place, is estimated analytically incorrect, due to the fact that the formula:

$$c(2, 1) = \prod_{p \in P} \left(1 - \frac{1}{p \cdot (p-1)} \right)$$

It is not true, because it includes all primes and among them those primes for which $a = 2$ is not a primitive root [5]. An important result is the creation of a computer model of the process of forming classes $P(a, 1), \dots, P(a, i), \dots$. For any values of $a > 1$, the interactions between the classes Table 2 and Table 3 are investigated (as a continuation). The first estimates were $c(a, i)$ for $i \geq 2$, and it was established that the statement that $c(a, i)$ is proportional to $1/i^2$ is absolutely false [1]. Obtaining the results is the basis for further deepening research on the Artin's hypothesis using analytical methods. In accordance with the developed mathematical model for the formation of primes on the base $a > 1$ and the calculated values of the generalized constants $c(a, i)$ for $i \geq 1$, as a result of computer simulation it was established that the generalized hypothesis is true. Tables 1, 2, 3 show the values of the Artin's constants, the relationships between classes, the dynamics of the formation of classes and its properties on the set of all primes P . The first column of Table 2 contains the values of the Artin's constants for the antiderivatives of the set $\{2, 3, 5, 6, 7, 8, 9, 10, 11, 12\}$. Actually, the modeling of $P(a, i)$ classes was carried out for many:

$$a \in \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 29, 53\}$$

Numbers $a \in \{4,9,16\}$ as squares of other numbers according to Fermat's theorem [2] cannot be primitive (primitive) roots of $p \in P$, and, accordingly, of residue groups $(\mathbb{Z}/p\mathbb{Z})^*$ modulo p . Particular attention was paid to the numbers $\{5,13,17,29,53\}$ due to the fact that they belong to the class of numbers of the Chebyshev's type [3] that is, they have representations $p = 4k + 1$, while $p \in P$, and the number n is a natural number. According to Chebyshev's assumption, the behavior of these numbers in residue classes modulo a prime number should differ from other primes.

To solve the problem of modeling classes of primes for a given base and evaluating the generalized constants of Artin's $c(a, i)$, an Excel-based software package was created that allows you to extend the modeling process to any natural numbers $a > 1$ and any set of consecutive primes whose power is a multiple of 500,000. This is the number of primes was chosen for the reason that it is statistically represented and provides an accurate representation of the dynamic processes of the formation of classes $P(a, i)$. Table 1 shows a fragment of the modeling process for $a \in \{2,3,5,8,12\}$ values. $a = 2$ is included in this set for the reason that it can be verified that the estimate [5,6] is different from the exact value. The difference begins with the third decimal place. This fact is important due to the fact that expression (4), although from an asymptotic point of view is close to the exact value of $c(2)$, nevertheless, it does not take into account all the features of the formation of classes $P(a, 1)$ for $a = 2$. The number $a = 5$ is included in the analysis of the dynamics, because $a = 5 = 4 \cdot 1 + 1$ is the smallest Chebyshev's number, which is as sensitive as possible to the established fact that all $P(5, 10k + 5)$ classes for $k \geq 0$ are empty. This is true for all Chebyshev's numbers. The proof of this fact is of a number-theoretic nature, and therefore, is excluded from consideration. The number $a = 8$ is included in the representation of dynamics for the reason that the dynamic properties of the classes $P(8, i)$ are radically different from the other classes studied. In particular, it was established that if $a = 8$ is the primitive root of $p \in P$, then $a = 2$ is also the primitive root of the same prime number. Conversely, if $a = 2$ is the primitive root of $p \in P$, then $a = 8$ will be either the same primitive root of p or $p \in P(8, 3)$. This is completely new information about the generalized Artin's constants; the developed information technologies have become the basis of fundamentally new results from

modern number theory, and as a consequence of modern cryptography. The numbers $a = 3$ and $a = 12$ are included in table 1 for the reason that $P(3,1)$ and $P(12,1)$ contain 0.8 common primes, while for any other pairs of sets $P(a_l,1)$ and $P(a_s,1)$, the total fraction of common primes is 0.4 for $l \neq s$ from the considered set of values. These two facts are obtained on the basis of the analysis of the information presented in table 3. This result was also based on the methods of modern number theory and probability theory.

Table 1. The quantity of prime numbers into intervals for $a=2,3,5,8,12$

Interval / a	2	3	5	8	12
0 – 0.5	187111	187011	196980	112331	187013
0.5 – 1.0	186912	186948	196836	112075	187057
1.0 – 1.5	186953	186960	197030	112175	187040
1.5 – 2.0	186846	186856	196894	112201	186958
2.0 – 2.5	187410	186896	196720	112345	186792
2.5 – 3.0	186711	186777	196957	112042	186767
3.0 – 3.5	187096	186926	197025	112335	187157
3.5 – 4.0	186975	187176	196942	112283	186984
4.0 – 4.5	187197	187148	196543	112296	187317
4.5 – 5.0	186713	186796	196689	121919	186721
5.0 – 5.5	186828	187013	197050	112093	187005
5.5 – 6.0	187197	186771	196790	112362	186936
6.0 – 6.5	186881	187116	196851	112226	187056
6.5 – 7.0	187065	187214	196478	112093	187122
7.0 – 7.5	187039	186718	196957	112236	187050
7.5 – 8.0	187045	186756	196764	112128	187161
8.0 – 8.5	187299	186805	196840	112187	186594
8.5 – 9.0	186663	187050	196583	111967	187144
9.0 – 9.5	186874	187156	196795	112133	186976
9.5 – 10.0	187034	187072	197083	111993	186947

In conclusion, by returning attention to table 1c of another theory of vision. The essence of a fundamentally new fact is that wherever 500,000 primes $p \in P$ are selected for any $a >$, the number of primes in classes ranges from no more than 500, which is no more than a thousandth of them. This means that on any set of consecutive primes we obtain an estimate of the Artin's constants up to the fifth decimal place. Statistical summation of values over the entire set of the first ten million primes made it possible to obtain estimates of the constants $c(a,1)$ to the eighth decimal place.

It follows that the methods of computer simulation of the processes of forming classes of primes $P(a,1), P(a,2), \dots, P(a,i), \dots$ and estimation of constants $c(a,1), c(a,2), \dots, c(a,i), \dots$ are the basis for the development of information technologies in modern both pure and applied mathematicians.

Table 2 shows the values of the estimates of the generalized Artin's constants for the marked set of values of a , which were studied as classifiers of the set of all primes. Table 2 shows a fragment of the entire huge number of obtained estimates of the Artin's constants. The first column contains estimates of Artin's constants in its original form. They belong to primitive roots. The numbers 4, 9, 16 are not primitive roots, since according to Fermat's theorem they, like the squares of other numbers, cannot be primitive roots. However, they can be classifiers of primes as roots of subgroups of residues modulo primes. An interesting fact is that they can be used to build pseudo-random number generators. In addition, the diskette logarithm problem can be considered on their basis.

Table 2. The distribution of prime numbers in 1 to 10 classes in the generalized Artin's conjecture

a	$P(a,1)$	$P(a,2)$	$P(a,3)$	$P(a,4)$	$P(a,5)$	$P(a,6)$	$P(a,7)$	$P(a,8)$	$P(a,9)$	$P(a,10)$
2	0,374	0,280	0,066	0,046	0,018	0,049	0,008	0,035	0,007	0,014
3	0,373	0,299	0,066	0,056	0,019	0,033	0,008	0,014	0,007	0,015
4	0	0,560	0	0,093	0	0,099	0	0,070	0	0,028
5	0,393	0,265	0,070	0,066	0	0,047	0,009	0,016	0,007	0,028
6	0,374	0,280	0,066	0,074	0,018	0,049	0,008	0,014	0,007	0,014
7	0,374	0,282	0,066	0,068	0,018	0,050	0,008	0,017	0,007	0,014
8	0,224	0,168	0,199	0,028	0,011	0,149	0,005	0,021	0,022	0,008
9	0	0,598	0	0,112	0	0,066	0	0,028	0	0,030
10	0,374	0,280	0,066	0,071	0,018	0,049	0,008	0,016	0,007	0,014
11	0,374	0,281	0,066	0,069	0,018	0,050	0,008	0,017	0,007	0,014
12	0,374	0,299	0,066	0,056	0,018	0,033	0,009	0,014	0,007	0,015
13	0,376	0,278	0,067	0,069	0,019	0,049	0,009	0,017	0,007	0,014
14	0,373	0,280	0,066	0,070	0,018	0,049	0,008	0,017	0,007	0,014
15	0,373	0,279	0,066	0,070	0,018	0,050	0,008	0,017	0,007	0,015
16	0	0,374	0	0,186	0	0,066	0	0,140	0	0,018
17	0,375	0,279	0,066	0,069	0,019	0,049	0,009	0,017	0,007	0,014
29	0,374	0,280	0,066	0,070	0,018	0,049	0,008	0,017	0,007	0,014
53	0,374	0,280	0,066	0,070	0,019	0,049	0,009	0,017	0,007	0,014

An interesting result is the equality of the constants $c(2,1) = c(3,1) = c(6,1) = c(7,1) = c(10,1) = \dots = c(15,1) = c(17,1) \dots$ up to one thousandth, although $c(8,1)$ and $c(5,1)$ are radically different. On the basis of modern number theory and the theory of random processes, the validity of such results is proved. Evidence of these allegations of remoteness is built only on the basis of data obtained by computer simulation. When analyzing the data, an assumption arose that the constructed classes for primitive roots have common primes. Table 3 shows the results of the analysis of sets of classes $P(a,1)$ for all pairs of primitive roots that were obtained using the constructed filter system. It turned out that all pairs of primitive roots have the same number of common primes with great accuracy. However, the classes $P(3,1)$ and $P(12,1)$ have exactly twice as many primes. This fact

is strictly mathematically justified. Note that other sets of primes of the form $P(a, k)$ with values greater than unity were not the object of even a brief analysis, since in order to obtain their exact values it is necessary to increase the number of primes analyzed, at least by an order of magnitude. This is due to the fact that they are found much less frequently in the set of primes.

Table 3. The intersection of prime numbers with $a=2..16$ and $a=2..10$

a\base	2	3	4	5	6	7	8	9	10
2	0,3740	0,1473	0	0,1619	0,1474	0,1473	0,2243	0	0,1328
3	0,1473	0,3739	0	0,1619	0,1474	0,1500	0,1020	0	0,1474
4	0	0	0	0	0	0	0	0	0
5	0,1619	0,1619	0	0,3937	0,1620	0,1620	0,1120	0	0,1620
6	0,1474	0,1474	0	0,1620	0,3741	0,1474	0,1020	0	0,1474
7	0,1473	0,1500	0	0,1620	0,1474	0,3741	0,1019	0	0,1474
8	0,2243	0,1020	0	0,1120	0,1020	0,1019	0,2243	0	0,0919
9	0	0	0	0	0	0	0	0	0
10	0,1328	0,1474	0	0,1620	0,1474	0,1474	0,0919	0	0,3741
11	0,1474	0,1483	0	0,1620	0,1474	0,1476	0,1020	0	0,1473
12	0,1473	0,2947	0	0,1619	0,1474	0,1500	0,1020	0	0,1474
13	0,1492	0,1493	0	0,1639	0,1492	0,1493	0,1033	0	0,1493
14	0,1474	0,1474	0	0,1619	0,1499	0,1474	0,1020	0	0,1473
15	0,1473	0,1327	0	0,1619	0,1474	0,1471	0,1020	0	0,1474
16	0	0	0	0	0	0	0	0	0

When analyzing the data, an assumption arose that the constructed classes for primitive roots have common primes. Table 3 shows the results of the analysis of sets of classes $P(a, 1)$ for all pairs of primitive roots that were obtained using the constructed filter system. It turned out that all pairs of primitive roots have the same number of common primes with great accuracy. However, the classes $P(3, 1)$ and $P(12, 1)$ have exactly twice as many primes. This fact is strictly mathematically justified. Note that other sets of primes of the form $P(a, k)$ with values greater than unity were not the object of even a brief analysis, since in order to obtain their exact values it is necessary to increase the number of primes analyzed, at least by an order of magnitude. This is due to the fact that they are found much less frequently in the set of primes.

CONCLUSION

Based on the analysis of the processes of formation of classes of primes for any bases, fundamentally new information technologies for solving complex mathematical problems by the methods of modern experimental mathematics were created. The correctness of the developed approach and computational efficiency are proved. A generalized theory of Artin's hypothesis has been developed which its classical version is a very special case. Estimates of the Artin's constants for bases greater than two are

obtained, and the statistical validity of the estimates obtained is proved. A detailed analysis of the classes of primes is carried out and the foundations of effective methods for the structural analysis of classes are created. It is proved that a new method for modeling the dynamics of the formation of classes of primes and a description of their properties creates the basis for constructing more advanced models of pseudo-prime generators, the development of new methods of information protection in modern cryptography, opens up new possibilities for constructing models of nonlinear dynamic systems.

REFERENCES

- [1]. Ambrose D. (2014). On Artin's Primitive Root Conjecture. Dissertation zur Erlangung des mathematisch -naturwissenschaftlichen Doctorgrades "Doctor rerum naturalium" der Georg-August-Universität Göttingen. – 169 p.
- [2]. Artin E. (1982). Collected papers. Edited by Serge, Lang and T. John, Springer-Verlag, New York. – 467 p.
- [3]. Crandall R., Pomerance C. (2005). Prime Numbers A Computational Perspective. Springer, Portland. – 659 p.
- [4]. Manin Yu. I., Panchishkin A. A. (2009). Introduction to the modern theory of numbers. MTSNMO, Moscow. – 547 p.
- [5]. Hooley C. (1973) Application of sieve methods to the theory of numbers. Cambridge, London, q. – 234 p.
- [6]. Moree P. (2012). Artin's Primitive root conjecture a survey, arXiv: math/0412262v2. – 86 p.
- [7]. Vostrov G., Opiata R. (2019). A generalized probabilistic model of computer proof of the Artin's hypothesis, International Symposium Computer Data Analysis and Modeling Stochastic Processes, Minsk. – P. 240– 247.
- [8]. Zagier D., (1977). First 50 million prime numbers, Math Intell.– P. 42–71.

КОМПЬЮТЕРНОЕ МОДЕЛИРОВАНИЕ ПРОЦЕССОВ РАЗВИТИЯ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ В ДИНАМИЧЕСКИХ ПРОЦЕССАХ ФОРМИРОВАНИЯ КЛАССОВ ОБОБЩЕННОГО ГИПОТЕЗА АРТИНА

Востров Г., Опиата Р.

Исследована взаимосвязь процессов формирования классов простых чисел в обобщенной гипотезе Артина и теории информации, и, как следствие, информационных технологий. Доказано, что вероятностные методы теории информации и информационных технологий являются основой для построения компьютерных моделей классов простых чисел в соответствии с обобщенной гипотезой Артина. Разработаны методы расчета артинских констант и установлена сходимость оценок констант по вероятности для ограничения значений. Созданы основы теоретико-числового анализа констант Артина и родственных классов.

Ключевые слова. Обобщенные классы Артина, постоянные Артина, классовые вероятности, устойчивость оценок постоянных Артина, сходимость по вероятности

УПРАВЛІННЯ СМУГОВИМ ЧАСТОТНО-ЗАЛЕЖНИМ КОМПОНЕНТОМ ІНФОРМАЦІЙНО-УПРАВЛЯЮЧОЮ СИСТЕМОЮ ДЛЯ УСУНЕННЯ ДЕТОНАЦІЇ ДВИГУНА ВНУТРІШНЬОГО ЗГОРАННЯ

Ухіна Г., Теплечук А., Кірьязов Ю., Ситніков В.

У роботі розглянуті підходи усунення детонації двигуна внутрішнього згорання. Виявлено недолік базових рішень, для підвищення ефективності та економічності системи запропоновано використовувати систему керування на основі зміни коефіцієнта надлишку повітря.

Це дозволило зменшити апаратні витрати, габаритні розміри, і як наслідок зменшити собівартість системи в цілому. Головні переваги цієї системи: універсальність, підвищення економічності двигуна і використання одного смугового цифрового фільтру. Ідея керування при зміні коефіцієнта надлишку повітря полягає у наступному.

При появі детонації система керування збільшує коефіцієнта надлишку повітря, це переводить роботу двигуна в режим збіднення робочої суміші. При цьому виникає падіння потужності двигуна, що призводить до зниження навантаження на двигун, що приведе не тільки до зникнення детонації, а і до економії палива і зниженню CO, NOX та СН. Оскільки основною задачею смугового фільтрує розпізнавання детонації і власних шумів двигуна на різних частотах, знайдені формули для розрахунку нових значень коефіцієнтів цифрового смугового фільтру другого порядку. Таке рішення дозволило зменшити похибки при роботі блоку виділення детонації на початкових етапах зародження детонації, що дає змогу оперативніше приймати рішення на усунення небажаної детонації.

Ключові слова. Коефіцієнт надлишку повітря, датчики детонації, обробка сигналів, кут випередження запалювання, цифровий смуговий фільтр, перебудова коефіцієнтів фільтру.

ВСТУП

Сучасний стан розвитку гібридних та електричних автомобілів не знімає проблеми підвищення економічності бензинового двигуна внутрішнього згорання і зниження токсичних газів у вихлопах автомобіля, що є однією із важливих задач при розробці та експлуатації автомобіля. Несумісність цих показників приводить до ускладнення системи керування двигуном.

Підвищення економічності залежить від значення коефіцієнта надлишку повітря α [1]

де G_T – вага палива, яке подається у циліндри ДВЗ,

G_B – вага повітря, яке подається на такті впуску у циліндри ДВЗ,

L – стехіометричний склад пальної суміші (на 1 кг палива приходить 14,5 кг повітря).

Для роботи двигуна в режимі економічності палива коефіцієнт надлишку повітря α повинен мати значення $\alpha = 1,1 \div 2,0$, а для роботи двигуна в режимі повної потужності N_e^{\max} та максимального крутного моменту $M_e^{\max} - \alpha = 0,8 \div 0,9$, що підвищує ккд (η_e) двигуна. Однак при роботі двигуна в названих режимах при використанні неякісного палива в циліндрах з'являються режими згорання з високими швидкостями порядку (1000-2000) м/с, які приводять до зносу робочих поверхонь поршневої групи та газорозподільного механізмів [2-10], рис.1. Ці режими згорання, які одержали назву детонація, характеризуються наступними параметрами:

- швидкістю детонації;
- масовою швидкістю продуктів реакції за ударною хвильою;
- розподіленою температурою, тиском та густиною суміші безпосередньо за ударною хвильою;
- термодинамічними властивостями палива та його продуктів реакції, які залежать від фізико-хімічних властивостей.



Рис.1 – Наслідки детонації

ДЕТОНАЦІЯ ТА СПОСОБИ ЇЇ УСУНЕННЯ

Для вимірювання швидкості детонації використовують фотографічні камери, які дозволяють безперервно реєструвати переміщення полум'я газоподібних продуктів реакції та ударних хвиль, що виникають у середині та навколо заряду, який детонує. Це дозволяє вивчати процеси хімічних реакцій, які визивають детонації.

Для усунення цих режимів запропонуються різні підходи. Так, наприклад, є пропозиції охолоджувати паливноповітряну суміш перед подачі її до циліндрів двигуна [11], або обробляти цю суміш акустичною хвилею [12]. Але частіше цю проблему усувають за допомогою зменшення кута випередження запалювання або зменшення подачі палива. Так, наприклад, на рис. 2, показані індикаторні діаграми процесів у циліндрах двигуна внутрішнього згорання (ДВЗ).

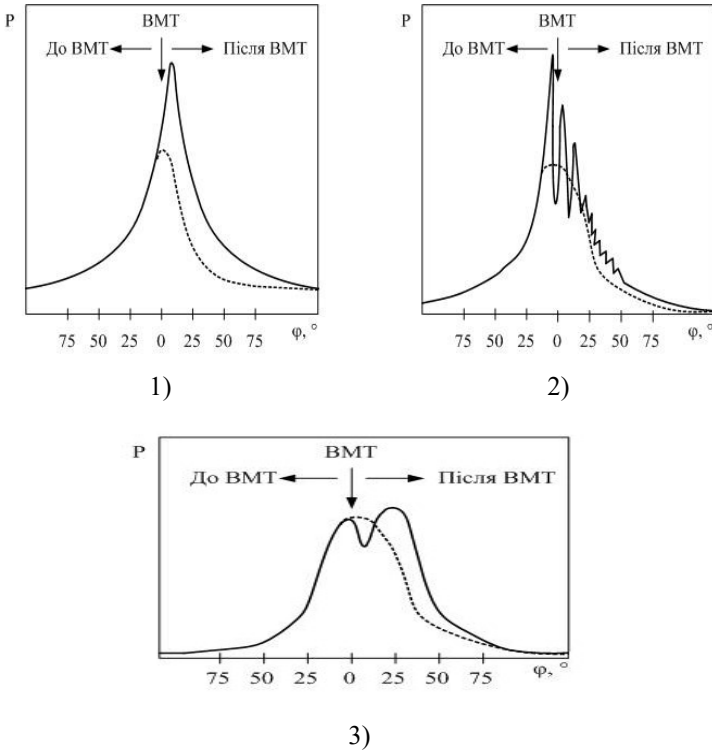


Рис. 2 – Індикаторні діаграми процесів у циліндрах ДВС
 1 – при оптимальному запалюванні; 2 – при згоранні з детонацією;
 3 – при пізньому запалюванні, де VMТ – верхня мертва точка

У сучасних автомобілях використовують систему електронного керування упорскуванням – BOSH MotronicME7.1.1, що дозволяє реалізувати високу потужність двигуна W12 при мінімальних витратах пального за рахунок узгодження режимів його роботи з умовами експлуатації. Електронне регулювання кута випередження запалювання узгоджується з сигналами від датчиків детонації, яких встановлюється

чотири штуки на дванадцять циліндрів, рис.3. По цим сигналам електронний блок управління двигуном розпізнає циліндр, що детонує.

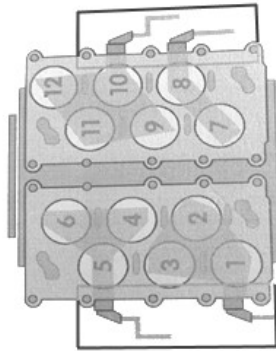


Рис. 3 – Схема встановлення датчиків детонації на блоці циліндрів двигуна W12.

В загальному випадку датчик детонації встановлюється [10] на блоці циліндрів і розпізнає коливання [5], які виникають при детонації. Обробка цих сигналів у електронному блоці керування дозволяє, в залежності від ступеня детонації, виробляти команди на зміну кута випередження запалювання. Загальна структурна схема системи контролю детонації, яка використовується у цих системах, наведена на рис. 4.

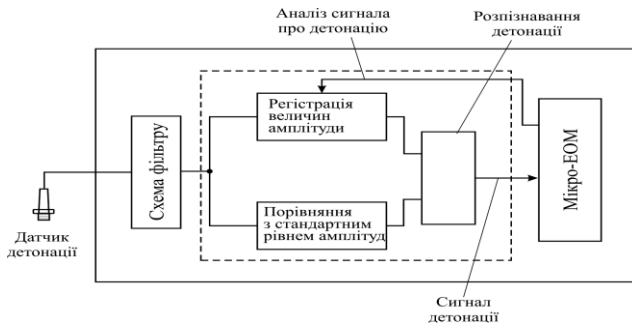


Рис. 4 – Структурна схема системи контролю детонації

Для детонації характерна поява у високочастотній частині спектру частот складової з високою амплітудою, рис. 2, яка виділяється за допомогою смугового фільтру. Шляхом віднімання поточного сигналу і середньої величини амплітуди сигналу датчика при відсутності детонації виконується розпізнавання детонації. Ступінь детонації визначається підрахунком числа амплітуд, величина яких перевищує стандартну

величину, характерну для початку детонації. Після розпізнавання детонації в залежності від її ступеня проводиться зменшення кута випередження запалювання. Якщо після цього детонація зникає, кут випередження запалювання поступово підвищується, тобто керування кутом випередження запалювання проводиться таким чином, щоб він був близький до детонаційного краю [2].

У стандартних блоках керування двигуном, автомобілів Impreza (концерн Subaru) реалізована активна система регулювання, яка дозволяє в нормальному режимі роботи двигуна вносити в кут випередження запалення поправки величиною приблизно від 3° до $+12^\circ$. Так, для турбированих двигунів нові блоки керування здатні оперувати істотно великими поправками. Застосування аналогових систем керування та мікросхем на їх основі не дозволило підвищити перешкодозахищеність та якість керування. Для підвищення цих показників використовують комбіновані аналого-цифрові системи керування. Так, наприклад, для виявлення детонації в двигунах внутрішнього згорання автомобільні концерни застосовують сигнальні процесори детонаційного шуму [13]. Ці процесори здійснюють аналогову обробку сигналів, але застосування цієї мікросхеми передбачає установку на її виході зовнішнього АЦП і мікроконтролера, який керує коефіцієнтами фільтрів мікросхеми та виробляє сигнали керування до системи запалювання (рис. 5).

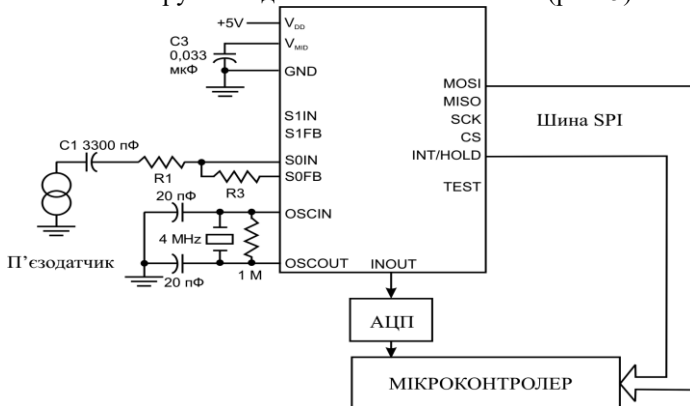


Рис. 5 – Схема системи контролю детонації на основі сигнального процесора детонаційного шуму HIP9010AB

Сигнал з датчика детонації в момент відсутності детонації надходить на один канал мікросхеми (для виявлення власних шумів двигуна), а в момент детонації сигнал надходить на інший канал мікросхеми. Обидва сигнали надходять на програмований інтегратор, у якому відбувається виділення сигналу детонації і його інтегрування. Підсилений сигнал надходить до АЦП, з виходу якого подається на

мікроконтролер. Мікроконтролер приймає інформацію, обробляє її і подає сигнали керування на електронний блок запалювання. Мікроконтролер змінює коефіцієнти фільтрів мікросхеми відповідно до частоти обертів колінчастого вала двигуна (рис. 6). Розглянуті вище рішення для усунення детонації ДВЗ базуються в основному на керуванні кутом випередження запалювання, як найбільш прості рішення, але не зовсім ефективні та економічні.

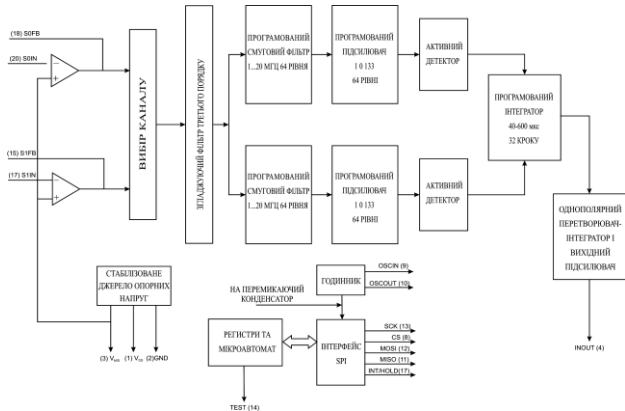


Рис. 6 – Структурна схема сигнального процесора NIP9010AB

УСУНЕННЯ ДЕТОНАЦІЇ НА ОСНОВІ ЗМІНИ КОЕФІЦІЄНТА НАДЛИШКУ ПОВІТРЯ

Сучасний розвиток спеціалізованих комп'ютерних та інформаційно-керуючих систем дозволяє побудувати іншу систему керування на основі зміни коефіцієнта надлишку повітря α . Розробка такої системи дозволила зменшити апаратні витрати, габаритні розміри, і як наслідок зменшити собівартість системи в цілому. Головні переваги цієї системи: універсальність, підвищення економічності двигуна і використання одного смугового цифрового фільтра.

Ідея керування при зміні коефіцієнта надлишку повітря α полягає у наступному. При появі детонації система керування збільшує α до $1,1 \div 2,0$, це переводить роботу двигуна в режим збіднення робочої суміші. При цьому виникає падіння потужності двигуна, що змусить водія перейти на нижчу передачу, а автомобіль, обладнаний автоматичною коробкою передач, зробить це сам не зменшуючи комфорту подорожі і тим самим знизити навантаження на двигун, що приведе не тільки до зникнення детонації, а і до економії палива і зниженню CO, NOX та CH.

При побудові такої системи спрощується її структурна організація, яка близька до типової, рис 7.

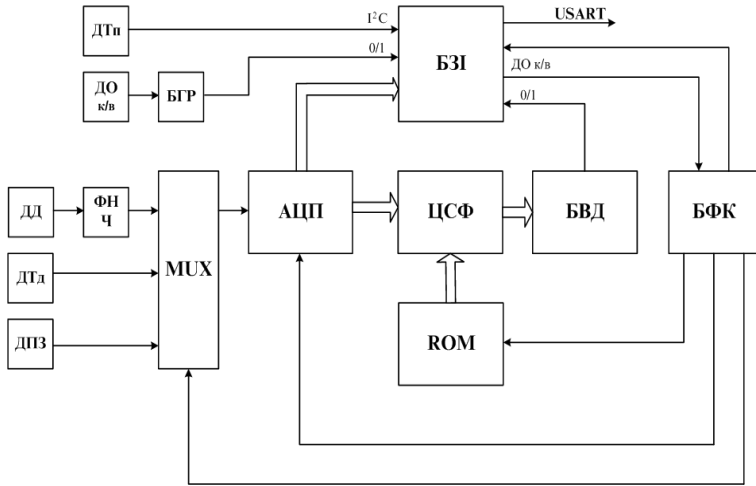


Рис. 7 – Структурна схема системи відстеження детонації

ДТп – датчик температури повітря; ДО к/в – датчик обертів колінчастого валу; ГР – блок гальванічної розв’язки; БЗІ – блок збору інформації; ДД – датчик детонації; ДТд – датчик температури двигуна; ДПЗ – датчик положення дросельної заслінки; MUX – мультиплексор; АЦП – аналого-цифровий перетворювач; ЦСФ – цифровий смуговий фільтр; БВД – блок виявлення детонації; БФК – блок формування команд; ROM – блок пам’яті.

За командами блока формування команд (БФК) мультиплексор (МПС) підключає до АЦП необхідний датчик і оцифрований сигнал надходить до блоку збору інформації (БЗІ). Після БЗІ всі дані зняті з датчиків в цифровому вигляді за протоколом USART подаються в пристрій обробки інформації. Таким чином, пристрій знімає дані з усіх датчиків, які використовуються загальною системою управління коефіцієнтом α та кутом випередження запалення.

Сигнал у цифровому вигляді з датчика детонації (ДД) надходить на вхід цифрового смугового фільтра (ЦСФ). Після фільтрації сигналу блок виявлення детонації (БВД) аналізує одержаний сигнал, і результат передає до блоку збору інформації (БЗІ). Застосування датчика обертів (ДО) і ДД дозволило точно встановити моменти виникнення детонації і моменти її відсутності. Це дало можливість застосувати цифровий фільтр, який у моменти відсутності детонації відстежує власні шуми двигуна, а в моменти прояву детонації – суму сигналів детонації та

власних шумів двигуна. Різниця цих значень дозволяє виявити наявність детонації. Слід відмітити, що смуговий фільтр повинен відстежувати частоту обертів колінчастого вала двигуна. У цьому випадку є проблема в побудові смугового фільтру що перебудовується.

ДОСЛІДЖЕННЯ ВЛАСТИВОСТЕЙ І СТРУКТУР ЦИФРОВОГО ФІЛЬТРУ

Оскільки основною задачею смугового фільтрує розпізнавання детонації і власних шумів двигуна на різних частотах, то чим точніше задані коефіцієнти передавальної функції ЦСФ, тим точніше фільтр виділить сигнал, зв'язаний з детонацією. Тому найбільша увага приділялась дослідженню властивостей і структур цифрового фільтру при відстеженні частоти обертів колінчастого валу, що пов'язано зі змінами коефіцієнтів передавальної функції. Передавальна функція ЦСФ має вигляд:

$$H(z^{-1}) = \frac{a_0 + a_1 z^{-1} + a_2 z^{-2} + a_3 z^{-3} + a_4 z^{-4}}{1 + b_1 z^{-1} + b_2 z^{-2} + b_3 z^{-3} + b_4 z^{-4}}$$

В роботі [10] запропоновано підхід на основі попереднього розрахунку з заданим кроком коефіцієнтів передавальної функції. В цьому випадку зовнішня пам'ять (ПЗУ) зберігає розраховані коефіцієнти цифрового фільтру, які по команді блоку формування команд (БФК) передаються на цифровий смуговий фільтр (ЦСФ). Діапазон робочих частот знаходиться в межах між $f_H = 20$ Гц та $f_B = 200$ Гц, при смузі пропущення (заданий крок) \square Гц. Слід відмітити, що ці параметри обумовлені режимами роботи двигуна і появою детонації. Так при частоті обертання колінчастого валу 600 об/хв (режим холостого ходу) центральна частота смугового фільтра дорівнює 20 Гц. В даному режимі роботи двигуна детонація неможлива (по визначенню) і система аналізує власні шуми двигуна. Так, для центральної частоти $f_0 = 100$ Гц і частоти дискретизації $f_d = 4000$ Гц цифровий смуговий фільтр має АЧХ H_{id} , яка наведена на рис. 7.

У даному діапазоні робочих частот необхідно розрахувати коефіцієнти для $N = \frac{f_B - f_H}{\Delta f} = \frac{200 - 20}{2} = 90$ фільтрів.

Така організація не дозволяє виконувати плавне регулювання. Тому проводились дослідження можливості плавної перебудови ЦСФ.

Однак, при плавному регулюванні можливі випадки виходу за межі стійкості фільтру [14].

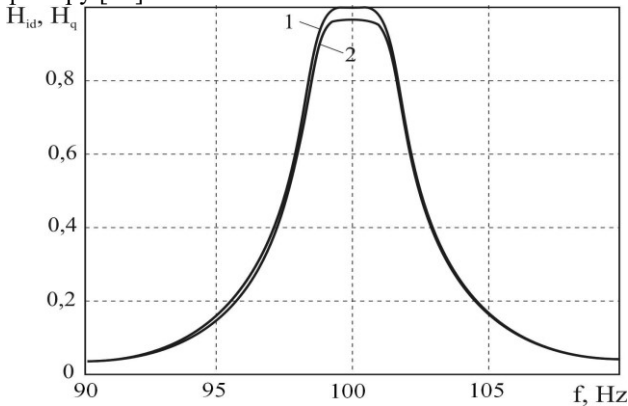


Рис. 8 – АЧХ смугового фільтру з ідеальними (1) та з округленими (2) коефіцієнтами на частоту $f_0 = 100$ Гц

Для зручності корекції частотної характеристики цифровий смуговий фільтр складається з двох смугових фільтрів Баттерворта першого порядку, які описуються передавальною функцією другого порядку. Для перебудови цифрового смугового фільтру другого прядку знайдені формули для розрахунку нових значень коефіцієнтів фільтру. Так, коефіцієнти знаменника b_1 та b_2 перераховуються від залежності частоті обертів колінчастого валу та необхідної смуги частот, яка налагоджується при підготовки двигуна. Так, коефіцієнт b_1 для частоти зрізу $\bar{\omega}_1$ визначається наступним чином

$$b_1 = -(1 + b_2) \cos(\bar{\omega}_1) + (1 - b_2) \sin(\bar{\omega}_1) \frac{\sqrt{1 - c^2}}{c}. \quad (1)$$

На основі частот зрізу смугового фільтру є можливість визначити величину коефіцієнту знаменника b_2 як

$$b_2 = \frac{(\cos(\bar{\omega}_2) + \cos(\bar{\omega}_1)) - 2c^2 \cos(\bar{\omega}_2) - 2c\sqrt{1 - c^2} \sin \cos(\bar{\omega}_2)}{(\cos(\bar{\omega}_2) + \cos(\bar{\omega}_1)) - 2c^2 \cos(\bar{\omega}_1) - 2c\sqrt{1 - c^2} \sin \cos(\bar{\omega}_1)}. \quad (2)$$

Однак, для реалізації таке представлення залежностей коефіцієнтів знаменника b_1 та b_2 від частот зрізу і рівня пульсації c не зовсім вдале

для реалізації на мікропроцесорній техніці. Введемо деяку фіктивну величину ξ так, щоб $c = \cos\left(\frac{\xi}{2}\right)$, або фіктивна величина дорівнює $\xi = 2\arccos(c)$. В цьому випадку, співвідношення для коефіцієнтів мають наступний вигляд

$$b_1 = -(1 + b_2) \cos(\bar{\omega}_1) + (1 - b_2) \sin(\bar{\omega}_1) \operatorname{tg}\left(\frac{\xi}{2}\right);$$

$$b_2 = \frac{2 \cos(\bar{\omega}_2) + \cos(\bar{\omega}_1) - \cos(\bar{\omega}_2 - \xi)}{\cos(\bar{\omega}_2) + 2 \cos(\bar{\omega}_1) - \cos(\bar{\omega}_1 - \xi)}. \quad (3)$$

Таким чином, аналіз АЧХ цифрових смугових фільтрів показує, що при розробці смугових цифрових частотно-залежних компонент в залежності від заданих частот зрізу і рівня пульсацій C можна однозначно знайти значення коефіцієнтів чисельника і знаменника.

Коефіцієнти, які представляються з обмеженою розрядністю, обрані так, щоб максимум зваженої різниці частотних характеристик ідеального та реального фільтрів був мінімальним.

Проведений аналіз дозволяє значно полегшити вирішення поставленого завдання розробки цифрових смугових частотно-залежних компонент інформаційно-управляючої комп'ютерної системи з можливістю корекції і перебудови характеристик компоненти і системи в цілому.

Для реалізації цифрового смугового фільтру обрана програмована логічна інтегральна схема (ПЛІС) SPARTAN3 XC3S400-4-PQ20S фірми Xilinx [15], та обраний мікроконтролер фірми ATMEL MEGA128, яка поєднує в собі аналогову та цифрову частини схеми.

Для мікроконтролера фірми ATMEL MEGA128 розроблений алгоритм та написана програма обробки датчиків та системи керування коефіцієнтом надлишку повітря α . Це дало змогу зменшити апаратну частину системи керування, так як усі датчики (рис. 7) підключаються до АЦП, який знаходиться у мікроконтролері, а також зменшити час обробки даних з АЦП тому, що він знаходиться на одній шині даних з процесором.

Для побудови цифрового фільтру другого порядку виконано каскадне з'єднання двох смугових фільтрів першого порядку. Це

зроблено для більш зручної настройки та перебудови фільтру в цілому. Граф фільтру наведено на рис. 9.

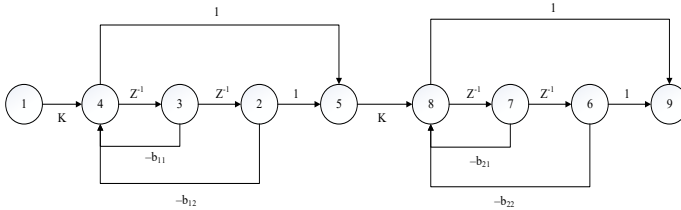
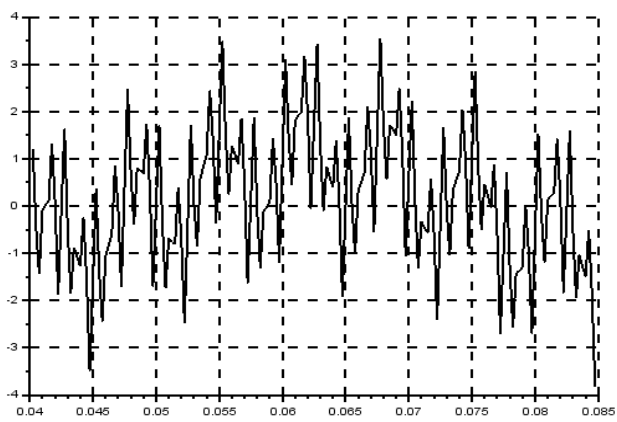


Рис. 9 – Сигнальний граф цифрового смугового фільтру другого порядку побудованого на основі каскадного з’єднання смугових фільтрів першого порядку

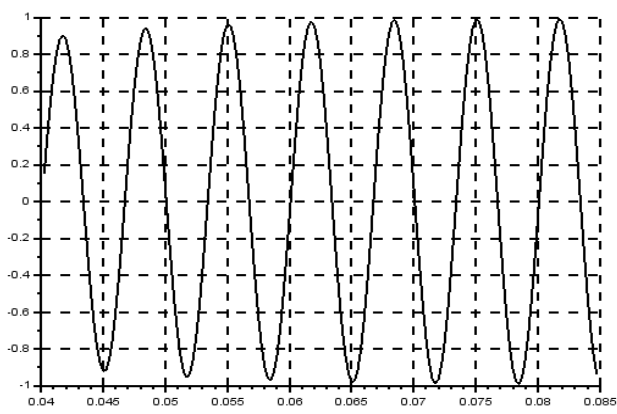
На основі сигнального графу складена система рівнянь роботи фільтру на основі якої побудовано алгоритм обчислень:

$$\begin{aligned}
 x_1[i] &= s(i); \\
 x_2[i] &= x_3[i-1]; \\
 x_3[i] &= x_4[i-1]; \\
 x_4[i] &= kx_1[i] - b_{12}x_2[i] - b_{11}x_3[i]; \\
 x_5[i] &= x_2[i] + x_4[i]; \\
 x_6[i] &= x_7[i-1]; \\
 x_7[i] &= x_8[i-1]; \\
 x_8[i] &= kx_5[i] - b_{22}x_6[i] - b_{21}x_7[i]; \\
 x_9[i] &= x_6[i] + x_8[i].
 \end{aligned}$$

Фрагмент моделювання даного фільтру показано на рис. 10



а)



б)

Рис 10 – Фрагмент вхідного (а) та вихідного (б) сигналів цифрового смугового фільтру з центральною частотою $f_0 = 150$ Гц

ВИСНОВКИ

При обробці сигналів датчиків у системі керування коефіцієнтом надлишку повітря α таке рішення дозволило зменшити похибки при роботі блоку виділення детонації на початкових етапах зародження детонації, що дає змогу оперативніше приймати рішення на усунення небажаної детонації.

Крім того, при розробці такої системи вдалося зменшити апаратні витрати, габаритні розміри, і як наслідок зменшити собівартість системи в цілому. Головні переваги цієї системи: універсальність, підвищення економічності двигуна і використання одного смугового цифрового фільтру.

ЛІТЕРАТУРНІ ДЖЕРЕЛА

- [1] Акимов А.В., Акимов О.А. Электрооборудование автомобилей / Под ред. Ю.П. Чижкова. М.: Транспорт, 1999. – 384 с.
- [2] Почему детонирует двигатель. URL: <http://krutimotor.ru/detonaciya-topliva-v-dvigatеле/>
- [3] Детонация двигателя. URL: <http://krutimotor.ru/detonaciya-dvs/>
- [4] Причины детонации двигателя при выключении зажигания и запуске. URL: <https://znanieavto.ru/dvs/detonaciya-dvigatelya.html>
- [5] Что такое детонация двигателя. URL: <https://pricurivatel.ru/detonaciya-dvigatelya.html>
- [6] Особенности технического обслуживания и диагностики электронных систем управления двигателем. URL: https://studref.com/423793/tehnika/osobennosti_tehnicheskogo_obslyzhvaniya_diagnostiki_elektronnyh_sistem_uchpravleniya_dvigatелеm
- [7] Детонация двигателя, датчик детонации. URL: <https://carspec.info/detonaciya>
- [8] Датчик детонации. URL: <http://krutimotor.ru/datchik-detonacii/>
- [9] Детонация в двигателях внутреннего сгорания. URL: <https://www.chem21.info/info/351591/>
- [10] Бадерко І.В., Дікусар К.В., Полонська Н.О., Ситніков В.С., Теплечук А.М. Електронне управління характеристиками бензинового двигуна внутрішнього згорання для усунення детонації. *Вісник Чернівецького ун-та: Комп'ютерні системи та компоненти*, 2011. – Т.2. – Вип. 1. – С. 92– 100.
- [11] Способы для предотвращения детонационного сгорания. Патент № RU 2636253. URL: <http://www.findpatent.ru/patent/263/2636253.html>
- [12] Способ предотвращения детонационного сгорания. Патент № RU 2146016. URL: <http://ru-patent.info/21/45-49/2146016.html>
- [13] Александр Зыбайло. Сигнальный процессор обнаружения детонации в двигателе внутреннего сгорания НР9010АВ. URL: http://www.compitech.ru/html.cgi/arhiv/00_04/stat_34.htm

[14] Ukhina H., Sytnikov V. Express-analysis of stability of fourth order frequency-dependent rearrangeable components International Conference on Electronics and Information Technology, EIT 2016 – Conference Proceedings 7500998

[15] Малахов В.П., Ситников В.С, Ядвичук П.В. Реализация элементов быстродействующего цифрового фильтра на ПЛИС. Технология и конструирование в электронной аппаратуре, 2004. – № 4. – С. 3 – 8.

FREQUENCY-FREQUENCY COMPONENT MANAGEMENT INFORMATION AND CONTROL SYSTEM TO REMOVE DETONATION OF ENGINE INTERNAL

Ukhina G., Teplechuk A., Kiryazov Y., Sitnikov V.

Approaches to eliminating internal combustion engine detonation are considered in the paper. A lack of basic solutions has been identified, and it is suggested to use a control system based on changes in the excess air factor to improve the efficiency and economy of the system. This made it possible to reduce hardware costs, overall dimensions and, as a consequence, reduce the cost of the system as a whole. The main system benefits are versatility, increased engine efficiency and the use of a single bandpass filter. The idea behind controlling the excess air ratio is as follows. When a detonation occurs, the control system increases the coefficient of excess air, this puts the engine into the lean mode. This results in a decrease in engine power, which leads to a decrease in the load on the engine, which will lead not only to the disappearance of detonation, but also to fuel savings and reduce CO, NOX and CH. Since the main purpose of the bandpass filter is to detect detonation and intrinsic motor noise at different frequencies, formulas have been found to calculate new values of second order digital bandwidth coefficients. The frequency response analysis shows that in the development of bandwidth digital frequency-dependent components, depending on the specified cutoff frequency and the level of ripple, you can uniquely find the values of the numerator and denominator coefficients. To construct a second-order digital filter, a cascade connection of two first-order bandpass filters is made. This is for easier setup and restructuring of the filter as a whole. A graph and the equations system of filter operation were constructed, on the basis of which the calculations algorithm was implemented. This solution made it possible to reduce the errors during the operation of the detonation unit at the initial detonation stages, which makes it possible to make a more promptly decision to eliminate unwanted detonation. In addition, the development of such a system managed to reduce hardware costs, overall dimensions, and as a consequence, reduce the cost of the system as a whole.

Keywords. Excess air ratio, detonation sensors, signal processing, ignition timing angle, digital bandpass filter, filter coefficients adjustment.

ІНФОРМАЦІЙНА ПІДТРИМКА ТЕХНІЧНОГО ОБСЛУГОВУВАННЯ СИСТЕМ ЗАЛІЗНИЧНО-ВОДНОГО СПОЛУЧЕННЯ

Казмиренко Ю., Дрозд О., Єгольников О.

Сучасний розвиток комп'ютерних технологій у обслуговуванні та експлуатації транспортних засобів створює передумови для реалізації нових механізмів управління системами технічної діагностики і моніторингу, застосування яких дозволить у автоматичному режимі контролювати технічний стан та формувати інформаційні повідомлення щодо відхилень показників від нормативних. Створення спеціалізованої інформаційної системи для вирішення завдань оцінювання технічного стану систем залізнично-водного сполучення є актуальним та спрямовано на удосконалення методу технічного діагностування і моніторингу пристроїв залізничної автоматики. Авторами запропоновано механізми проектування інформаційної системи, для чого проаналізовано існуючі комп'ютерні діагностичні системи, визначено недоліки їх роботи, причини виникнення та наслідки, до яких вони призводять. В роботі застосовано об'єктно-орієнтований підхід та використано класичні технології проектування баз даних.

Особливості експлуатації інформаційних систем в умовах залізнично-водного сполучення полягають у поєднанні функцій залізничного транспорту з водним, що враховується при побудові структури бази даних, в основу якої покладено завдання мінімізації дублювання даних і спрощення процедур їх обробки та поновлення. Програмне забезпечення дозволяє ефективно вирішувати дослідницькі та експлуатаційні завдання. Результати роботи можуть бути застосовані для тестування пристроїв залізничної автоматики на шляхах залізнично-водного сполучення.

Ключові слова. Залізнично-водне сполучення, системи технічного діагностування і моніторингу, інформаційна підтримка проекту.

ПОСТАНОВКА ПРОБЛЕМИ ТА ЇЇ ЗВ'ЯЗОК З НАУКОВИМИ ПРОГРАМАМИ

Успішна експлуатація систем залізнично-водного сполучення на припортових мережах безпосередньо пов'язана із впровадженням на проблемних ділянках систем технічного діагностування і моніторингу (СТДМ), що не можливо без інформаційної підтримки проектів. На даний час в Україні застосовані методи діагностування є морально застарілими. Розвинути цей напрямок можливо в результаті створення принципово нового наукового підходу до діагностування і моніторингу роботи пристроїв залізничної автоматики: їх безаварійна експлуатація вимагатиме застосування системного підходу з розробленням нової інформаційно-пошукової системи для обслуговування. Сучасний розвиток комп'ютерних технологій для систем моніторингу створює

передумови для реалізації нових механізмів управління засобами технічної діагностики, що відповідає пріоритетним напрямкам розвитку залізничного транспорту, визначених у Стратегії розвитку транспорту на період до 2020 р. (розпорядження Кабінету Міністрів України від 16.12.2009 р. № 1555-р) та Закону України «Про пріоритетні напрями інноваційної діяльності в Україні» № 3715-VI від 08.09.2011 р. (редакція від 05.12.2012 р.), ст. 4 «Стратегічні пріоритетні напрями інноваційної діяльності на 2011–2021 рр., п. 2 «Освоєння нових технологій високотехнологічного розвитку транспортної системи, ракетно-космічної галузі, авіа- та суднобудування, озброєння і військової техніки».

АНАЛІЗ ОСТАННІХ ДОСЛІДЖЕНЬ І ПУБЛІКАЦІЙ

Головними ознаками існуючих інформаційних діагностичних систем [1,2] є багатоаспектність, багатоструктурність, невизначеність їх поведінки, ієрархічність будови, структурна подібність, надлишковість основних елементів і підсистем, зав'язків між ними, багатоваріантність реалізації функцій управління на кожному з цих рівнів, територіальна розподіленість та мобільність компонентів, наявність режиму функціонування у реальному часі. За думкою автора цих робіт існування різних, але споріднених за функціональними можливостями програмних комплексів, призводить до порушення у координації процесами управління і моніторингу, а в умовах обмеження у фінансуванні та часі виникає необхідність знехтування вузькоспеціалізованих систем.

В роботі [3] запропоновано метод комп'ютерної діагностики визначення несправностей, пов'язаних з механічною та електричною частиною стрілочних переводів. Розроблене програмне забезпечення дає змогу відновлювати сигнал та будувати його часову характеристику, а також розкласти його спектральні складові для проведення аналізу отриманих даних. Підвисити функціональні можливості технологій моніторингу пристроїв залізничної автоматики можливо шляхом застосування Інтернет-систем [4], як це прийнято для судноплавства [5,6] і авіації [7], в основу створення яких покладено доменно-орієнтоване програмне забезпечення [9], розроблене на основі доменного аналізу [8]. Автори розглядають інформаційну підтримку як складну систему підтримки прийняття рішень (СППР), основними етапами створення якої є: технічне проєктування інформаційної технології; розробка методів доменного аналізу, теорії прийняття рішень, аналітичної обробки за технічним станом; проведення дослідної експлуатації ІТ-системи. Доцільність та перспективи застосування СППР на залізничному транспорті наведено у роботах [1,2,10,11]. Метою їх впровадження є підвищення оперативності, обґрунтованості і ефективності управлінського апарату за рахунок використання інформаційних технологій та оперативного формування на їх основі комплексної

оперативної інформації. Проте невирішеними залишаються питання створення інформаційних систем (ІС) для технічного діагностування і моніторингу пристроїв залізничної автоматики з урахуванням особливостей їх експлуатації та адаптації у мережах залізнично-водного сполучення.

МЕТА РОБОТИ

Встановити та дослідити механізми проектування інформаційної системи для технічного обслуговування мереж залізнично-водного сполучення.

Для досягнення поставленої мети в роботі необхідно вирішити наступні **завдання**:

1) проаналізувати існуючі діагностичні системи, які на даний час застосовуються для технічного обслуговування пристроїв залізничної автоматики; визначити недоліки їх роботи, встановити причини виникнення та наслідки, до яких вони призводять;

2) визначити цілі та задачі створення та експлуатації інформаційних систем для обслуговування мереж залізнично-водного сполучення;

3) надати техніко-економічне обґрунтування щодо необхідності впровадження інформаційних систем у мережах залізнично-водного сполучення.

ОБ'ЄКТ ДОСЛІДЖЕННЯ

Процеси обробки і систематизації даних пристроїв залізничної автоматики у мережах залізнично-водного сполучення.

ПРЕДМЕТ ДОСЛІДЖЕННЯ

Механізми та моделі проектування інформаційної системи для технічного обслуговування мереж залізнично-водного сполучення.

МЕТОДИ ДОСЛІДЖЕНЬ

В роботі застосовано об'єктно-орієнтований підхід та використано класичні технології проектування баз даних [12–14]. Інформаційну базу досліджень становлять нормативна документація Укрзалізниці, законодавча база України; наукові праці вітчизняних і зарубіжних вчених з розробки методів та систем технічного діагностування і моніторингу пристроїв залізничної автоматики [1–11], роботи в області обслуговування мереж залізнично-водного сполучення [15], системного аналізу та інформаційних технологій [16–20]. В основу техніко-економічного обґрунтування проекту покладено визначення соціально-економічних чинників застосування комплексу технічних і програмних засобів.

ВИКЛАД ОСНОВНОГО МАТЕРІАЛУ

Аналіз побудови та експлуатації інформаційних систем, які застосовуються для діагностування пристроїв залізничної автоматики.

В основу проектування інформаційної системи (ІС) покладено аналіз сучасних діагностичних систем (табл. 1), який пов'язує недоліки їх роботи, причини виникнення та наслідки, до яких вони призводять. Передбачливий технічний результат від застосування методів діагностування полягає у можливості оцінки реального технічного стану пристроїв залізничної автоматики та прийняття рішення щодо можливості їх подальшої експлуатації.

Таблиця 1. Аналіз сучасних інформаційних діагностичних систем
[складено авторами]

Недоліки роботи систем	Причини виникнення	Наслідки
Тимчасові затримки та помилки в управлінні	Невірні вирішення задач аналізу технічного стану	Відмови, невиконання задач, створення аварійних ситуацій
Відхилення від нормативних даних	Невиконання задач оцінювання та контролю функціонального стану пристроїв залізничної автоматики	Нештатні аварійні ситуації
Порушення функціональності технологій моніторингу	Порушення роботи датчиків, некомпетентність персоналу	Розлад у системі управління
Відсутність взаємодії інформаційних систем і систем моніторингу	Наявність різних, несумісних ІС, що функціонують на різних програмно-апаратних платформах	Відсутність єдиних механізмів контролю за інформацією

Вихідні данні включають у себе:

- інформацію про комплектування встановленого обладнання;
- інформацію про якість його роботи, яка надходить з:
 - журналу обліку роботи, технічного обслуговування та ремонту, обліку дефектів та відмов,
 - графіків та планів технічного обслуговування;
 - протоколів відомчих та Державних перевірок;
- інформацію про працездатність залізничної автоматики та суміжних з ними систем;

- інформацію з журналів обліку вимушених зупинок через пристрої залізничної автоматики за кількісними характеристиками з аналізом прийнятих рішень щодо їх усунення, для чого застосовуються: диспетчерські та змінні журнали за певний період обслуговування;
- інформацію про кількість та стан запасних інструментів та приладів;
- протоколи обстеження в процесі діагностування.

Цілі і задачі створення і експлуатації інформаційних систем для обслуговування мереж залізнично-водного сполучення.

Особливості експлуатації ІС в умовах залізнично-водного сполучення полягають у поєднанні функцій залізничного транспорту з водним, як це показано на діаграмі прецедентів (рис. 1), побудованій за допомогою універсальної мови моделювання (Unified Modelling Language або UML) [21], що дає можливість створювати програмні системи меншого розміру шляхом використання загальних механізмів, забезпечуючи необхідну економію засобів вираження. Побудована діаграма встановлює взаємний зв'язок між диспетчерськими службами залізничного терміналу та порту. До функціональних обов'язків диспетчера залізничного терміналу відносяться огляд розташування вагонів, їх графічна інтерпретація, огляд каргоплану судна, підтвердження заявок; введення заявок на подачу вагонів здійснюється техніком станції. Диспетчер порту збирає і вводить дані про судна та умови рейсу.

Головним призначенням систем технічної діагностики і моніторингу є контроль показників пристроїв залізничної автоматики, який повинен здійснюватися у автоматичному режимі. Інформаційні повідомлення щодо відхилень показників від нормативних формуються технологом. Одержані дані зазнають оброблення та використовуються для складання практичних рекомендацій. При проєктуванні нових інформаційних систем необхідно враховувати особливості умов експлуатації пристроїв залізничної автоматики у мережах залізнично-водного сполучення, такі як погодні умови (температура повітря в день та вночі; температура води, солоність та вологість повітря, швидкість вітру тощо). Ці дані заносяться до інформаційної системи, а потім формалізуються у вигляді електронного журналу. Як вихідні також застосовуються аналітичні дані роботи датчиків, експертна оцінка інформації у вигляді відхилень від нормативних.

На рис. 2 наведено складену структурну схему застосування інформаційних ресурсів для оцінювання технічного стану систем залізнично-водного сполучення, в основу якої покладено виконаний ґрунтовний аналіз діагностичних систем (табл. 1).

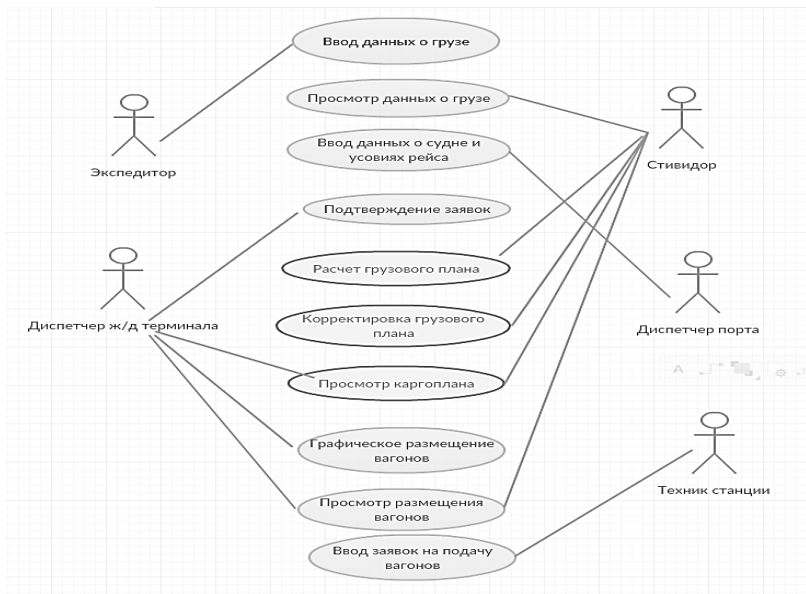


Рис.1 Діаграма прецедентів

Основні етапи проектування інформаційної системи включають у себе:

- вибір технології та засобів моделювання;
- розроблення функціональної моделі системи; побудову концептуальної та логічної моделей;
- розроблення математичної моделі;
- вибір інструментів розробки;
- побудову діаграм станів та діаграм класів;
- тестування програмного забезпечення системи.

Для розроблення інформаційної системи застосовано мову SQL (Structured Query Language) [12], Web-додатки написані мовою програмування PHP [13], перевагою якої є здатність генерувати не тільки HTML-документи, а й зображення різних форматів: JPEG, GIF, PNG, файли PDF та FLASH. В основу проектування структури бази даних покладені завдання мінімізації дублювання даних і спрощення процедур їх обробки та поновлення [14]. Робота з базою даних можлива у двох режимах: адміністратора та користувача.

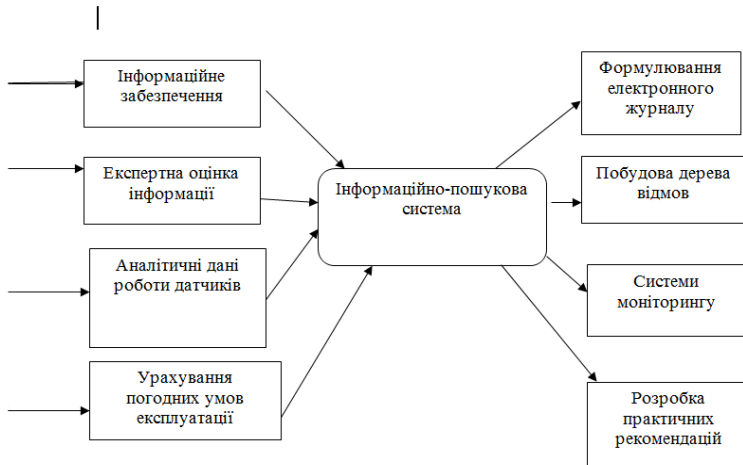


Рис. 2 – Структурно-логічна схема застосування інформаційних ресурсів для оцінювання технічного стану систем залізнично-водного сполучення

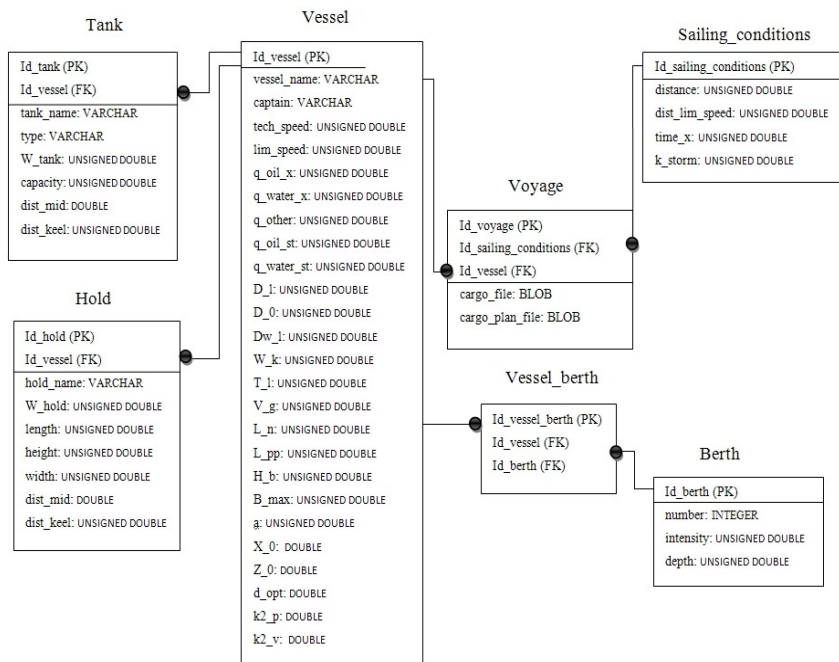


Рис. 3 – Фізична модель даних

Введення і редагування інформації здійснюється адміністратором, користувачеві надаються функції швидкого пошуку та порівняння, сортування показників за назвами і датою, перегляду повної інформації, яку представлено у вигляді електронного журналу. Клієнтська частина системи управління базами даних забезпечує інтерфейс зв'язку користувача з базою даних: перетворює запити в команди запитів до серверної частини. Доступ до бази даних здійснюється через web-сервер Apache. У рамках системи пошук може бути здійснений за запитом, що включають в себе конкретні числові показання електронних датчиків. Крім функцій перегляду, інформаційна система дає змогу здійснювати додавання даних і перегляд вмісту таблиць. Фізичну модель даних наведено на рис. 3.

ТЕХНІКО-ЕКОНОМІЧНЕ ОБҐРУНТУВАННЯ ЩОДО ВПРОВАДЖЕННЯ ІНФОРМАЦІЙНОЇ СИСТЕМИ У МЕРЕЖАХ ЗАЛІЗНИЧНО-ВОДНОГО СПОЛУЧЕННЯ ТА ПЕРСПЕКТИВИ ПОДАЛЬШИХ ДОСЛІДЖЕНЬ

Застосування систем технічного діагностування і моніторингу залізничних сполучень є для України новим, ще неапробованим напрямом. Їх впровадження обмежено через економічні чинники, які пов'язані з улаштуванням та обслуговуванням. Проте визначення технічного стану об'єкту, пошук місць, у яких спостерігаються несправності, є важливою експлуатаційною задачею, зокрема для мереж залізнично-водного сполучення, успішне вирішення якої визначає безпеку руху. Економічна ефективність від створення ІС для технічного обслуговування обумовлена очікуваним зростанням ефективності використання технічних засобів, зменшенням а поступовим виключенням людського фактору та скороченням витрат на заробітну платню та електроенергію.

Подальші дослідження пов'язані з реалізацією запропонованих механізмів створення інформаційної системи та її впровадженням у мережі залізнично-водного сполучення.

Наукове значення роботи. Одержані результати спрямовані на вирішення науково-прикладної проблеми інформаційного забезпечення систем діагностування і моніторингу мереж залізнично-водного сполучення.

Практичне значення роботи полягає у розробленні заходів з управління системами технічної діагностики і моніторингу, що дозволить удосконалити існуючі методи технічного обслуговування пристроїв залізничної автоматики у мережах залізнично-водного сполучення. Результати роботи можуть бути застосовані для тестування пристроїв залізничної автоматики.

ВИСНОВКИ

1. Виконано аналіз існуючих комп'ютерних діагностичних систем, які на даний час застосовуються для технічного обслуговування залізничної автоматики; визначено, що головним недоліком вважається відсутність взаємодії інформаційних систем з системами моніторингу.

2. Визначено та сформульовано основні етапи проектування нової ІС для обслуговування систем технічного діагностування і моніторингу у мережах залізнично-водного сполучення, в основу якої покладено завдання мінімізації і дублювання даних, спрощення процедур їх обробки і поновлення.

3. В основу техніко-економічного обґрунтування щодо необхідності впровадження нових інформаційних систем у мережах залізнично-водного сполучення покладено необхідність створення засобів аналізування та обробки інформації отриманих даних, що сприятиме зменшенню та виключенню втручання персоналу з подальшою економією витрат на заробітну платню та електроенергію.

ЛІТЕРАТУРНІ ДЖЕРЕЛА

[1] Мирошник М. А. Проектирование компьютерных систем с интеллектуальной диагностической инфраструктурой // РАДИОТЕХНИКА : Всезкр. межвед. науч.-техн. сб., 2011. – Вып. 164. – С. 190 –197.

[2] Мирошник М. А. Разработка интеллектуальной диагностической инфраструктуры в распределенных компьютерных системах // Информационно-керуючі системи на залізничному транспорті, 2015. – № 3. – С. 3 – 9.

[3] Дистанционное диагностирование состояния стрелочных переводов по временной характеристике и спектральному составу токовой кривой // Наука та прогрес транспорту. Вісн. Дніпропетр. нац. ун-ту залізн. трансп. ім. акад. В. Лазаряна – Д. 2015. – № 2 (56). – С. 39 – 57.

[4] Елисеєв В. А. Направление развития спутникового мониторинга железнодорожного транспорта // Интерактивная наука, 2016. – № 8. – С. 62 – 68.

[5] Варбанец Р. А. Интернет-мониторинг эксплуатационных режимов судовой энергетической установки // Вісник Одеського національного морського університету, 2011. – № 32. – С. 123 –127.

[6] Вычужанин В.В. Информационное обеспечение мониторинга и диагностирования технического состояния судовых энергоустановок // Вісник одеського національного морського університету, збірник наукових праць, 2012. – № 36. – С. 111 –124.

[7] Хижняк В. В. Інтелектуальна технологія управління експлуатацією авіаційної техніки на основі доменного аналізу // Наука і техніка Повітряних Сил Збройних Сил України, 2016. – № 4 (25). – С. 24 – 29.

- [8] Kang Kyo C. Future-Oriented Domain Analysis // Feasibility Study: Software Engineering Institute, 1990. – P.1.161.
- [9] Neighbors J.V. The Draco Approach to Constructing Software from Reusable Components // IEEE Trans. On Softw. Eng., 1984. – № 10 (3). – P. 564 –576.
- [10] Моисеенкова Д. А. Задача принятия управленческих решений на железнодорожном транспорте // Известия Петербургского университета путей сообщения, 2007. – Вып. 4 (13). – С.5 – 16.
- [11] Цуриков А. Н. Автоматизированная информационная система поддержки принятия управленческих решений и рассылки оповещений в условиях чрезвычайной ситуации на железнодорожной станции // Вагонный парк. – Харьков, изд-во «Подвижной состав», 2014. – № 1 (82). – С. 41 – 44.
- [12] Beauliev A. Learning SQL // Copyright 2009: O'Reilly Media. Printed in the United States of America – 335 p.
- [13] М. Зандстра. РНР-объекты, шаблоны и методики программирования; пер. с англ. С. Н. Тригуб. – М. : Вильямс, 2011. – 528 с.
- [14] Дейт К. Дж. Введение в системы баз данных; пер. с англ. К. А. Птицына. – М. : Вильямс, 2005. – 1328 с.
- [15] Верлан А. І. Підвищення ефективності перевезень у залізнично-водному сполученні за рахунок розвитку приватних припортових станцій : дис. канд. наук : 05.22.20. – Дніпропетровськ, 2015. – 178 с.
- [16] Саати Т. Л. Принятие решений. Метод анализа иерархий. – М.: Радио и связь, 1993. – 450 с.
- [17] Вентцель Е. С. Исследование операций: задачи, принципы, методология. – М.: Наука, 1980. – 207 с.
- [18] Кулинич А.А. Компьютерные системы моделирования когнитивных карт: подходы и методы // Проблемы управления,– 2010. – № 2. – С. 2 –16.
- [19] Коваленко И. И. Экспертные технологии поддержки принятия решений. – Николаев : Иллион, 2013. – 216 с.
- [20] Колтунский Ю. Ю. Разработка системы диагностики электрической централизации с применением вычислительных средств // Електромагнітна сумісність та безпека на залізничному транспорті, 2016. – № 11. – С.76 – 80.
- [21] Язык UML. Руководство пользователя / Г. Буч, Д. Рамбо, И Якобсон [и др.] – 2-е изд. : Пер.с англ. – М. : ДМК Пресс, 2006. – 496 с.

THE INFORMATION SUPPORT OF MAINTENANCE OF RAILWAY-MARINE GOODS

Kazymyrenko Y., Drozd O., Yeholnikov O.

The advanced development of computer technologies in the maintenance and operation of vehicles creates the preconditions for the implementation of new mechanisms to control technical diagnostics and monitoring systems. The application of these mechanisms will make it possible to automatically control the technical condition and formulate information messages about deviations. The specialized information system development for solving the problems of assessing the technical state of railway-waterway connection systems is relevant. It aims at improving the method of technical diagnostics and monitoring of railway automation devices. The authors propose mechanisms for designing the information system, for which the existing diagnostic systems are analyzed, the disadvantages of their operation, their causes and consequences are identified. An object-oriented approach and traditional database design technologies are used in the paper. The peculiarities of IT systems operation in railway-waterway connection mean consolidation of the functions of railway transport with water one, which is considered in the database structure construction, based on the task of minimizing duplication of data and simplifying the procedures for their processing and updating. The software allows to solve research and operational tasks efficiently. The achieved results can be applied to the testing of railway automation devices on the railway-waterway routes.

Keywords. Railway-waterway connection, technical diagnostics and monitoring systems, project information support

ДІАГНОСТИКА РИЗИКІВ АВАРІЙНИХ СИТУАЦІЙ ЗА ДОПОМОГОЮ ГІБРИДНИХ ЕКСПЕРТНИХ СИСТЕМ

Коновалов С., Єгошина Г.

У даній статті представлений комплексний метод діагностики ризиків аварійних ситуацій, який ґрунтується на використанні гібридних експертних систем. Всі ризики, що загрожують складній системі можна структурувати за допомогою бази знань, де зазначатимуться значення змінних при аварійній ситуації, а також дії для усунення цієї ситуації з урахуванням часу і вартості при їх застосуванні. Різні загрози, які можуть становити небезпеку для складних технічних систем в тій чи іншій ймовірності, представлені графічно за допомогою діаграми Ішікави. Також при діагностиці були враховані різні види надійності складних технічних систем, які розраховувалися при тій чи іншій ймовірності роботи систем. Крім цього враховується розрахунок і всіляких ризиків, які впливають як на надійність, так і на працездатність. При цьому сама працездатність розраховується з урахуванням даних змінних, в тому числі і ризиків всіх технічних вузлів складових частин складної технічної системи, а остаточно обчислюється за допомогою нейронної мережі, яка враховує також і типи самих даних (лінгвістичні та нелінгвістичні). Сама нейронна мережа здатна навчатися, застосовуючи для цього злиття таких методів як метод зворотного поширення помилки і рекурентного методу. Для постановки остаточного діагнозу працездатності складної технічної системи застосовується параметричне діагностування двигуна, яке засноване на порівнянні математичних моделей даної самої системи, що діагностується з математичною моделлю еталонної складної технічної системи. В результаті були розроблені алгоритми для діагностики ризиків за допомогою гібридних експертних систем. Один з них показує більш конкретно аналіз працездатності за даними аварійності системи при використанні нечіткої логіки гібридної експертної системи. Інший описує більш загальну діагностику ризиків у вигляді проходження інформації про стан системи до її повної обробки гібридною експертною системою.

Ключові слова. Гібридна експертна система, база знань, нейронна мережа, нечітка логіка, діаграма Ішікави, надійність, працездатність, діагностика, ризики.

ВСТУП

Останнім часом проблема управління ризиками дуже актуальна. Ризик - це складна, іноді нерозв'язна задача, особливо коли це стосується складних технічних систем.

Управління ризиками являє собою безліч системно організованих процедур, що призводять величину ризику в певні межі. Під ризиком розуміється діяльність, яка пов'язана з подоланням невідомості в ситуації, коли вибір неминучий, тобто існує можливість оцінити

ймовірність досягнення передбачуваного результату, невдачі і відхилення від необхідної мети як кількісно, так і якісно [1].

Проблема протиаварійного керування шляхом зниження ризиків ще досить гостра. Це пов'язано з нестачею графічного представлення ймовірних факторів небезпеки. Популярні засоби представлення різних небезпек не забезпечують в повному обсязі прийняття швидкого рішення для безпечної експлуатації судна. Але при цьому розроблений новий підхід до безпеки управління, який представлений кластерно-ймовірною інформаційною моделлю. Дана модель показує загрози і ризики при роботі складних технічних систем судна.

Безпечний режим роботи в складних технічних системах (СТС) досягається шляхом своєчасного запобігання виникнення ризиків аварійних ситуацій. Ефективно цього можна досягти при використанні гібридних експертних систем (ГЕС) [2, 3].

МЕТА ДОСЛІДЖЕННЯ

Запобігання виникненню аварійних ситуацій в складних технічних системах, а також швидка протидія таким, якщо запобігти по тій або іншій причині не виходить. Все це повинно бути досягнуто за допомогою розробки комплексного методу для діагностики ризиків аварійних ситуацій за допомогою гібридних експертних систем.

ОБ'ЄКТ ДОСЛІДЖЕННЯ

Складна технічна система, яка діагностується в умовах змінних і екстремальних техногенних та природних впливів.

ПРЕДМЕТ ДОСЛІДЖЕННЯ

Технічні характеристики, оцінка різних ризиків, а також причини відмов функціональних елементів СТС.

МЕТОДИ ДОСЛІДЖЕННЯ

Були використані експериментально-теоретичні методи, в тому числі, аналіз, синтез, моделювання. Застосовано системний підхід до подання СТС як ієрархічної системи взаємопов'язаних і взаємодіючих підсистем.

ОСНОВНИЙ МАТЕРІАЛ ДОСЛІДЖЕННЯ

Той чи інший ризик аварійної ситуації можна структурувати, тобто розкласти на складові дані СТС, при певному значенні яких він виникає. Всі ці дані знаходяться в базі знань (БЗ) ГЕС (табл. 1).

Таблиця 1. База знань ГЕС

Аварійна ситуація	Значення змінних СТС				Дії для усунення аварійної ситуації	Витрачений час	Вартість
	z_1	z_2	...	z_k			
as_1	B	C	...	H	U_1	t_1	C_1
as_2	H	BC	...	C	U_2	t_2	C_2
...
as_d	B	HC	...	C	U_d	t_d	C_d

У табл. 1. B – значення «високе», BC – значення «вище середнього», C – значення «середнє», HC – значення «нижче середнього», H – значення «низьке», k – кількість змінних, d – кількість аварійних ситуацій.

Дана БЗ поповнюється за допомогою нейронної мережі (НМ), здатної до навчання, застосовуючи для цього злиття методів зворотного поширення помилки і рекурентного методу [4]. Також для графічного представлення класифікації ризиків для ГЕС застосовується діаграма Ішікави [5]. Дана діаграма є оригінальним підходом графічного моделювання прикладів загроз і ризиків технічної безпеки.

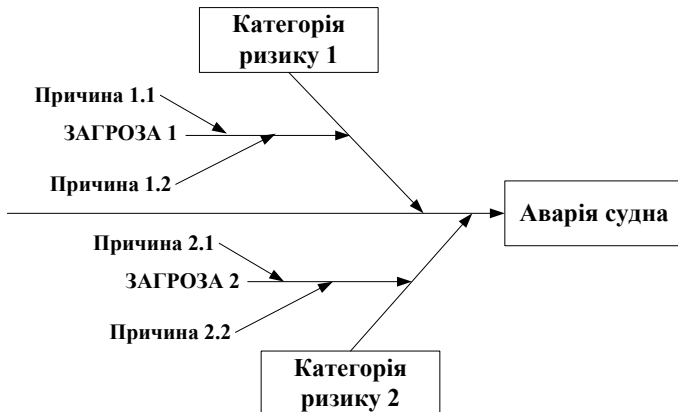


Рис. 1 – Приклад фрагмента діаграми Ішікави

Представлена діаграма схожа на скелет риби (рис. 1). Аварійна ситуація в цілому позначена центральною стрілкою по горизонталі, а різні ризики представлені стрілками, які спрямовані до центральної стрілки по горизонталі, при цьому нахил йде в ліву сторону. Рівень аналізу стає більш глибоким шляхом додавання стрілок, які впливають на ризики факторів загроз. Подальша деталізація з поділом факторів загроз робиться, поки не будуть виявлені справжні причини загроз роботи СТС.

Етапи побудови діаграми Ішікави:

I. визначаються різноманітні причини і загрози, що впливають на аварійність;

II. систематизуються причинні фактори, а також смислові розділи;

III. ранжуються і оцінюються загрози і причини всередині цих розділів;

IV. отримана структура аналізується;

V. визначаються і віднімаються ті причинні, які неможливо змінити;

VI. ігноруються несуттєві причини і загрози.

Переваги моделі діаграми: наочність показу зв'язків між аварійною ситуацією та її причиною; можливість провести аналіз ланцюга факторів загроз, які впливають на проблему аварійності; наочність взаємозв'язків між різноманітними причинами і порівняння їх відносної важливості; отримання інформації для прийняття рішень при аварійній ситуації.

Недолік діаграми Ішікави полягає в складності застосування при високій ймовірності плутанини в складній технічній системі, а також громіздкість графічного представлення зображення всіх «гілок», але це компенсується присутністю програмного забезпечення, за допомогою якого можливий приватний розгляд зображення будь-якої гілки.

Ризик найчастіше оцінюють при розрахунку таких понять, як надійність і працездатність СТС. При технічному обслуговуванні, а також ремонті оцінюють експлуатаційну (реальну) надійність СТС. Експлуатаційна надійність СТС - надійність, обумовлена в реальних умовах експлуатації, враховуючи комплексний вплив зовнішніх і внутрішніх аварійних факторів, пов'язаних як з кліматичними, так і з географічними особливостями експлуатації, реальними режимами роботи функціональних систем СТС, а також різними умовами, при яких вони обслуговуються. Виробнича надійність СТС - це надійність, яку можна забезпечити в даних певних умовах їх застосування при ідеальній, тобто абсолютно надійній, технічній експлуатації СТС.

Проектна надійність СТС переважно визначається надійністю структури СТС (виробничою надійністю) і надійністю технічної експлуатації (експлуатаційною надійністю).

Проектна надійність СТС формується такою за допомогою подібної структурної моделі:

$$N = NE \quad (1)$$

де N – подія, що показує досягнення результатів для поставлених завдань при експлуатації СТС;

E – подія, що показує безвідмовність експлуатації СТС під час виконання цих завдань.

Ймовірність події N :

$$P(N) = P(N/E)P(E) \quad (2)$$

де $P(N)$ – повна ймовірність виконання СТС поставленого завдання при експлуатації;

$P(N/E)$ – умовна ймовірність виконання СТС поставленого завдання при експлуатації, обчислена при безвідмовній експлуатації СТС, умовний ризик;

$P(E)$ – ймовірність безвідмовної експлуатації СТС при виконанні поставленого завдання, мінімальний ризик.

Подія E є складною, вона представлена у вигляді добутку декількох подій:

$$E = E_1 E_2 E_3 E_4 \quad (3)$$

де E_1 – відсутність в СТС факторів небезпеки під час проектування;

E_2 – відсутність в СТС факторів небезпеки, пов'язаних з виробництвом;

E_3 – відсутність в СТС факторів небезпеки при монтажі;

E_4 – відсутність факторів небезпеки в експлуатаційно-технічній документації для обслуговування СТС.

Покладаючись на вищенаведену модель можна записати:

$$P(E) = P(E_1)P(E_2/E_1)P(E_3/E_1E_2)P(E_4/E_1E_2E_3) \quad (4)$$

де $P(E_1)$ – ймовірність того, що проект СТС безпечний;

$P(E_2/E_1)$ – ймовірність того, що СТС виготовлена безпечно, при відсутності факторів небезпеки в проектній документації;

$P(E_3/E_1E_2)$ – ймовірність того, що СТС безпечно зібрали і змонтували, за умови безпечного проектування і виготовлення СТС;

$P(E_4/E_1E_2E_3)$ – ймовірність того, що експлуатаційно-технічна документація безпечна, з урахуванням відсутності факторів ризику проекту, виготовлення та монтажу СТС. Ввівши дані значення:

$$M = P(N); M_{\Pi} = P(N/E); M_E = P(E) \quad (5)$$

Уявімо цю модель в такому вигляді:

$$M = M_{\Pi} M_E \quad (6)$$

де M , M_{Π} і M_E – загальні значення проектної, виробничої та експлуатаційної надійності СТС відповідно.

При визначенні оцінок значень в певний момент часу виробничої M_{Π} , експлуатаційної M_E і проектної M надійності СТС знаходять також оцінки в певний момент часу їх середніх квадратичних відхилень $\sigma_{M_{\Pi}}$, σ_{M_E} , σ_M .

Також обчислюються різні ризики, які можуть вплинути на показання даних працездатності СТС, що може відбитися на діагностиці ГЕС її стану. Зокрема працездатність СТС залежить від наступних ризиків: ризик відмови, ризик помилки з сигналом [6, 7, 8].

R – ризик відмови СТС обчислюється за такою формулою:

$$R = f(P, U, Q) \quad (7)$$

де P – ймовірність аварії СТС;

U – збиток від аварії СТС;

Q – вага ризику.

За формулою теорії надійності, ймовірність аварії СТС P :

$$P(t) = 1 - \exp\left[-\int_0^t \lambda(\tau) d\tau\right] \quad (8)$$

де $\lambda(t)$ – залежна від часу інтенсивність відмов вузлів.

Також ймовірність ланцюжка подій, що приводить до ризику аварійної ситуації, визначається формулою [9]:

$$P(as) = P(s_0) \prod_{i=1}^N P_{s_{i-1}}(s_i) \quad (9)$$

де $P(as)$ – ймовірність аварійної ситуації;

$P(s_0)$ – ймовірність вихідної події ланцюга, що призводить до аварійної ситуації;

$P_{s_{i-1}}(S_i)$ – умовна ймовірність проміжної події S_i за умови виконання попередньої події ланцюжка.

Ризик відмови СТС R визначається як сума ризиків для всіх технічних вузлів r_i СТС:

$$R = \sum_{i=1}^m r_i \cdot u_i \quad (10)$$

де u_i – збиток через аварію i -го вузла СТС;

m – загальна кількість технічних вузлів СТС;

r_i – ризик відмови для технічного вузла i :

$$r_i = v_i \cdot p_i(t) \quad (11)$$

де $p_i(t)$ – ймовірність аварії технічного вузла i в момент часу t ;

v_i – вага ризику для технічного вузла i .

Критичність k_i (ймовірність переходу системи в відмову в тому випадку, коли даний елемент вийшов з ладу) показана наступною формулою:

$$k_i = P(S_n / S) \quad (12)$$

де $P(S_n / S)$ - ймовірність переходу СТС в несправний стан S_n з справного S , за умови, що вузол i вийшов з ладу.

При цьому ризик відмови технічного вузла з урахуванням критичності r_i^k знаходиться:

$$r_i^k = k_i \cdot r_i = k_i \cdot v_i \cdot p_i(t) \quad (13)$$

Ризик аварії СТС, з огляду на критичність R^k :

$$R^k = \sum_{i=1}^m r_i^k \cdot u_i \quad (14)$$

Ризик помилки з сигналом R^S , у випадках його спотворення, чи некоректного прийняття при передачі даних з одного вузла до іншого:

$$R^S = \sum_{i=1}^z l_i \cdot p(S_j / S_i) \quad (15)$$

де S_j - помилковий сигнал;

S_i - правильний сигнал;
 l_i - втрати при помилку;
 $p(S_j / S_i)$ - ймовірність помилки;
 Z - кількість сигналів.

Виходячи з вищесказаного, залежність працездатності СТС W від ризиків знаходиться:

$$W = f(R^k, R^S) \quad (16)$$

При оцінці ризику вирішується завдання класифікації. Працездатність СТС має на увазі сукупність даних різноманітних змінних, які належать до всіх технічних вузлів в складі системи, яка діагностується (рис. 2).

Працездатність кожного технічного вузла СТС, як і всілякі його ризику знаходиться, з огляду на параметри цих вузлів і їх типів даних (лінгвістичні і нелінгвістичні). У зв'язку з цим працездатність кожного технічного вузла знаходиться, з огляду на автоматизовану експертну БЗ, в якій проводяться розрахунки для змінних, а також їх аналіз. Після цього, покладаючись на ці дані, нейронна мережа виробляє остаточний розрахунок загальної працездатності СТС [10]:

$$W_m = N(F_{m1}(x_{m11}, \dots, x_{m1n}), \dots, F_{mh}(x_{mh1}, \dots, x_{mhk}), y_{m11}, y_{m12}) \quad (17)$$

де $N(\)$ – нейромережева функція;
 x, y – змінні параметри технічного вузла СТС;
 h – кількість функцій для заданого вузла;
 n, k – кількість змінних для кожної функції.

Крім цього для постановки діагнозу працездатності СТС застосовується параметричне діагностування двигуна [11], яке засноване на порівнянні математичних моделях даної СТС з математичною моделлю еталонної СТС (рис. 3). В результаті діагностика ризиків на основі даних СТС представлена у вигляді двох алгоритмів: аналізу працездатності за аварійністю (рис. 4) та аналізу технічного стану СТС за працездатністю (рис. 5). На рис. 4. зображений алгоритм, що описує діагностику працездатності СТС на предмет її аварійності. На вході розміщується сама працездатність СТС. Спочатку працездатність надходить на блок нечіткої логіки (НЛ), де шляхом її порівняння зі значеннями працездатності з БЗ з'ясовується, виходить її значення за рамки норми. Якщо значення працездатності СТС виходить за рамки норми, блок НЛ видає значення «1», тобто ситуація аварійна. Тоді проходить перевірка наявності подібної ситуації в БЗ.

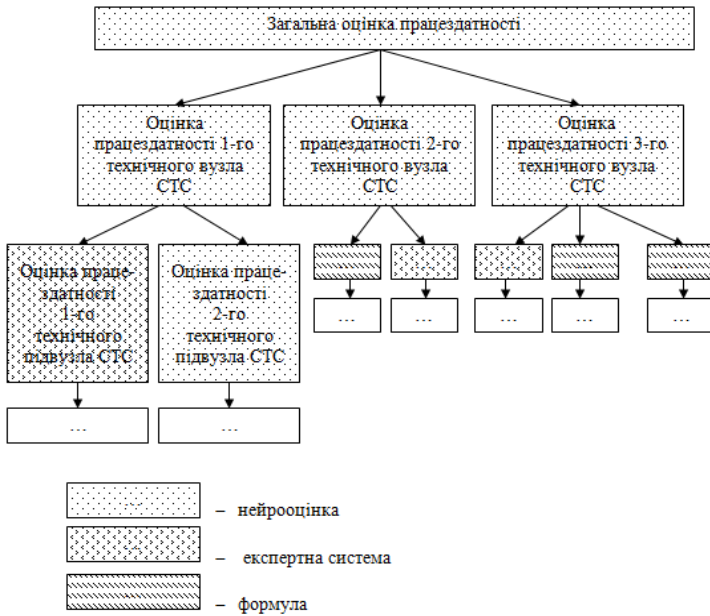


Рис. 2 – Структура моделі оцінки працездатності СТС



Рис. 3 – Схематичне представлення діагностики СТС за допомогою ГЕС

Після знаходження причин аварії і дій для її ліквідації всі дані передаються особі, що приймає рішення (ОПР). Якщо значення працездатності СТС не виходить за рамки норми, блок НЛ видає значення «0», тобто аварійна ситуація відсутня. Після цього відбувається прогнозування технічного стану СТС. Дані прогнозування передаються ОПР. На виході отримуємо перевірену СТС на приклад аварійності. На рис. 5. показана друга стадія діагностики СТС, як аналіз технічного стану СТС за отриманою працездатністю. На вході розміщується дані СТС, які проходять моніторинг, після якого знаходиться працездатність СТС. По

цій працездатності проводиться діагностика технічного стану СТС. В процесі цієї діагностики перевіряється аварійна ситуація в СТС.

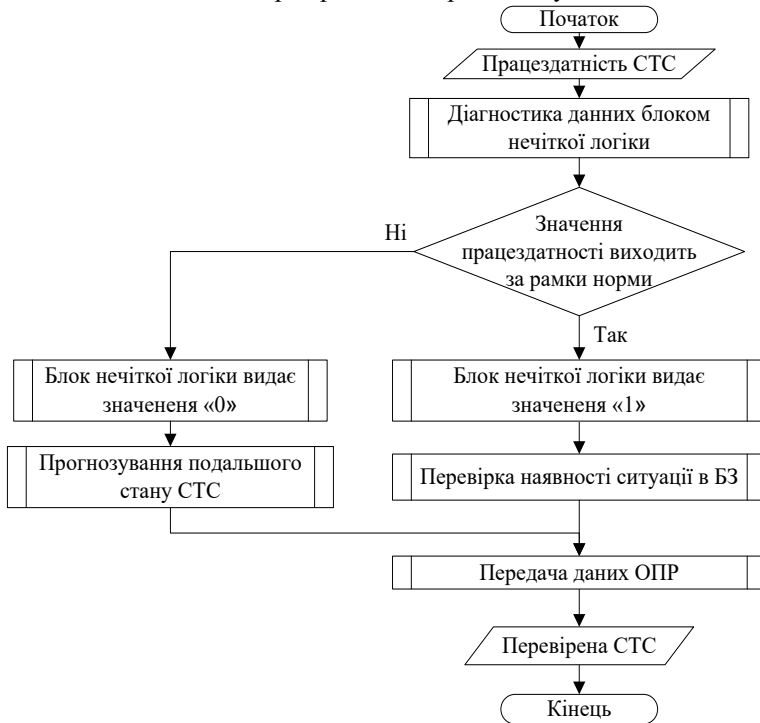


Рис. 4 – Алгоритм аналізу працездатності за аварійністю

Якщо аварійної ситуації немає, то проводиться прогнозування технічного стану СТС. Якщо аварійна ситуація присутня, то перевіряється, є дана ситуація в БЗ. Якщо дана ситуація в БЗ є, то в БЗ відразу знаходяться причини і дії для запобігання цій ситуації. Якщо ж такої ситуації немає в БЗ, то здійснюється пошук причини аварійної ситуації по технічним вузлам.

Спочатку йде пошук по верхній структурі СТС, тобто з технічних вузлів, і перевіряється, працездатність якого вузла або декількох вузлів негативно вплинули на загальну працездатність, що призвело до аварійної ситуації. Після знаходження потрібного вузла або вузлів, йде пошук на нижньому рівні, тобто розглядаються різні критерії працездатності обраного вузла або вузлів, значення яких негативно вплинули на працездатність СТС, відхилившись від норми в більшу або меншу сторону.

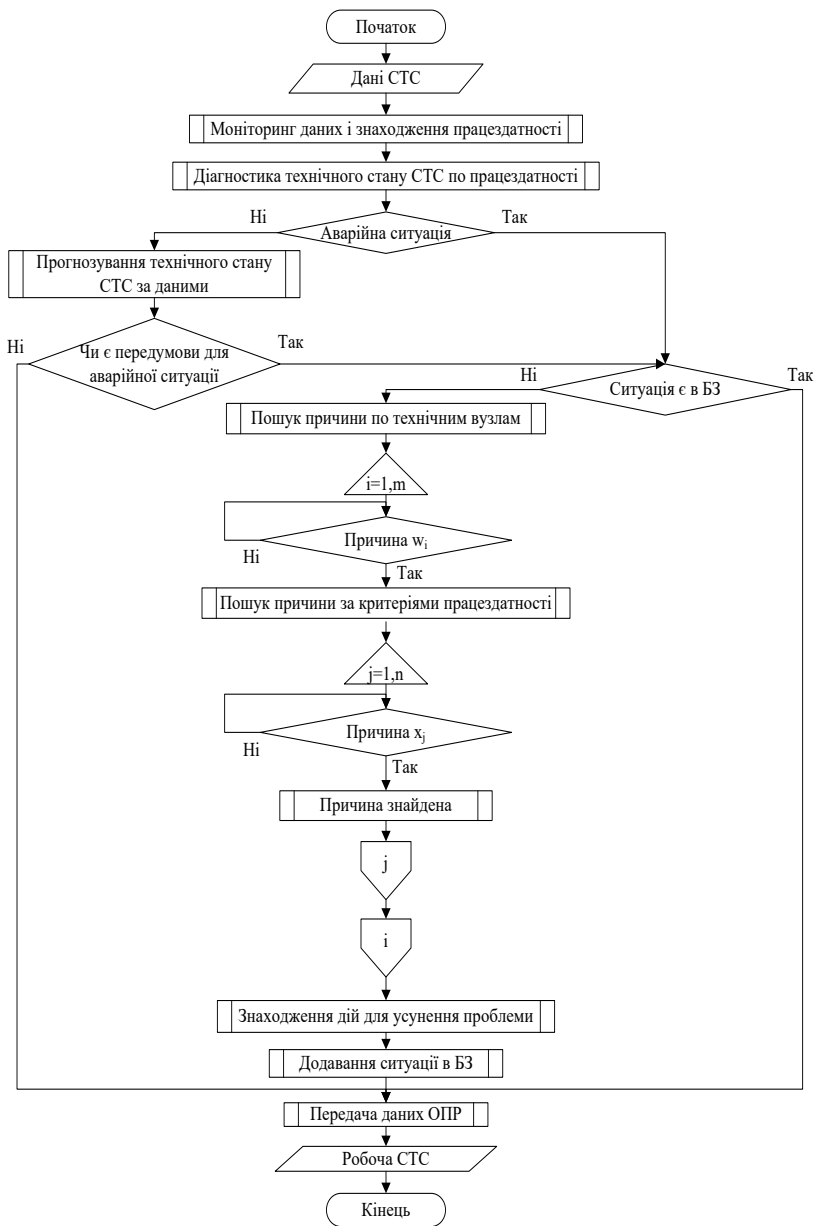


Рис. 5 – Алгоритм аналізу технічного стану СТС за працездатністю

В результаті знаходження причин аварійної ситуації йде знаходження дій для усунення даної ситуації (зміни значень параметрів, які виходять за рамки норми і т.п.). Після цього дана ситуація, її причини та рішення додаються в БЗ для подальшого уніфікації діагностики аварійних ситуацій. Всі ці дані потім передаються оператору, тобто особі, що приймає рішення. У підсумку, на виході отримаємо проаналізовану робочу СТС.

ВИСНОВКИ

Розроблений метод діагностики ризиків аварійних ситуацій на складних технічних системах за допомогою ГЕС забезпечує можливість отримувати досить точні діагностичні дані про працездатність СТС через великий діапазон обчислюваних даних, які можна охопити різними розрахунками, застосовуючи при цьому різні методики діагнозу гібридної експертної системи: бази знань, нейронні мережі, нечітка логіка. Все це сприяє кращому запобіганню виникнення аварійних ситуацій у складних технічних системах.

ЛІТЕРАТУРНІ ДЖЕРЕЛА

- [1] Махитько В.П. Оценка рисков и факторов опасности в системе безопасности полетов воздушных судов // Известия Самарского научного центра Российской академии наук, 2017. – том 19. – № 4(2). – С. 192 – 197.
- [2] Nourian R. A fuzzy expert system for mitigation of risks and effective control of gas pressure reduction stations with a real application // Journal of loss prevention in the process industries, 2019. – № 59. – P. 77 – 90.
- [3] Lopez C. A scenario-based modeling method for controlling ECM performance // Expert systems with applications, 2018. – № 97. – P. 253 –265.
- [4] Коновалов С.Н. Особенности разработки искусственной нейронной сети гибридной экспертной системы // Штучний інтелект, Київ: ІППШ «Наука і освіта», 2018. – № 2(80). – С. 139 – 143.
- [5] Вильский Г.Б. Информационные риски судоходства в диаграмме Исикавы // Науковий вісник Херсонської державної морської академії, 2016. – № 1(14). – С. 19 – 25.
- [6] Vychuzhanin V., Rudnichenko N. Devising a method for the estimation and prediction of technical condition of ship complex systems. Eastern-European Journal of Enterprise Technologies, Information and controlling systems, 2016. – 6/9 (84). – P. 4 – 11.
- [7] Вычужанин В.В., Бойко В.Д. Оценка живучести судовых технических систем // Журнал университета водных коммуникаций. — СПб.: ГУМРФ имени адмирала С. О. Макарова, 2013. – (Вып. 2) . – С.38 – 45.
- [8] Коновалов С.Н., Вычужанин В.В. Информатизация противоаварийного управления сложными техническими системами //

Інформатика та математичні методи в моделюванні. – Одеса: ОНПУ, 2017. – том 7. – № 4. – С. 265 – 275.

[9] Серебровский А.Н. Подход к созданию базы знаний экспертной системы оценки, прогноза и анализа ситуаций на объектах повышенной опасности / А.Н. Серебровский. // Математичні машини і системи, 2009. – № 4. – С. 58 – 66.

[10] Konovalov S.N., Vychuzhanin V.V. Method for antifault control of complex technical systems // Розвиток транспорту, збірник наукових праць, Одеса: ОНМУ, 2017. – № 1(1). – С. 45 – 59.

[11] Шишулин А.В. Возможность построения нейро-нечетких экспертных систем диагностики авиационных газотурбинных двигателей // Научный Вестник МГТУ ГА, 2005. – № 85. – С. 79 – 81.

DIAGNOSIS OF RISKS OF EMERGENCY SITUATIONS BY USING HYBRID EXPERT SYSTEMS

Konovalov S., Yehoshina G.

This article presents a comprehensive method for diagnosing emergency risks, which is based on the use of hybrid expert systems. All risks that threaten a complex system can be structured using the knowledge base, which will indicate the values of variables in an emergency, as well as actions to eliminate this situation taking into account time and cost when applied. Various threats that may be dangerous to complex technical systems in one way or another are represented graphically using the Ishikawa diagram. Also, during diagnostics, various types of reliability of complex technical systems were taken into account, which were calculated for one or another probability of operation of the systems. In addition, the calculation of all kinds of risks that affect both reliability and performance are taken into account. At the same time, the working capacity itself is calculated taking into account these variables, including the risks of all the technical nodes of the components of a complex technical system, and is finally calculated using a neural network, which also takes into account the types of data themselves (linguistic and non-linguistic). The neural network itself is able to learn using a fusion of methods such as the back propagation method of error and the recurrence method. To make a final diagnosis of the operability of a complex technical system, parametric engine diagnostics is used, which is based on a comparison of the mathematical models of this most diagnosed system with the mathematical model of the reference complex technical system. As a result, algorithms were developed for diagnosing risks using hybrid expert systems. One of them shows more specifically the analysis of health according to the system crash data when using the fuzzy logic of a hybrid expert system. Another describes a more general risk diagnosis in the form of passing information about the state of the system before it is completely processed by a hybrid expert system.

Keywords. Hybrid expert system, knowledge base, neural network, fuzzy logic, Ishikawa diagram, reliability, performance, diagnostics, risks.

АЛГОРИТМЫ ПОЛНОГО ПЕРЕБОРА В СХЕМАХ ВЕТВЛЕНИЯ РЕШЕНИЯ КОМБИНАТОРНЫХ ЗАДАЧ С ПСЕВДОБУЛЕВЫМИ ФУНКЦИЯМИ

Трофимчук А., Васянин В., Ушакова Л.

В статье приводятся полезные сведения для разработчиков алгоритмов и программ об использовании кодов Грея для решения комбинаторных задач с псевдодобулевыми функциями. В качестве примера эффективности применения этих кодов рассматривается решение двух комбинаторных задач с булевыми переменными с полным перебором вариантов решения. Представлены результаты экспериментального исследования, которые показывают, что коды Грея можно практически применять в схемах ветвления, например, в методе ветвей и границ, когда количество переменных в узлах ветвления решающего алгоритма не превышает 35.

Ключевые слова. Коды Грея, задачи комбинаторной оптимизации, время решения задачи

ВВЕДЕНИЕ

В статье рассматривается применение двоично-отраженных кодов Грея для решения комбинаторных задач с псевдодобулевыми функциями (полиномами от булевых переменных). Приводится рекурсивный алгоритм Эрлиха для генерации последовательности строк n -разрядных кодов Грея, в которой каждая следующая строка отличается от предыдущей только одним разрядом (битом). На примерах решения 0-1 задачи о ранце (0-1 Knapsack Problems) [1] и задачи выбора пропускных способностей дуг коммуникационной сети с ограничением на время задержки потоков [2] показано, как эти коды можно использовать для эффективного вычисления значений целевой функции и ограничений.

Цель статьи состоит в том, чтобы показать разработчикам алгоритмов и программ как можно применять коды Грея в различных схемах разветвления решающего алгоритма, например, в методе ветвей и границ, когда количество двоичных (булевых) переменных в узлах дерева ветвления не больше 35.

Методика исследований основана на проведении вычислительного эксперимента решения вышеуказанных задач предложенным алгоритмом перебора вариантов решений с частичным и полным пересчетом значений целевой функции и ограничений.

В 1953 г. физик Фрэнк Грей (Frank Gray) получил патент на изобретение двоично-отраженных n -разрядных кодов, которые и были названы его именем [3]. Изначально эти коды применялись в кодово-импульсной модуляции для управления различными электромеханическими переключателями и методе аналоговой передачи цифровых сигналов. В настоящее время коды Грея используются для выявления и исправления ошибок в системах связи, управления различными цифровыми датчиками, кодирования номеров дорожек в жестких накопителях компьютеров и пр. Кроме того известно о применении кодов Грея для решения комбинаторных задач «Ханойская башня» и «Китайские кольца» [4]. Подробнее о кодах Грея можно узнать в книге Д. Кнута [5].

ЗАДАЧА О РАНЦЕ И ЗАДАЧА ВЫБОРА ПРОПУСКНЫХ СПОСОБНОСТЕЙ ДУГ

Математическая формулировка первой задачи заключается в следующем. Задан набор n предметов, для каждого из которых известна стоимость $c_i \in Z^+$ и вес $a_i \in Z^+$, $i = \overline{1, n}$. Требуется так загрузить ранец предметами, чтобы его суммарная стоимость была максимальной

$$\max \sum_{i=1}^n c_i x_i \quad (1)$$

и выполнялось ограничение на суммарный размер ранца $W \in Z^+$

$$\sum_{i=1}^n a_i x_i \leq W, \quad (2)$$

$$x_i \in \{0, 1\}, i = \overline{1, n}. \quad (3)$$

Предполагается, что $a_i \leq W$, $i = \overline{1, n}$ и $\sum_{i=1}^n a_i > W$.

Вторая задача заключается в выборе пропускных способностей дуг из заданного набора дискретных целочисленных значений при ограничении на максимальное время задержки потоков, которая актуальна при распределении потоков в многопродуктовых коммуникационных сетях. В этой задаче задержки потоков t_{kl} на дугах определяются как $t_{kl} = f_{kl} / (w_{kl} - f_{kl})$, $kl \in E$, а ограничение на время задержки потоков в сети имеет вид $t_{av} = 1 / U_{\Sigma} \sum_{kl \in E} f_{kl} / (w_{kl} - f_{kl}) \leq T_{\max}$.
Здесь $f_{kl} \in Z^+$ - фиксированное значение потока по дуге $kl \in E$, E - множество дуг сети, $w_{kl} \in Z^+$ - пропускная способность дуги $kl \in E$,

T_{\max} — максимальное время задержки потоков в сети, $U_{\Sigma} = \sum_{ij \in S} u_{ij}$ -

суммарный поток в сети, $u_{ij} \in Z^+$ - величина потока из узла i в узел j , S - множество пар индексов корреспондирующих узлов в сети. При приближении величины потока на дугах к их пропускным способностям задержки увеличиваются и, следовательно, могут возникать перегрузки в сети. Суть задачи заключается в том, что при фиксированных потоках требуется так выбрать пропускные способности дуг из заданного набора целых чисел, чтобы выполнялось ограничение на время задержки потоков, и достигался минимум некоторой целевой функции. Такая задача возникает не только в сетях передачи данных, но и в транспортных сетях при распределении потоков по критерию минимума стоимости сети и заданном ограничении на время задержки потоков [6-8]. Управляя параметром T_{\max} для максимальной задержки, администратор сети передачи данных или диспетчер транспортной сети может обеспечить требуемый ему резерв пропускной способности каналов связи или грузоподъемности транспортных средств при прогнозируемых колебаниях величины потоков на заданных промежутках времени. Уменьшение параметра T_{\max} (увеличение резерва) приводит к удорожанию сети, но уменьшает вероятность перераспределения потоков и технического переоснащения каналов связи или парка транспортных средств при увеличении потоков и угрозе возникновения перегрузок в сети. Увеличение параметра T_{\max} дает возможность уменьшить пропускную способность каналов связи или грузоподъемность транспортных средств и стоимость сети, но увеличивает риск перераспределения потоков и вероятность модернизации сети.

Математическая модель задачи формулируется следующим образом. Пусть задана связная ориентированная сеть $G(N, E)$ с множеством узлов N , $n = |N|$ и множеством дуг E , $e = |E|$, где n и e соответственно количество узлов и дуг сети. Будем считать, что сеть такова, что для каждой прямой дуги kl , ($k < l$) существует обратная lk , ($l > k$). Для сети передачи данных дуга представляет коммутированную линию связи, состоящую из одного или пучка элементарных каналов. Для транспортной сети дуга отождествляется с маршрутом транспортного средства, концевые узлы которого совпадают с начальным и конечным узлами дуги. Сеть может содержать петли и параллельные дуги, поскольку допускаются циклические и повторяющиеся маршруты, а также маршруты с одинаковыми концевыми узлами. На сети задана целочисленная матрица потоков

$U = \left\| u_{ij} \right\|_{n \times n}$, где u_{ij} - величина потока (количество единиц потока) из узла i в узел j в некоторых транспортных блоках заданного размера. Пусть w_{kl} , $kl \in E$ - искомые пропускные способности дуг сети в транспортных блоках, $w_{kl} \in \{w_1, w_2, \dots, w_\alpha\}$, w_i , $i = \overline{1, \alpha}$ - упорядоченные по возрастанию целые положительные числа; d_{kl} , $kl \in E$ - длины дуг; $C_{kl}(w_{kl}, d_{kl}) \in R^+$, $kl \in E$ - дискретные стоимости дуг, такие, что $C_{kl}(w_i, d_{kl}) \leq C_{kl}(w_{i+1}, d_{kl})$, $i = \overline{1, \alpha-1}$; $f_{kl} = \sum_{ij \in S} u_{ij}^{kl}$, $kl \in E$ - фиксированные суммарные потоки в транспортных блоках, протекающие по дугам сети, где u_{ij}^{kl} - поток транспортных блоков из i в j , проходящий по дуге kl .

Требуется найти минимальное значение функции стоимости сети

$$\min_{w_{kl}} \sum_{kl \in E} C_{kl}(w_{kl}, d_{kl}), \quad w_{kl} \in \{w_1, w_2, \dots, w_\alpha\} \quad (4)$$

при ограничении:

$$\frac{1}{U_\Sigma} \sum_{kl \in E} \frac{f_{kl}}{w_{kl} - f_{kl}} \leq T_{\max}, \quad w_{kl} > f_{kl}, \quad \forall kl \in E. \quad (5)$$

Задачу (4), (5) представим в виде задачи о ранце с булевыми переменными и мультивыбором (0-1 Multiple-choice Knapsack Problem, 0-1 МСКР). Пусть $c_{ij} \in R^+$ - дискретные стоимости дуг i с пропускной способностью $w_{ij} \in \{w_1, w_2, \dots, w_\alpha\} \in Z^+$ и длиной d_i , $j = \overline{1, \alpha}$, $i = \overline{1, e}$; $t_{ij} = f_i / (w_{ij} - f_i)$, $w_{ij} > f_i$, $j = \overline{1, \alpha}$, $i = \overline{1, e}$ - задержки потоков на дугах; f_i - поток по дуге i , $i = \overline{1, e}$. Положим $x_{ij} = 1$, если для дуги i выбрана пропускная способность w_{ij} , $j = \overline{1, \alpha}$, $i = \overline{1, e}$ и $x_{ij} = 0$ в противном случае.

Требуется найти

$$\min \sum_{i=1}^e \sum_{j=1}^{\alpha} c_{ij} x_{ij} \quad (6)$$

при ограничениях

$$\frac{1}{U_\Sigma} \sum_{i=1}^e \sum_{j=1}^{\alpha} t_{ij} x_{ij} \leq T_{\max}, \quad (7)$$

$$\sum_{j=1}^{\alpha} x_{ij} = 1, \quad i = \overline{1, e}, \quad (8)$$

$$x_{ij} \in \{0, 1\}. \quad (9)$$

Здесь искомые пропускные способности w_{ij} соответствуют x_{ij}^* - оптимальному решению задачи (6)-(9).

Нетрудно видеть, что любую индивидуальную задачу, сформулированную в виде (4), (5), можно за время $O(e\alpha)$ преобразовать в соответствующий экземпляр задачи (6)-(9). Для этого необходимо построить две матрицы размером $e \times \alpha$, строки которых соответствуют дугам, столбцы — набору дискретных пропускных способностей, а в качестве элементов матриц принимаются стоимости дуг c_{ij} и задержки на дугах t_{ij} . Справедливо и обратное преобразование. В работе [2] доказано, что оптимизационная задача, сформулированная в виде (4), (5) является NP-трудной, и для ее решения предложено два приближенных алгоритма, основанных на аппроксимации дискретных функций стоимости линейными, и на методе последовательного анализа вариантов. Было показано, что оба алгоритма на заключительной стадии работы сужают область допустимых решений до двух значений $w_{ij} \in \{w_j, w_{j+1}\}$, $j = \overline{k, \alpha - 1}$, $i = \overline{1, e}$, и при полном переборе 2^e вариантов позволяют получить точное решение. Запишем сокращенную задачу в виде

$$\min \sum_{i=1}^e \sum_{j=1}^2 c_{ij} x_{ij} \quad (10)$$

при ограничениях

$$\frac{1}{U_\Sigma} \sum_{i=1}^e \sum_{j=1}^2 t_{ij} x_{ij} \leq T_{\max}, \quad (11)$$

$$\sum_{j=1}^2 x_{ij} = 1, \quad i = \overline{1, e}, \quad (12)$$

$$x_{ij} \in \{0, 1\}. \quad (13)$$

где c_{ij} , $j = \overline{1, 2}$, $i = \overline{1, e}$ - дискретные стоимости дуг; $t_{ij} = f_i / (w_{ij} - f_i)$, $j = \overline{1, 2}$, $i = \overline{1, e}$ - задержки потоков на дугах; f_i - поток по дуге i , $i = \overline{1, e}$.

Задачу (10) - (13) можно преобразовать в 0-1 задачу о ранце:

$$\min \sum_{i=1}^e (c_{i1} + \Delta c_i x_i) \quad (14)$$

при ограничениях

$$\frac{1}{U_{\Sigma}} \sum_{i=1}^e (t_{i1} - \Delta t_i x_i) \leq T_{\max}, \quad (15)$$

$$x_i \in \{0, 1\}, \quad i = \overline{1, e}, \quad (16)$$

где $\Delta c_i = c_{i2} - c_{i1}$, $t_{i1} = f_i / (w_{i1} - f_i)$, $\Delta t_i = f_i / (w_{i1} - f_i) - f_i / (w_{i2} - f_i)$, $i = \overline{1, e}$, $\frac{1}{U_{\Sigma}} \sum_{i=1}^e t_{i1} > T_{\max}$, $\frac{1}{U_{\Sigma}} \sum_{i=1}^e t_{i2} \leq T_{\max}$.

Все сформулированные выше задачи являются NP-трудными, т.е. в общем случае для их решения не существует точных полиномиальных алгоритмов [9]. В книгах [10, 11] можно найти описание трех точных псевдополиномиальных алгоритмов решения задач (1)-(3) и (14)-(16), основанных на методах ветвей и границ и динамического программирования с использованием Лагранжевой и LP релаксации. Листинги программ этих алгоритмов (exрknap, minknap и combo) на C++ приведены на странице D. Pisinger в Интернете (www.diku.dk/~pisinger/). В этих же книгах приведены также полностью полиномиальные приближенные схемы решения задачи 0-1 КР (FPTAS — Fully Polynomial Time Approximation Scheme) — алгоритмы, которые за полиномиальное время от размера входа задач и $1/\varepsilon$, позволяют получить $(1-\varepsilon)$ для задачи на максимум или $(1+\varepsilon)$ для задачи на минимум гарантированное приближенное решение, где ε — сколь угодно малое положительное число.

АЛГОРИТМ ПОЛНОГО ПЕРЕБОРА НА ОСНОВЕ КОДОВ ГРЕЯ

Для полного перебора вариантов решения задач (1)-(3) и (14)-(16) будем использовать алгоритм генерации последовательности двоично-отраженных n -разрядных кодов Грея, предложенный Эрлихом [12]. Алгоритм дает возможность в процессе решения эффективно вычислять значения целевой функции (1) или (14) и ограничения (2) или (15). Двоично-отраженный (зеркальный, рефлексивный) код Грея определяется по следующим рекурсивным правилам:

$$B_0 = "", \quad B_{n+1} = 0B_n 1B_n^r, \quad n = 0, 1, 2, 3, \dots, \quad (\text{разряды добавляются слева}) \text{ или}$$

$$B_0 = "", \quad B_{n+1} = B_n 0B_n^r 1, \quad n = 0, 1, 2, 3, \dots, \quad (\text{разряды добавляются справа}).$$

где $B_0 = ""$ — пустая строка, B_n — бинарная последовательность Грея, состоящая из n - битовых строк, $0B_n$ и $B_n 0$ — последовательность B_n с префиксом 0 в начале и конце каждой строки, $1B_n^r$ и $B_n^r 1$ — последовательность B_n в обратном порядке с префиксом 1 в начале и

конце каждой строки. Поскольку последняя строка в B_n эквивалентна первой строке в B_n^r , то ясно, что на каждом шаге B_{n+1} изменяется ровно один бит, если B_n обладает тем же свойством. С каждым шагом длина строк увеличивается на 1, а их количество — вдвое. Таким образом n -разрядный код Грея есть упорядоченная циклическая последовательность 2^n n -разрядных строк, в которой последовательные строки различаются только одним битом (разрядом). В алгоритмах полного перебора для вычисления значений различных функций с помощью кодов Грея удобно эти коды представлять упорядоченным списком номеров разрядов, изменяющих свое значение на противоположное при переходе от текущей строки к следующей. Такую последовательность переходов P_n можно определить по следующим рекурсивным правилам: $P_1 = 1$, $P_n = P_{n-1}, n, P_{n-1}$, $n = 2, 3, 4, \dots$. Длина последовательности P_n равна $2^n - 1$, а нумерацию разрядов в последовательности можно выполнять справа налево или наоборот. Например, для $n = 4$ и начальной строки 0000 $P_4 = 1, 2, 1, 3, 1, 2, 1, 4, 1, 2, 1, 3, 1, 2, 1$, а ее длина равна $2^4 - 1 = 15$. Соответствующая последовательность бинарных строк при нумерации разрядов слева направо имеет вид: 0000, 1000, 1100, 0100, 0110, 1110, 1010, 0010, 0011, 1011, 1111, 0111, 0101, 1101, 1001, 0001. При нумерации разрядов справа налево получим перевернутую последовательность, соответствующую первому рекурсивному определению B_{n+1} . Интересно заметить, что если построить граф, вершины которого соответствуют бинарным последовательностям длины n , а ребра соединяют две вершины отличающиеся только одним разрядом, то такой граф представляет двоичный n -мерный куб. При этом построенная бинарная последовательность соответствует гамильтонову пути в таком графе.

Для генерации последовательности переходов P_n в бинарной строке $B = \| b_i \|$, $i = \overline{1, n+1}$, определим вектор указателей $P = \| p_i \|$, $i = \overline{1, n+2}$, который имитирует стек для рекурсивного определения P_n . Оптимальное решение x_i^* , $i = \overline{1, n}$ будем сохранять в векторе $B^{opt} = \| b_i^{opt} \|$, $i = \overline{1, n+1}$, где: если $b_i^{opt} = 0$, то $x_i^* = 0$; если $b_i^{opt} = 1$, то $x_i^* = 1$; $i = \overline{1, n}$. Размерности векторов P , B и B^{opt} увеличены соответственно на две и одну единицу из-за специфики работы алгоритма. Знак « \leftarrow » означает операцию присваивания.

Алгоритм ОРТ1 с частичным пересчетом целевой функции и ограничения для задачи (1)-(3)

1. $CSUM \leftarrow c_i$; $ASUM \leftarrow a_i$; $CSUMOPT \leftarrow 0$; $B \leftarrow 0$; $B^{opt} \leftarrow 0$.
2. Для $\{i \mid i = \overline{1, n+2}\}$ выполнить $p_i \leftarrow i$.
3. $i \leftarrow 1$.
4. Пока $i < n+1$ выполнять шаги 5-7.
5. Если $b_i = 0$, то $CSUM \leftarrow CSUM - c_i$; $ASUM \leftarrow ASUM - a_i$, иначе $CSUM \leftarrow CSUM + c_i$; $ASUM \leftarrow ASUM + a_i$.
6. Если $ASUM \leq W$, то если $CSUM > CSUMOPT$ выполнить: $CSUMOPT \leftarrow CSUM$; $B^{opt} \leftarrow B$; $ASUMOPT \leftarrow ASUM$.
7. $i \leftarrow p_i$; $b_i \leftarrow 1 - b_i$; $p_1 \leftarrow 1$; $p_i \leftarrow p_{i+1}$; $p_{i+1} \leftarrow i+1$.
8. Конец алгоритма. Вывести значения: $\max \sum_{i=1}^n c_i x_i^* = CSUMOPT$;
 $\sum_{i=1}^n a_i x_i^* = ASUMOPT$; W ; B^{opt} .

В варианте решения задачи с полным пересчетом целевой функции и ограничения (**алгоритм ОПТ2**) шаг 1 заменится на

1. $CSUMOPT \leftarrow 0$; $B \leftarrow 0$; $B^{opt} \leftarrow 0$,
а шаг 5 заменится на
5. $CSUM \leftarrow 0$; $ASUM \leftarrow 0$. Для $\{j \mid j = \overline{1, n}\}$ выполнить $CSUM \leftarrow CSUM + c_j * b_j$; $ASUM \leftarrow ASUM + a_j * b_j$.

Для алгоритма с частичным пересчетом целевой функции и ограничения для задачи (14)-(16) псевдокод будет таким:

1. $T_\Sigma \leftarrow 0.0$; $C_\Sigma \leftarrow 0.0$; $D_\Sigma \leftarrow 0.0$.
2. Для $\{i \mid i = \overline{1, e}\}$ выполнить: $T_\Sigma \leftarrow T_\Sigma + t_{i1}$; $C_\Sigma \leftarrow C_\Sigma + c_{i1}$;
 $D_\Sigma \leftarrow D_\Sigma + \Delta c_i$.
3. $D_\Sigma \leftarrow C_\Sigma + D_\Sigma$; $T_s \leftarrow T_{\max} \times U_\Sigma$; $B \leftarrow 0$; $B^{opt} \leftarrow 0$; $C_\Sigma \leftarrow C_\Sigma + \Delta c_1$;
 $T_\Sigma \leftarrow T_\Sigma - \Delta t_1$.
4. Для $\{i \mid i = \overline{1, e+2}\}$ выполнить $p_i \leftarrow i$.
5. Пока $i < e+1$ выполнить пп. 6-8.
6. Если $b_i = 0$, то $C_\Sigma \leftarrow C_\Sigma - \Delta c_i$; $T_\Sigma \leftarrow T_\Sigma + \Delta t_i$, иначе $C_\Sigma \leftarrow C_\Sigma + \Delta c_i$;
 $T_\Sigma \leftarrow T_\Sigma - \Delta t_i$.
7. Если $T_\Sigma \leq T_s$, то если $C_\Sigma < D_\Sigma$ выполнить: $D_\Sigma \leftarrow C_\Sigma$; $B^{opt} \leftarrow B$;
 $t_{av} \leftarrow T_\Sigma / U_\Sigma$.
8. $i \leftarrow p_i$; $b_i \leftarrow 1 - b_i$; $p_1 \leftarrow 1$; $p_i \leftarrow p_{i+1}$; $p_{i+1} \leftarrow i+1$.
9. Конец алгоритма. Вывести значения: $\min C_\Sigma = D_\Sigma$; t_{av} ; B^{opt} .

ЭКСПЕРИМЕНТАЛЬНОЕ СРАВНЕНИЕ АЛГОРИТМОВ ОРТ1 И ОРТ2

Сравнение быстродействия алгоритмов ОРТ1 и ОРТ2 проводилось для задач (1) - (3) и (14) - (16) на примерах, сгенерированных датчиком псевдослучайных чисел. Здесь приводятся результаты решения только для задачи (1)-(3), так как для задачи (14)-(16) были получены аналогичные по характеристикам данные. Для всех размерностей задачи (1) - (3), которая изменялась от $n = 5$ до $n = 36$, стоимости предметов c_i и их веса a_i генерировались в диапазоне от 5 до 10 и от 1 до 20 соответственно. При проведении эксперимента проверялась также точность решения задачи (1) - (3) «жадным» эвристическим алгоритмом ОРТ3 с временной сложностью $O(n \log n)$. Алгоритм основан на предварительном упорядочении предметов в заданном наборе по невозрастанию их удельных стоимостей c_i / a_i , $i = \overline{1, n}$ с последующим выбором предметов в ранец до тех пор, пока выполняется ограничение на размер ранца W . Хорошо известно, см. например, [11,13], что решения $ОРТ3 = \max \sum_{i=1}^n c_i x_i^*$, полученные «жадным» алгоритмом, могут отличаться от оптимальных не более чем в два раза, при выборе окончательной стоимости ранца из условия $\max \sum_{i=1}^n c_i x_i^* = \{\max \sum_{i=1}^n c_i x_i^*, \max c_i\}$, где $\max c_i$, $i = \overline{1, n}$ — максимальная стоимость предмета в заданном наборе.

На рис. 1 показано время решения задачи (1) - (3) (с точностью до двух знаков после запятой) на ПК с тактовой частотой 2.66 GHz для алгоритмов ОРТ1 и ОРТ2 с частичным и полным пересчетом целевой функции и ограничения. Как видно из рис. 1, алгоритм ОРТ1 может применяться для практических расчетов в схемах ветвления, когда количество переменных в узлах дерева ветвления не превышает 35 (для $n = 35$ время счета около 8 минут). Алгоритм ОРТ1 для вариантов 5-10 быстрее алгоритма ОРТ2 в среднем в 7 раз.

На рис. 2 приведены результаты решения задачи (1) - (3) точным алгоритмом ОРТ1 и «жадным» алгоритмом ОРТ3 (результаты решений алгоритмов ОРТ1 и ОРТ2 естественно совпадают).

Значения $\sum_{i=1}^n a_i x_i^* = ASUMOPT$ показаны для алгоритма ОРТ1. Значения целевой функции, полученные «жадным» алгоритмом, отличаются от оптимальных в пределах DEL% от 0 до 12,5 %. Алгоритм ОРТ3 имеет полиномиальную оценку временной сложности и его можно применять на практике для решения задачи (1) - (3) большой размерности, когда лицу принимающему решение необходимо

достаточно быстро получить приближенное значение целевой функции при ограниченных вычислительных ресурсах.

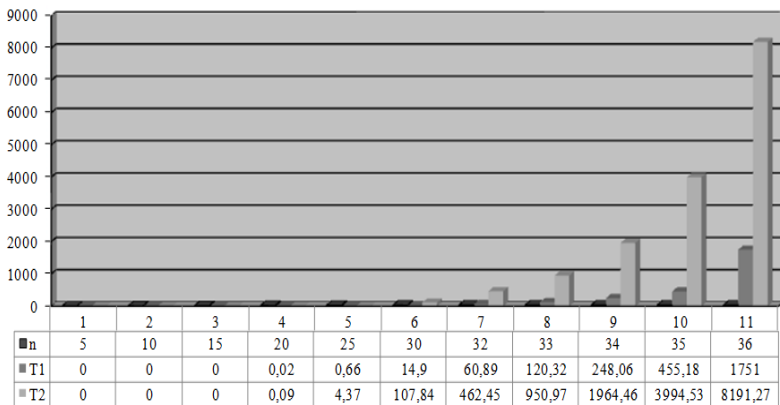


Рис.1 – Время решения задачи в сек. С частичным (T1, алгоритм OPT1) и полным (T2, алгоритм OPT2) пересчетом целевой функции и ограничений

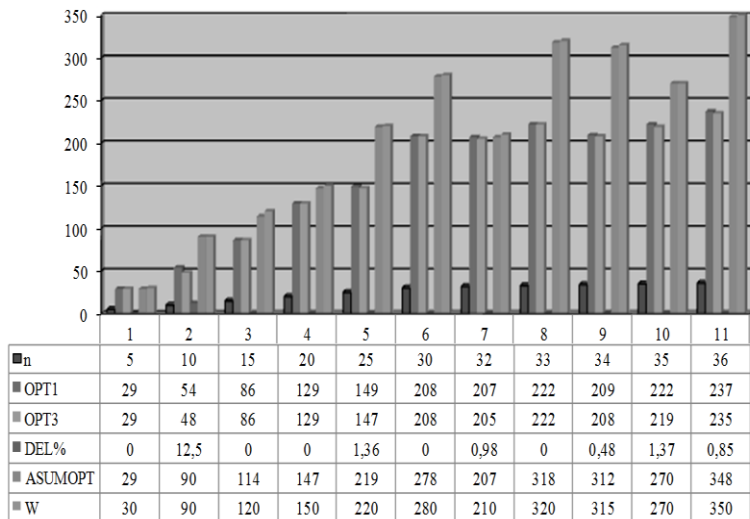


Рис.2 – Результаты оптимального (OPT1) и приближенного (OPT3) решения задачи

Так, например, задача (1)-(3) была решена для $n = 10000$ с теми же границами изменения значений c_i и a_i за 0,11 сек. При этом OPT3 = 70598, ASUMOPT = 103985 при $W = 104000$.

Все программы написаны на языке Фортран в среде Microsoft Developer Visual Studio и могут быть адаптированы для работы в системе параллельного программирования Intel® Parallel Studio XE 2018, в которую вошли последние версии компиляторов C/C++ и Фортран (<https://software.intel.com/ru-ru/try-buy-tools>).

ВЫВОДЫ

В статье показано, как можно применять на практике двоично-отраженные коды Грея для решения комбинаторных задач с псевдодулевыми функциями, когда количество переменных в узлах дерева ветвления решающего алгоритма не превышает 35. На примере решения двух комбинаторных задач с булевыми переменными продемонстрирована вычислительная эффективность предложенного алгоритма полного перебора решений с ограниченным пересчетом значений целевой функции и ограничений, и он может применяться на практике в различных схемах разветвления решающего алгоритма.

СПИСОК ЛИТЕРАТУРЫ

- [1] Васянин В.А., Ушакова Л.П. Коды Грея в задачах комбинаторной оптимизации // Математичне моделювання в економіці, 2019. – № 1-2. – С. 63 – 69.
- [2] Trofymchuk O.M., Vasyanin V.A. Choosing the Capacity of Arcs with Constraint on Flow Delay Time // Cybernetics and Systems Analysis, 2019. – Vol. – 55. – Issue 4. – P. 561 –569. DOI: <https://doi.org/10.1007/s10559-019-00165-0>
- [3] Gray F. Pulse code communication // U.S. Patent 2632058, March 17, 1953.
- [4] Гарднер М. Математические головоломки и развлечения: 2-е изд., испр. и дополн. / Пер. с англ. – М.: «Мир», 1999. – 447 с.
- [5] Knuth D.E. The Art of Computer Programming. Volume 4A / Combinatorial Algorithms, Part 1. - Addison Wesley Longman, Inc., 2011. – 933 p.
- [6] Trofymchuk O.M., Vasyanin V.A. Simulation of Packing, Distribution and Routing of Small-Size Discrete Flows in a Multicommodity Network // Journal of Automation and Information Sciences, 2015. – Vol. – 47. – Issue 7. – P. 15 – 30.
- [7] Vasyanin V.A. Problem of Distribution and Routing of Transport Blocks with Mixed Attachments and Its Decomposition // Journal of Automation and Information Sciences, 2015. – Vol. – 47. – Issue 2. – P. 56 – 69.

- [8] Васянин В.А., Трофимчук А.Н., Ушакова Л.П. Экономико-математические модели задачи распределения потоков в многопродуктовой коммуникационной сети // Математичне моделювання в економіці, 2016. – № 2. – С. 5 – 21.
- [9] Garey M.R., Johnson D.S. Computers and Intractability: A Guide to the Theory of NP-Completeness. - W. H. Freeman & Co. New York, NY, USA, 1979. – 338 p.
- [10] Martelo S., Toth P. Knapsack problems: algorithms and computer implementations. - Great Britain: Wiley, 1990. – 296 p.
- [11] Kellerer H., Pferschy U., Pisinger D. Knapsack Problems. - Springer-Verlag Berlin Heidelberg, 2004. – 548 p.
- [12] Bitner J.R., Ehrlich G., Reingold E.M. Efficient Generation of the Binary Reflected Gray Code and its Applications // Comm. ACM, 1976. – 19. – P. 517 – 521.
- [13] Кузюрин Н.Н., Фомин С.А. Эффективные алгоритмы и сложность вычислений. - М: МФТИ, 2008. – 326 с.

FULL SEARCH ALGORITHMS IN THE SCHEMES OF BRANCHING OF COMBINATORIAL PROBLEMS WITH PSEUDOBOOLEAN FUNCTIONS

Trofymchuk O., Vasyanin V., Ushakova L.

The article provides useful information for developers of algorithms and programs on the use of Gray codes for solving combinatorial problems with pseudoBoolean functions. As an example of the effectiveness of the use of these codes, the solution on two combinatorial problems with Boolean variables with a full search of the solutions is considered. The results of an experimental study are presented, which show that Gray codes can be practically applied in branching schemes, for example, in the branch and bound method, when the number of variables in the branch nodes of the decision algorithm does not exceed 35.

Keywords. Gray codes, combinatorial optimization problems, problem solving time

ИНФОРМАЦИОННАЯ ТЕХНОЛОГИЯ ПОСТРОЕНИЯ ПОВЕДЕНЧЕСКИХ МОДЕЛЕЙ МУЛЬТИАГЕНТНЫХ РОБОТОТЕХНИЧЕСКИХ СИСТЕМ

Кораблев В.А., Мазурок Т.Л.

Данное исследование имеет целью выведение комплекса задач и систематизацию подходящих для их решения методологий в ходе выработки единого теоретического базиса для моделирования мультиагентных роботизированных систем (МАРС). Оно имеет характер экспериментально-размышления, предметом которого постепенно становятся потребность в подобных системах, необходимые технологические решения и оптимальный путь реализации подобных систем на примере избранной задачи. Таким образом, будут рассмотрены существующие проблемы, которые могут быть решены благодаря внедрению подобных систем, их текущие методы решений и альтернатива. Описаны как оптимальные, для внедрения в будущий производственный процесс, существующие теоретические наработки внутри дисциплины, так и рекомендуемые направления для исследовательских работ с указаниями на конкретные моменты, как использование инклинометрической техники, либо внедрения методов когнитивного интеллекта, так и с ссылками на работы конкретных специалистов, что разбирались в данных вопросах. Будут приведены данные собственных теоретических разработок с учетом интернационального опыта и синтеза смежных дисциплин, что представляют собой выведение доктрины для стандартизации МАРС, в случае их использования в реальных условиях, для задач по разрешениям кризисов, связанных с опасностью жизни человека. Приложены рекомендации к аппаратной и к математической, программной частям подхода к реализации подобных систем с указанием сильных и слабых сторон избранных методов. Сделан упор на централизованную систему управления агентами, протоколы связи, требуемое сенсорное оборудование и прочее.

Ключевые слова. Мультиагентная система, мультиагентная роботизированная система, когнитивный интеллект, централизованный подход, поведенческая модель.

ВВЕДЕНИЕ

Термин «агент» встречается практически повсеместно, если говорить о научных, либо прикладных дисциплинах. Также, как различны методы этих родов деятельности, так в них различается и значение термина «агент», кроме следующих моментов: способность к реакции на факторы среды, возможность влияния, изменяющего среду, возможность взаимодействия с прочими агентами. Эта основа перекочевала и в агентно-ориентированный подход в программировании, где дополнилась спецификами, детерминированными этой дисциплиной.

Применение агентно-ориентированного подхода – следующий логический шаг развития робототехники, объясняемый преимуществом данной дисциплины у кибернетики и информатики. Что после некоторых манипуляций можно представить в виде лакановского преобразования виртуального в реальное, но в прикладном смысле.

Любой подобный переход обязательно сопровождается рядом ограничений, вызванными свойствами среды, как известными, так и скрытыми, как условными, так и безусловными. Это приводит к сужению спектра методик реализации данного подхода. Так, в нашем случае, из среды, где единственным ограничением является время, мы переходим в реальный мир с реальными проблемами.

Часто последний факт опускается в угоду возможности теоретической спекуляции, но при действительной необходимости реализации проекта придется столкнуться со следующими факторами: непосредственная задача, квалификации операторов, угрозы среды, физические законы, ориентирование в условиях неполной картины мира, критические ситуации, поломки, размерность устройств и прочее, что по большей части упирается в текущий уровень доступных технологий.

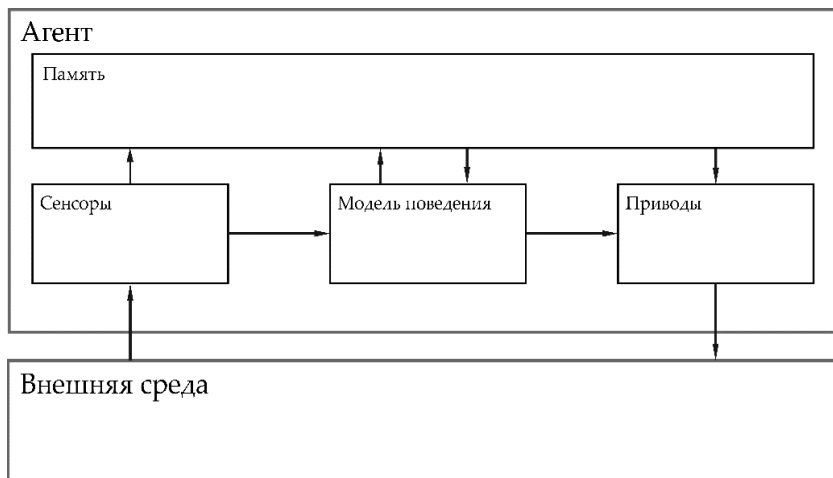


Рис. 1 – Структурная схема агента МАРС

Многоагентные системы или мультиагентные системы - это направление искусственного интеллекта, для решения сложной задачи или проблемы использует системы, состоящие из множества взаимодействующих агентов. Это направление искусственного

интеллекта, активно развивается и в настоящее время еще находится в стадии становления.

В многоагентных системах весь спектр задач по определенным правилам распределяется между всеми агентами, каждый из которых считается членом организации или группы. Распределение задач означает присвоение каждому агенту некоторой роли, сложность которой определяется исходя из возможностей агента.

Для организации процесса распределения задачи в многоагентных системах создается или система распределенного решения проблемы, или децентрализованный искусственный интеллект. В нашем варианте процесс декомпозиции глобальной задачи и обратный процесс композиции найденных решений происходит под управлением некоторого единого «центра». При этом многоагентная система проектируется строго сверху вниз, исходя из ролей, определенных для агентов и результатов разбиения глобальной задачи на подзадачи.

МАРС можно рассматривать как один из вариантов реализации мультиагентных систем (МАС), и, следовательно, каждый робот-агент должен обладать следующими свойствами: активность, способность к организации и реализации действий; реактивность, способность воспринимать состояние среды; автономность и относительная независимость от окружающей среды; общительность, что вытекает из необходимости решать свои задачи совместно с другими агентами и обеспечивается развитыми протоколами коммуникации; целеустремленность, которая предусматривает наличие собственных источников мотивации.

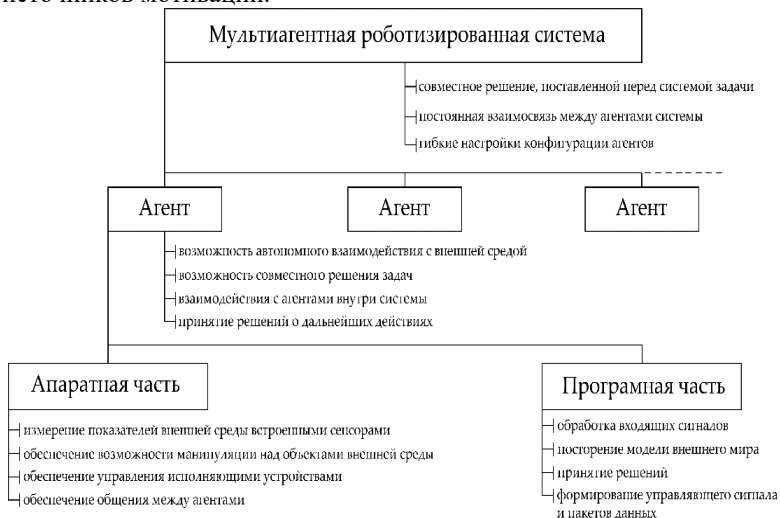


Рис. 2 – Структурная схема МАРС

В рамках данного доклада описывается вариант МРС, состоящий из унифицированных единиц, то есть множества идентичных агентов. Каждый агент представляет собой совокупность аппаратно-программных компонентов.

Аппаратная часть проводит измерения параметров внешней среды сенсорами, что обеспечивает возможности влияния на окружающую среду, управление исполнительными устройствами и возможность "общения" между агентами.

Программная часть выполняет задачи: анализа входных сигналов с сенсоров, а также построения модели внешнего мира, принятия решений, формирования управляющих сигналов и пакетов передачи данных.

МЕТОДИКА ИССЛЕДОВАНИЯ

Для начала следует определиться со спектром проблем, на которых стоит заострить внимание.

В каждый условный момент формируется некое множество задач, решения которых еще не сформированы, а решения, что не теряют актуальности для уже известных задач являются традиционными конструктами. Удобство последнее поясняется лишь за счет привычки. То есть необходимо бороться с неопределенностью, в прочем, как и с традицией, хоть последнее – заранее проигрышный вариант, что ведет к необходимости компромисса.

Для примера рассмотрим несколько вариантов возможного внедрения робототехнической системы, но стоит помнить, что специфика мультиагентного подхода подразумевает применение технологии групповой робототехники, то есть наделяет машины некоторой степенью автономии:

1. Система, заменяющая человека/группу людей на рабочем месте с малой степенью ответственности;
2. Система, заменяющая человека/группу людей на рабочем месте с высокой степенью риска и угрозы безопасности индивида;
3. Система, для работы в экстремальных условиях, заменяющая удаленного оператора;
4. Система, что дублирует/взаимодействует с человеком обеспечивая тому безопасность на рабочем месте с высокой степенью риска/ошибки.

Первый случай – несущественен, из-за наличия готовых (действенных и безопасных) решений и фактора реакции общества.

Второй – решается переходом к третьему, в противном случае сталкивается с общественным недоверием.

Третий же сложен даже с этической точки зрения, так как подразумевает упразднение рабочего места, требующего высочайший

уровень квалификации и узкой подготовленности. Также третий случай подразумевает наличие кризисных ситуаций, пути решения которых могут быть найдены лишь эмпирически, то есть появляется необходимость давать машине время на обучение при действительном ущербе, либо дублировать ее тем же специалистом.

Эти варианты крайне обобщены и абстрагированы от множества специфических отраслей человеческой деятельности, но они безусловно являются наиболее распространенными.

Четвертый же подразумевает некий синтез между навыками, безусловно квалифицированными рабочими таких отраслей, как: армия, полиция, МЧС, медицина, тяжелое сырьевое производство и т.д.; и мощностями, коими владеет роботизированное устройство по определению.

Его то и стоит рассматривать в первую очередь. Для удобства, далее подобная система будет упоминаться, как автономный роботизированный инструментарий (АРИ).

Следующим этапом становится определение уровня централизации, диктуемой спецификой данной задачи. На практике централизованный и децентрализованный подходы в робототехнике предпочтительно рассматривать не как замкнутые доктрины, но как направления двунаправленной шкалы, где наша позиция зависит от весомости критериев, которые мы рассматриваем.

Так, тенденция к миниатюризации, что упирается в плато развития микропроцессоров, требует более централизованного подхода, так как приводит к уменьшению вычислительных мощностей отдельного агента. Механическая же простота задачи способствует унификации и приемам базового группового взаимодействия, что позволяет большую степень децентрализации.

В нашем случае мы также имеем штат квалифицированных сотрудников, что могут взаимодействовать со сложными компьютерными системами, что позволяет пойти на риск и перенести значительную часть высокоуровневой обработки данных в некий центральный контролирующий блок (ЦКБ). Такое решение налагает большие требования к системам связи между агентами, но также дает возможность виртуального моделирования среды и хода выполнения заданий, что сводит к минимуму количество «полевых» ошибок.

Восвобожденный полезный объем внутри конструкции физического агента следует применить для установки более эффективных сенсоров и средств быстрого (реактивного) реагирования на угрозы, как для робота, так и для людей в непосредственной близости.

Следовательно, в этом случае модель более склонна к централизации, что позволяет удешевить производство и облегчить

ремонт роботов, что полезно, учитывая специфику кризисов в описанных выше отраслях. Этого направления и будем придерживаться далее.

Итак, предлагается разработка специальной информационной технологии, что интегрируется в робототехнический комплекс для выполнения задач автоматизации и повышения эффективности его функционирования, путем построения поведенческих моделей мультиагентной системы (МАС) с использованием принципов централизации процессов анализа и управления, как составляющих виртуальной симуляции.

Стратегии централизованного управления предполагают сосредоточение всей совокупности командно-контрольных функций в едином блоке, что обеспечивает планирование и координацию целесообразных действий элементов группы агентов при решении общих прикладных задач. Соответствующая структура системы централизованного управления должна предусматривать наличие каналов двусторонней связи между командно-контролирующим органом и каждым из робототехнических устройств.

К числу основных преимуществ систем централизованного управления необходимо отнести существенное сокращение функциональной нагрузки на рядовых членов группы, когда множество важных и сложных по своей сути вопросов, как: анализ поставленной прикладной задачи; сбор, комплексирование и интерпретацию данных об особенностях текущей ситуации, рабочей обстановки и состояние внешней среды; планирования целесообразных действий и контроль их реализации конкретными исполнителями; целенаправленно переносятся на командно-контролирующий блок.

ЭКСПЕРИМЕНТАЛЬНАЯ ЧАСТЬ, АНАЛИЗ, ОБОБЩЕНИЕ И РАЗЪЯСНЕНИЕ СОБСТВЕННЫХ ДАННЫХ ИЛИ СРАВНЕНИЕ ТЕОРИЙ

Переходя к конкретике, следует выбрать отдельный случай внедрения подобной МАРС. Тут будет рассматриваться система, содействующая сотрудникам МЧС для разрешения кризисов в высотных зданиях, где количество этажей делает пожарные рукава и лестницы неэффективными, учитывая их максимальную длину. А также прочие трудности, возникающие в связи с нестандартными архитектурными решениями.

В такой ситуации, инструментом решения которой рассмотрим рой роботов, существенно расширяются возможности сотрудников МЧС. Унифицированные роботизированные единицы способны переносить снаряжение, а в случае необходимости объединяться в жесткие конструкции (предотвращающие обрушение подпорки, мосты, лестницы, пандусы и даже импровизированные тоннели, защищающие от огня и

обломков). Все это позволяет значительно снизить смертность, как среди пострадавших, так и сотрудников службы спасения.

Как известно, мультиагентную роботизированную систему (МАРС) можно рассматривать как один из вариантов реализации МАС, так что каждый робот-агент имеет все известные свойства агентов [1]. Системы управления такими сложными комплексами должны обеспечивать адаптивность робототехнических устройств к кругу решаемых задач, согласование выработки траекторий движения и др. Поэтому актуальной проблемой является повышение адаптивных свойств системы управления сложными робототехническими комплексами (СРК). Для полноценного функционирования таких систем необходимо совершенствование информационного обеспечения системы управления [2].

Итак, предлагается разработка специальной информационной технологии, что интегрируется в робототехнический комплекс, и предназначена для выполнения задач автоматизации, направленных на повышение эффективности функционирования СРК.

Такая информационная технология позволяет автоматизировать процесс построения поведенческих моделей мультиагентной системы, основанные на использовании принципов централизации процессов анализа и управления, как составляющей виртуальной симуляции (рис.3).

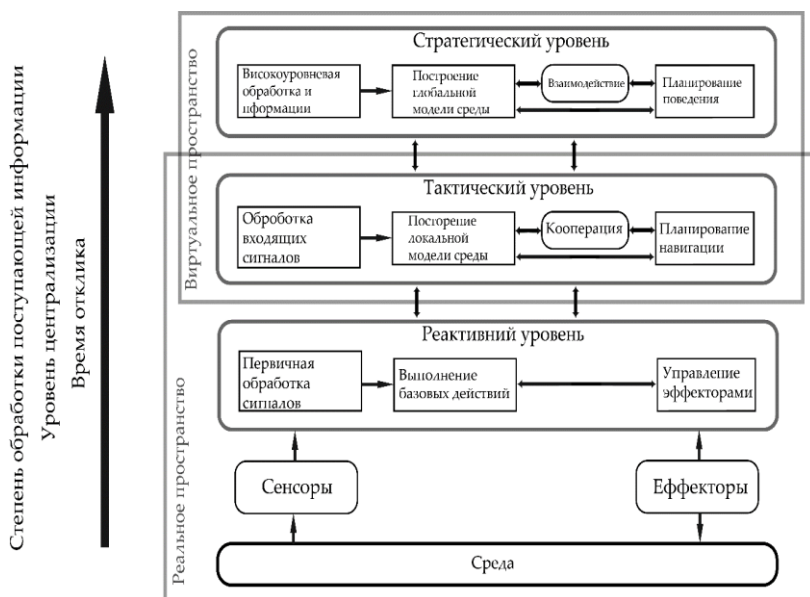


Рис. 3 – Структурная схема системы управления агента МАРС

Так, в информационной технологии, разрабатываемой предлагается расположить ЦКБ на удаленном сервере, и проводить стратегическое планирование внутри виртуальной среды, имитирующей реальное пространство (стратегический уровень).

Такой подход целесообразен при выполнении задач в среде с высокой степенью изучения, например, при наличии виртуализованных планов здания со всеми показателями (пути эвакуации, характеристики материалов, слабые и сильные элементы конструкции). Так, на основе первичного глобального формирования предполагаемого окружения и системы обратной связи с роботами-агентами, система позволяет динамично достраивать виртуальное пространство тождественное реальному и формировать алгоритм решения кризисной ситуации.

ЦКБ, благодаря высоким мощностям, может быстрее производить алгоритм решения задачи. Или же, если необходимо нестандартное решение, требующее эвристического подхода, есть возможность проработать с максимальной скоростью необходимое количество симуляций для получения решения приближенного к оптимальному, еще до непосредственного эмпирического эксперимента на местности [3].

Когда задача сформирована, она делится на подзадачи (тактический уровень) для каждой группировки роботов. Подзадачи имеют динамический характер, и они в значительной степени зависят от локальной модели пространства, сложившейся в реальном времени на основе данных сенсоров каждого агента, и предназначена для оперирования по обстоятельствам. На этом этапе происходит распределение задач навигации и манипуляции эффекторами агентов.

На данном уровне необходимо повысить коэффициент автономии. Необходимость в реакционных возможностях, в плане построения/перестройки маршрутов, либо порядка взаимодействий между агентами и объектами окружения, диктует потребность в использовании современных решений искусственного интеллекта [4].

Данная необходимость растет по мере повышения сложности проходимой местности и количества задач, что ставит оператор системы. Критические ситуации, для которых она подготавливается, априори несут в себе огромные требования по мультизадачной нагрузке группировок агентов, что может вызвать коллизии в протоколах их собственных, отдельных от непосредственного управляющего сигнала, каналах общения. Это потребует полевого тестирования и развития теперешних доктрин мультиагентных коммуникаций [5].

Более того, первые версии подобных систем необходимо создавать при обширной консультационной поддержке профессионалов, в отраслях которых оные будут задействованы. Это абсолютная необходимость, но она внедряет человеческий фактор, а точнее узкую область личного опыта конкретного специалиста, в подготовку машин,

от которых ожидается универсальность в подходе к выполнению задач. Первичное обучение нейросети может оказаться недостаточным, а значит необходимо внедрение адаптивных механизмов, и инструментов динамического полевого обучения [6]. А значит, потребуются новейшие технологии коммуникаций [7] и синхронизации симуляции реального пространства с данными протоколами [8].

Также для реализации подобного модуля необходимо интегрированное, как на уровне командного центра, так и отдельного агента, программное обеспечение, осуществляющее мягкие вычисления, и набор нестандартных для подобных систем датчиков, содержащих инклинометрические приборы высокой точности и скорости сбора данных [9].

Остается открытым вопрос организации реакционного поведения отдельного агента (реактивный уровень). Предполагается, что оптимально корректное выполнение задания агентом все время находится под угрозой, как внешней (механические препятствия, непосредственные угрозы и т.п.), так и внутренней (ошибка в навигации, повреждения самой единицы и т.п.). Эти факторы обуславливают необходимость определенного уровня автономности агента, что требует реализации малой когнитивной системы (МКС).

Здесь МКС будет ответственна за разработку всех возможных вариантов действий, которые должны быть сформированными в случае опасности для агента, оператора или постороннего человека, если нет соответствующей инструкции с более высокого уровня командной иерархии. Также во внекризисное время данная система будет отвечать за реактивное маневрирование агента, что необходимо для корректировки его положения в пространстве в соответствии с траекторией движения, отвечающей заданию с более высокого уровня [10].

ВЫВОДЫ

Подводя черту под всем вышесказанным, можно утверждать, что, учитывая сегодняшний технологический уровень (Так как элементом «прогресса» в науке должно пренебрегать из-за того, что следует рассматривать те факты, что есть, а не те, что должны быть.) модели МАРС направленные на решение специализированных под групповую робототехнику задач все еще стремятся в централизации.

Это обусловлено теперешним уровнем технологии нейросетей (что была деконструирована в угоду упрощения мат. аппарата) которую для оперирования в реальной среде следует подкреплять стоящими выше системами жесткого контроля, основанными на огромном массиве логических вентилей и, желательно, контролем со стороны оператора. Обусловлено доступностью необходимой по критерию

стрессоустойчивости аппаратуры на массовом рынке. Обусловлено безусловность наличия человеческого фактора в множестве формируемых решений.

Такая модель подразумевает многоуровневую структуру (стратегический, тактический и реактивный), где по нисходящей уменьшается сложность и глобальность принимаемых решений, но повышается отклик системы, что в конечном итоге приводит к системам балансировки и экстренной реакции на угрозы. Это требует включения на реактивном уровне технологии МКС, что будет обеспечивать имитацию инстинктивного (детерминированного свойствами отдельного агента) поведения.

Главный недостаток этой схемы заключается в ее потенциальной уязвимости, поскольку выход из строя ЦКБ неизбежно приводит к нарушению работоспособности системы в целом. Также проблемы могут возникнуть из-за экранирования рабочих участков и через сторонние препятствия для прохождения сигнала. При таком варианте возникает потребность в использовании сложных протоколов коммуникаций с дублирующими методами передачи данных и постоянной проверкой их актуальности, и борьбе с коллизиями.

Итогом же данного исследование была попытка индоктринации выработанной модели для дальнейшего развития МАРС внутри ее парадигмы, что должно направить мощности, как специалистов, так и меценатов на реальные рельсы.

СПИСОК ЛИТЕРАТУРЫ

- [1] Schelling T.C. Dynamic models of segregation // The Journal of Mathematical Sociology, 1971. – Vol. 1. – № 2. – P. 143 – 186.
- [2] Wooldridge M. An introduction to multiagent systems. – JOHN WILEY & SONS, LTD, 2002. – 484 p.
- [3] Sarkar A., Debnath N. Measuring complexity of Multi-Agent System architecture // IEEE 10th International Conference on Industrial Informatics, 2012. – P. 998 –1003.
- [4] Russell S., Norvig P. Artificial Intelligence: A Modern Approach. – 3rd. – Upper Saddle River, NJ, USA : Prentice Hall Press, 2009. – 1152 с.
- [5] Communicating Agents in Open Multi Agent Systems / T.R. Payne // Proceedings of 1st GSFC/JPL Workshop on Radical Agent Concepts (WRAC), 2002. – С. 365 – 371.
- [6] Xiao L., Greer D. Modeling, Auto-generation and Adaptation of Multiagent Systems // Proceedings of the Tenth CAiSE/IFIP8.1 International Workshop on Exploring Modeling Methods in Systems Analysis and Design (EMMSAD'05) . – Porto, Portugal, 06.2005. – С. 605 – 616.
- [7] Pitt J., Mamdani A. Communication Protocols in Multiagent Systems: A Development Method and Reference Architecture // Issues in Agent

Communication / под ред. F. Dignum, M. Greaves. — Berlin, Heidelberg: Springer Berlin Heidelberg, 2000. — С. 160 –177.

[8] Gehrke J.D., Schuldt A., Werner S. Quality criteria for multiagent-based simulations with conservative synchronisation // 13th ASIM Dedicated Conference on Simulation in Production and Logistics / под ред. M. Rabe. — Fraunhofer IRB Verlag, 2008. — С. 545 – 554.

[9] Рыжков И.В. Инклинометрические приборы. Конструкции и способы повышения точности // Saarbrucken, Deutschland: LAPLAMBERT Academic Publishing, 2016. — 274 с.

[10] Прокопчук Ю.А. набросок формальной теории творчества. - Днепр: Изд-во ПГАСА, 2017. — 452 с.

INFORMATION TECHNOLOGY FOR BUILDING BEHAVIORAL MODELS OF MULTIAGENT ROBOTECNICAL SYSTEMS

Korablev V.A., Mazurok T.L.

This study aims to derive a set of tasks and systematize methodologies suitable for their solution in the course of developing a unified theoretical basis for modeling multi-agent robotic systems (MARS). It has the character of a reflection experiment, the subject of which is gradually becoming the need for such systems, the necessary technological solutions and the optimal way to implement such systems using the example of a chosen task. Thus, existing problems that can be solved through the introduction of similar systems, their current methods of solutions and an alternative will be considered. It describes both optimal, for introducing into the future production process, existing theoretical developments within the discipline, and recommended areas for research with indications of specific points, both using inclinometric techniques, or introducing methods of cognitive intelligence, and with references to the work of specific specialists, that understood these issues. The data of our own theoretical developments will be presented taking into account international experience and the synthesis of related disciplines, which are the derivation of the doctrine for standardizing MARS, if used in real conditions, for tasks on resolving crises related to the danger of human life. Recommendations are attached to the hardware and mathematical, software parts of the approach to the implementation of such systems, indicating the strengths and weaknesses of the selected methods. Emphasis is placed on a centralized agent management system, communication protocols, required sensor equipment, etc.

Keywords. Multi-agent system, multi-agent robotic system, cognitive intelligence, centralized approach, behavioral model.

АВТОМАТИЗОВАНА ПІДГОТОВКА РАЦІОНАЛЬНИХ СХЕМ РОЗКРОЮ МАТЕРІАЛІВ ЗІ СКЛАДНОЮ КОНФІГУРАЦІЄЮ ЗОВНІШНІХ КОНТУРІВ НА ПЛОСКІ ГЕОМЕТРИЧНІ ОБ'ЄКТИ

Чупринка В., Грикун Д., Чупринка Н.

В роботі запропоновані математична модель та алгоритми, які були реалізовані в програмне забезпечення автоматизованого проектування раціональних схем розкрою матеріалів зі складною конфігурацією зовнішнього контуру на плоскі геометричні об'єкти. Це програмне забезпечення підвищує показники економії матеріалу і продуктивність праці. Для розробки математичної моделі задачі автоматизованого проектування раціональних схем розкрою матеріалів зі складною конфігурацією зовнішнього контуру на плоскі геометричні об'єкти було формалізовано її структурні компоненти: аналітичний опис зовнішнього контуру матеріалу зі складною конфігурацією зовнішнього контуру; аналітичний опис плоских геометричних об'єктів, для яких необхідно спроектувати раціональні схеми розкрою на матеріалі зі складною конфігурацією зовнішнього контуру; параметри, що однозначно визначають положення плоского геометричного об'єкту на площині; побудова щільних укладок для плоских геометричних об'єктів; побудова схеми розкрою в прямокутній області заданих розмірів за параметрами спроектованої щільної укладки; перенесення спроектованої схеми на матеріал зі складною конфігурацією зовнішнього контуру; інтерактивне коригування схеми розкрою, що спроектована в атоматичному режимі, для матеріалу зі складною конфігурацією зовнішнього контуру. Використавши розглянуті структурні компоненти було розроблено алгоритм та програмне забезпечення автоматизованого проектування раціональних схем розкрою матеріалів зі складною конфігурацією зовнішнього контуру на плоскі геометричні об'єкти

Ключові слова. Раціональний розкрій, матеріал, зовнішній контур, системне розміщення, площина, інтерактивне коригування

ВСТУП

У будь-якій галузі промисловості питання витрат матеріалу при виробництві завжди було дуже актуальним. Висока матеріалоємність та значна вартість використовуваних матеріалів роблять задачу мінімізації витрат особливо важливою для взуттєвої промисловості. При чому максимальна кількість втрат площі виникає при розкрої натуральних шкір. Тому у виробництво необхідно впроваджувати новітні технології, які базуються на застосуванні засобів автоматизованої обробки інформації та дозволяють підвищити показники економії матеріалу і продуктивність праці.

В більшості випадків процес підготовки раціональних схем розкрою листових матеріалів на деталі взуття виконуються вручну. Але

розвиток обчислювальної техніки та методів обчислювальної математики дозволяють виконувати ці етапи в автоматичному режимі.

Актуальність досліджень. Розробка методу проектування раціональних схем розкрою матеріалів зі складною конфігурацією зовнішнього контуру на плоскі геометричні об'єкти дозволило розробити математичне та програмне забезпечення для автоматизованого проектування раціональних схем розкрою натуральних матеріалів на деталей взуття. Це програмне забезпечення підвищує показники економії матеріалу і продуктивність праці.

Постановка задачі. Розробити математичну модель задачі автоматичного проектування раціональних схем розкрою матеріалів зі складною конфігурацією зовнішнього контуру на плоскі геометричні об'єкти. На основі запропонованої математичної моделі розробити алгоритми та програмне забезпечення для автоматизованого проектування раціональних схем розкрою.

Результати досліджень. Проектуванню раціональних схем розкрою матеріалів на плоскі геометричні об'єкти присвячено багато робіт. В роботах [1-3] представлені математичні моделі компактного розміщення опуклих плоских геометричних об'єктів. Але в більшості випадків деталі взуття є не опуклі плоскі геометричні об'єкти. В роботах [4-9] застосовуються генетичні алгоритми. Але ці алгоритми не завжди дають задовільний результат за обмежений час. Тому завданням цієї роботи є розробка методу автоматичного проектування раціональних схем розкрою матеріалів будь-якою конфігурацією зовнішнього контуру для плоских геометричних об'єктів з складною конфігурацією зовнішніх контурів. Але задача автоматичного проектування раціональних схем розкрою матеріалів зі складною конфігурацією зовнішнього контуру на плоскі геометричні об'єкти в такій постановці не розглядалася.

Для розв'язання цієї задачі необхідно математична постановка задачі, в якій сформулювати технологічні вимоги та обмеження.

Для розробки математичної моделі задачі автоматичного проектування раціональних схем розкрою матеріалів зі складною конфігурацією зовнішнього контуру на плоскі геометричні об'єкти необхідно формалізувати її структурні компоненти: аналітичний опис зовнішнього контуру матеріалу зі складною конфігурацією зовнішнього контуру; аналітичний опис плоских геометричних об'єктів, для яких необхідно спроектувати раціональні схеми розкрою на матеріалі зі складною конфігурацією зовнішнього контуру; параметри, що однозначно визначають положення плоского геометричного об'єкту на площині; побудова щільних укладок для плоских геометричних об'єктів; побудова схеми розкрою в прямокутній області заданих розмірів за параметрами спроектованої щільної укладки; перенесення спроектованої схеми на матеріал зі складною конфігурацією зовнішнього контуру;

інтерактивне коригування схеми розкрою, що спроектована в атоматичному режимі, для матеріалу зі складною конфігурацією зовнішнього контуру.

Для опису зовнішнього контуру матеріалу і дефектів будемо застосовувати кусково-лінійну апроксимацію. Зовнішня границя області Ω буде задаватися вершинами опукло-ввігнутого багатокутника $\{X_j^s, Y_j^s\}, j = 1..k$. Дефекти будуть задаватися також вершинами апроксимуючого багатокутника $\{Xp_j^l, Yp_j^l\}, j = 1..M(l); l = 1..r$.

де $M(l)$ – кількість вершин опукло-вгнутого апроксимуючого багатокутника для l -го дефекту;

r – кількість дефектів;

k – кількість вершин опукло-вгнутого апроксимуючого багатокутника для зовнішнього контуру матеріалу.

Для опису зовнішнього контуру плоских геометричних об'єктів $\{Q_i\}, i = 1, 2, \dots, m$ будемо також застосовувати кусково-лінійну апроксимацію. При такій апроксимації будь-який плоский геометричний об'єкт S_i буде однозначно визначений координатами вершин опукло-ввігнутого багатокутника, тобто $\{X_j^i, Y_j^i\}, j = 1, \dots, N(i)$.

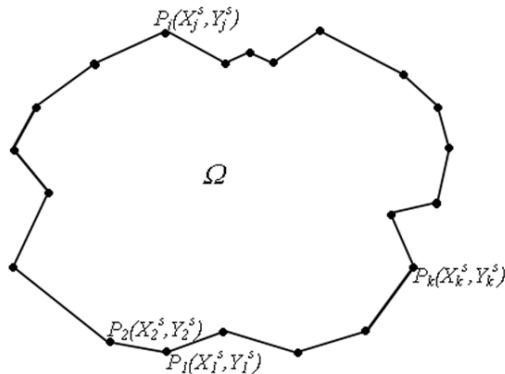


Рис. 1 – Апроксимація зовнішнього контуру матеріалу складної конфігурації

Параметри, які однозначно визначають положення плоского геометричного об'єкту на площині:

N_i – ідентифікатор, за яким однозначно визначається необхідний плоский геометричний об'єкт;

X_i, Y_i – координати полюсу плоского геометричного об'єкту S_i (будь-якої фіксованої точки на плоскому геометричному об'єкту відносно якого визначені координати апроксимуючого багатокутника для цього об'єкту) в системі координат, що пов'язана з площиною;

φ_i – кут повороту плоского геометричного об'єкту S_i відносно його основного положення.

ПОБУДОВА ЩІЛЬНОЇ УКЛАДКИ ПЛОСКИХ ГЕОМЕТРИЧНИХ ОБ'ЄКТІВ.

Розглянемо на площині об'єкти S_1 та S_2 . Нехай $\text{int } S = S - S^\wedge$, де S^\wedge – границя об'єкта S . Об'єкти S_1 та S_2 не перетинаються, якщо

$$\text{int } S_1 \cap \text{int } S_2 = 0 \quad (1)$$

Якщо одночасно виконується умова

$$S_1 \cap S_2 \neq 0 \quad (2)$$

то об'єкти S_1 та S_2 називаються щільно розміщеними.

Щільно розміщені об'єкти не мають спільних внутрішніх точок, але обов'язково мають спільні граничні точки.

Система об'єктів $S_i, i=1..p$, утворюють на площині укладку, якщо для кожної пари об'єктів із цієї системи виконуються умови їх взаємного неперетину:

$$\text{int } S_n \cap \text{int } S_m = 0, n \neq m, n, m = 1..p \quad (3)$$

та для будь-якого об'єкта $S_i, i=1..p$ знайдеться хоч один об'єкт S_q , де $q \in [1..p], q \neq i$, який дотикається до об'єкта S_i .

Розглянемо систему об'єктів $\bigcup_{n,m} S_1^{nm}$ та $\bigcup_{n,m} S_2^{nm}$, де $n, m = 0, \pm 1,$

$\pm 2, \pm 3, \dots \neq k, \dots$, які складаються із трансляції $S_1^{nm} = S_1 + na_1 + ma_2$ та $S_2^{nm} = S_2 + na_1 + ma_2 + g$ об'єктів S_1 та S_2 на вектори подвійної решітки $W = W(a_1, a_2, g)$ [10-11]. Якщо ця система є укладкою, то така укладка називається укладкою об'єктів S_1 та S_2 , виконаної по подвійній решітці $W = W(a_1, a_2, g)$. Подвійна решітка $W = W(a_1, a_2, g)$ в цьому випадку є допустимою для укладки об'єктів S_1 та S_2 . Подвійна решітка представляє собою дві однакові одинарні решітки $A_1 = A(a_1, a_2)$ та $A_2 = A(a_1 + g, a_2 + g)$, які зміщені одна відносно іншої на вектор зсуву решітки g . У вузлах решітки A_1 розміщаються об'єкти S_1 , а у вузлах решітки A_2 розміщаються об'єкти S_2 . Абсолютна величина визначника,

який складений із векторів решітки, називається визначником решітки W та позначається $\det W$, де [12]:

$$\det W = \begin{vmatrix} a_{1x} & a_{1y} \\ a_{2x} & a_{2y} \end{vmatrix} = |a_{1x}a_{2y} - a_{2x}a_{1y}|. \quad (4)$$

Щільність $\delta_s(W)$ решітчастої укладки можна характеризувати за допомогою співвідношення:

$$\delta_s(W) = (|S_1| + |S_2|) / \det W, \quad (5)$$

де $|S_1|$ та $|S_2|$ - відповідно площі плоского геометричного об'єкта S_1 та S_2 , $\det W$ - визначник решітки $W = W(a_1, a_2, g)$, за якою виконана укладка. Із наведеного співвідношення видно, що щільність $\delta_s(W)$ решітчастої укладки тим вища, чим менша площа паралелограма, сторонами якого є базові вектори решітки a_1 та a_2 .

В задачі побудови найщільнішої решітчастої укладки можна виділити наступні структурні компоненти:

- аналітичний опис умов взаємного неперетину деталей в укладці;
- аналітичний опис системи суміщення деталей в укладці;
- математичний опис множини допустимих розв'язків задачі;
- аналітичне представлення функції цілі.

Для аналітичного представлення умов взаємного не перетину деталей при їх суміщенні використаємо годограф вектор-функції щільного розміщення (ГВФЩР) [10-11]. Нехай S_1 та S_2 - два об'єкти, які зберігають постійну взаємну орієнтацію. Позначимо через O_1 та O_2 полюси об'єктів, які вибрані в довільних точках даних об'єктів. Тоді $X_1O_1Y_1$ та $X_2O_2Y_2$ - системи координат, які жорстко зв'язані з об'єктами S_1 та S_2 відповідно. Без обмеження відповідні координатні осі можна вважати направленими однаково. Припустимо, що об'єкт S_1 нерухомо закріплений на площині, а об'єкт S_2 - рухомий. Розглянемо множини можливих щільних положень об'єкта S_2 по відношенню до об'єкта S_1 . Кожне таке положення характеризується вектором $r_{12} = O_1O_2$. Вектор-функція, що ставить у відповідність щільному положенню об'єктів S_1 та S_2 значення вектора r_{12} при умові, що об'єкт S_1 нерухомий, називається годографом вектор-функції Γ_{12} щільного розміщення об'єкту S_2 відносно об'єкта S_1 .

Математичний опис множини допустимих розв'язків задачі проектування найщільнішої решітчастої укладки. Для математичного опису множини допустимих розв'язків задачі побудови найщільнішої

решітчастої укладки використаємо апарат ГВФЦР. Очевидно, що множиною допустимих розв'язків цієї задачі буде множина допустимих подвійних решіток $W^i = W(a^i_1, a^i_2, g^i)$, де $i=1, 2, \dots, q$. Визначивши область зміни параметрів a^i_1, a^i_2, g^i подвійних решіток W^i ми однозначно визначимо множину допустимих розв'язків. Подвійна решітчаста укладка однозначно визначається взаємним розміщенням трьох сусідніх рядів (рис. 2). За розміщення деталей в рядах відповідає вектор a^i_1 , за взаємне розміщення рядів відповідають вектори a^i_2 та g^i або вектори p^i та g^i . Вектор p^i визначає положення верхнього ряду, а вектор g^i визначає положення нижнього ряду відносно базового середнього ряду укладки. Як видно із рис. 2 $a^i_2 = B_1 A_1 = B_2 A_2 = \dots = B_5 A_5$, то тоді a^i_2 визначається через вектори p^i та g^i наступним чином $a^i_2 = p^i - g^i$.

Розглянемо множини допустимих розв'язків для задачі щільної укладки плоских геометричних об'єктів S_1 та S_2 , тобто для задачі побудови найщільнішої решітчастої укладки деталей виду S_1 та S_2 (деталі S_1 повернуті ряду на кут α та деталі S_2 повернуті ряду на кут β відносно їх вихідного положення). В цьому випадку необхідно впевнитись, що задача має розв'язок. Для цього побудуємо ГВФЦР $r = r_{11}(\theta)$ для деталі S_1 з самою собою та ГВФЦР $r = r_{22}(\theta)$ для деталі S_2 з самою собою. Знайдемо $\min_{r_{11}}, \min_{r_{22}}, \max_{r_{11}}, \max_{r_{22}}$ для яких виконуються наступні нерівності [13]:

$$\min_{r_{11}} \leq |r_{11}(\theta)| \leq \max_{r_{11}} \text{ та } \min_{r_{22}} \leq |r_{22}(\theta)| \leq \max_{r_{22}}.$$

Якщо

$$(\min_{r_{11}} \leq |r_{11}(\theta)| \leq \max_{r_{11}}) \cap (\min_{r_{22}} \leq |r_{22}(\theta)| \leq \max_{r_{22}}) \in \emptyset, \quad (6)$$

то задача не має розв'язку, інакше знаходимо множину векторів a^i_1 та кути α_i та β_i , для яких виконується рівність $r_{11}(\alpha_i) = r_{22}(\beta_i)$. Тоді $a^i_1 = r_{11}(\alpha_i) = r_{22}(\beta_i)$.

Повернемо плоский геометричний об'єкт S_1 на кут α , а плоский геометричний об'єкт S_2 на кут β та позначимо повернуті плоскі геометричні об'єкти через $S_{1\alpha}$ та $S_{2\beta}$.

Тоді середній ряд щільної укладки можна представити наступним чином:

$$\bigcup_n (S_{1\alpha} + n a^i_1), \text{ де } n=0, \pm 1, \pm 2, \dots \quad (7)$$

Навколо плоских геометричних об'єктів в середньому ряду побудуємо ГВФЦР Γ_{12} плоского геометричного об'єкту $S_{1\alpha}$ з плоским

геометричним об'єктом $S_{2\beta}$. Внутрішню область годографа позначимо через Φ_{12} . Тоді зону заборони для кінців векторів p^i та g^i визначимо наступним виразом:

$$\bigcup_n (\hat{O}_{21} + n a^i), \quad (8)$$

де $n=0, \pm 1, \pm 2, \dots$

Для кожного допустимого вектора a^i ми матимемо свої допустимі значення векторів p^i та g^i . Допустимі значення для векторів p^i та g^i вибираємо із умови, що кінці векторів p^i та g^i не повинні знаходитися у зоні заборони (8) та повинні знаходитися на годографі Γ_{12} . Крім того, вони повинні лежати по різні сторони від середнього ряду, тобто по різні сторони прямої $O_1O_2\dots O_n$, що проходить через полюси плоских геометричних об'єктів середнього ряду (рис. 2). Позначимо криві, на яких можуть знаходитись кінці векторів p^i та g^i через Pu та Pd (рис. 2). Ці криві періодичні, з періодом рівним довжині вектора a^i , тобто $A_1A_2=A_2A_3=A_3A_4=A_4A_5=B_1B_2=B_2B_3=B_3B_4=B_4B_5=O_1O_2=O_2O_3=O_3O_4=|a^i|$. Тому для області допустимих значень для векторів p^i та g^i достатньо обмежитись одним періодом. Визначивши область допустимих значень для векторів a^i , p^i та g^i ми однозначно визначили множину допустимих подвійних решіток $W^i=W(a^i, a^i, g^i)$, де $i=1, 2, \dots, q$, тобто область допустимих розв'язків задачі укладки для плоских геометричних об'єктів виду S_1 та S_2 .

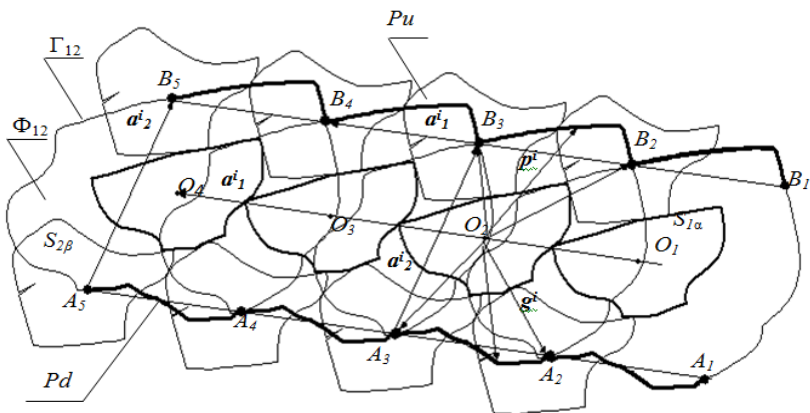


Рис. 2 – Области значень векторів p^i та g^i для задачі „Укладка Γ ”

Аналітичне представлення функції цілі для задачі проектування найцільнішої решітчастої укладки. Так як математичною моделлю

задачі проектування найщільнішої решітчастої укладки (укладка В, укладка Г) є подвійна решітка, то щільність $\delta_s(W)$ решітчастої укладки можна характеризувати за допомогою співвідношення (5). Площі $|S_1|$ та $|S_2|$ геометричних об'єктів S_1 та S_2 у цьому співвідношенні постійні, тому щільність $\delta_s(W)$ решітчастої укладки буде визначатись детермінантом решітки, значення якого дорівнює площі паралелограма, сторонами якого є базові вектори решітки a_1 та a_2 . Тоді із виразу (5) очевидно, що щільність $\delta_s(W)$ решітчастої укладки буде тим вища, чим менша площа цього паралелограму (4). Тобто цвілевою функцією буде детермінант $\det W$ подвійної решітки W .

Так як ми отримали аналітичний вигляд ГВФЦР Γ_{11} та Γ_{12} , а вектори $a_1=f_1(r_{11}(\theta))$ та $a_2=f_2(r_{12}(\theta))$, то знайдемо аналітичні вирази для векторів решітки a_1 та a_2 . Нехай ГВФЦР Γ_{11} та Γ_{12} мають наступний аналітичний вигляд:

$$\Gamma_{11}: \begin{cases} xg^{11} = (Xg_{i+1}^{11} - Xg_i^{11})t_i - Xg_i^{11} \\ yg^{11} = (Yg_{i+1}^{11} - Yg_i^{11})t_i - Yg_i^{11} \end{cases}, \quad (9)$$

де $i = 1, 2, \dots, n_g^{11}$ та $t_i \in [0, 1]$,

$$\Gamma_{12}: \begin{cases} xg^{12} = (Xg_{i+1}^{12} - Xg_i^{12})\tau_i - Xg_i^{12} \\ yg^{12} = (Yg_{i+1}^{12} - Yg_i^{12})\tau_i - Yg_i^{12} \end{cases}, \quad (10)$$

де $i = 1, 2, \dots, n_g^{12}$ та $\tau_i \in [0, 1]$.

Тоді вектори a_1 , p , g можна представити наступним чином:

$$a_1: \begin{cases} x_{a1} = (Xg_{i+1}^{11} - Xg_i^{11})t_i - Xg_i^{11} \\ y_{a1} = (Yg_{i+1}^{11} - Yg_i^{11})t_i - Yg_i^{11} \end{cases}, \quad (11)$$

де $i = 1, 2, \dots, n_g^{11}$ та $t_i \in [0, 1]$,

$$p: \begin{cases} x_p = (Xg_{j+1}^{12} - Xg_j^{12})\tau_j - Xg_j^{12} \\ y_p = (Yg_{j+1}^{12} - Yg_j^{12})\tau_j - Yg_j^{12} \end{cases}, \quad (12)$$

де $j = 1, 2, \dots, n_g^{12}$ та $\tau_j \in [0, 1]$,

$$g: \begin{cases} x_g = (Xg_{k+1}^{12} - Xg_k^{12})\tau_k - Xg_k^{12} \\ y_g = (Yg_{k+1}^{12} - Yg_k^{12})\tau_k - Yg_k^{12} \end{cases}, \quad (13)$$

де $k = 1, 2, \dots, n_g^{12}$ та $\tau_k \in [0, 1]$.

Звідси функція цілі матиме наступний вигляд:

$$\begin{aligned} \det W = [a_1 \times a_2] &= \begin{vmatrix} a_{1x} & a_{1y} \\ a_{2x} & a_{2y} \end{vmatrix} = \begin{vmatrix} a_{1x} & a_{1y} \\ x_p - x_g & y_p - y_g \end{vmatrix} = \\ &= |a_{1x}(y_p - y_g) - (x_p - x_g)a_{1y}| = |a_{1x}y_p - a_{1x}y_g - x_p a_{1y} + x_g a_{1y}| = \\ &= F(t_i, \tau_j, \tau_k, i, j, k) = \\ &= |A_{ij}t_i\tau_j + B_{ik}t_i\tau_k + (C_{ij} + D_{ik})t_i + E_{ij}\tau_j + F_{ik}\tau_k + L_{ijk}| \end{aligned} \quad (14),$$

де

$$\begin{aligned} A_{ij} &= (Xg_{i+1}^{11} - Xg_i^{11})(Yg_{j+1}^{12} - Yg_j^{12}) - (Yg_{i+1}^{11} - Yg_i^{11})(Xg_{j+1}^{12} - Xg_j^{12}), \\ B_{ik} &= (Yg_{i+1}^{11} - Yg_i^{11})(Xg_{k+1}^{12} - Xg_k^{12}) - (Xg_{i+1}^{11} - Xg_i^{11})(Yg_{k+1}^{12} - Yg_k^{12}), \\ C_{ij} &= Xg_j^{12}(Yg_{i+1}^{11} - Yg_i^{11}) - Yg_j^{12}(Xg_{i+1}^{11} - Xg_i^{11}), \\ D_{ik} &= Yg_k^{12}(Xg_{i+1}^{11} - Xg_i^{11}) - Xg_k^{12}(Yg_{i+1}^{11} - Yg_i^{11}), \\ E_{ij} &= Yg_i^{11}(Xg_{j+1}^{12} - Xg_j^{12}) - Xg_i^{11}(Yg_{j+1}^{12} - Yg_j^{12}), \\ F_{ij} &= Xg_i^{11}(Yg_{k+1}^{12} - Yg_k^{12}) - Yg_i^{11}(Xg_{k+1}^{12} - Xg_k^{12}), \\ L_{ijk} &= Xg_i^{11}Yg_j^{12} + Yg_i^{11}Xg_k^{12} - Xg_i^{11}Yg_k^{12} - Yg_i^{11}Xg_j^{12}. \end{aligned} \quad (15)$$

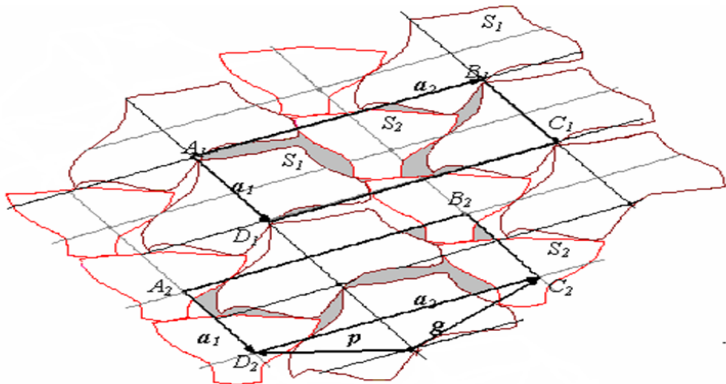


Рис. 3 – Щільна укладка для двох видів деталей

Як видно з рівнянь (14 -15) функція цілі $\det W$ є лінійна функція від трьох змінних $t_i \in [0,1]$, $\tau_j \in [0,1]$, $\tau_k \in [0,1]$ та трьох дискретних параметрів i, j, k ($i = 1, 2, \dots, n_g^{11}$, $j = 1, 2, \dots, n_g^{12}$, $k = 1, 2, \dots, n_g^{12}$) та має вигляд:

$$\det W = F(t_i, \tau_j, \tau_k, i, j, k) = \left| A_{ij} t_i \tau_j + B_{ik} t_i \tau_k + (C_{ij} + D_{ik}) t_i + E_{ij} \tau_j + F_{ik} \tau_k + L_{ijk} \right|. \quad (16)$$

Тоді локальний екстремум цієї функції може бути тільки на граничних значеннях змінних, тобто для

$$t_i = \begin{cases} 0 \\ 1 \end{cases}; \quad \tau_j = \begin{cases} 0 \\ 1 \end{cases}; \quad \tau_k = \begin{cases} 0 \\ 1 \end{cases}. \quad (17)$$

Для побудова схеми розкрою в прямокутній області заданих розмірів за параметрами спроектованої щільної укладки використаємо алгоритми, що запропоновані в роботі.[13]. Для автоматичного перенесення спроектованої схеми розкрою для прямокутної області заданих розмірів на матеріал зі складною конфігурацією зовнішнього контуру були вирішена наступна задача – визначення взаємного розміщення будь-якого плоского геометричного об'єкту із схеми розрою для прямокутної області заданих розмірів та матеріалу зі складною конфігурацією зовнішнього контуру, а саме: плоский геометричний об'єкт знаходиться повністю на матеріалі зі складною конфігурацією зовнішнього контуру; плоский геометричний об'єкт знаходиться зовні матеріалі зі складною конфігурацією зовнішнього контуру; плоский геометричний об'єкт перетинає матеріалі зі складною конфігурацією зовнішнього контуру. Для вирішення цієї задачі використаємо алгоритми, що запропоновані в роботі [14]. Для інтерактивного коригування схеми розкрою, що спроектована в атоматичному режимі, для матеріалу зі складною конфігурацією зовнішнього контуру необхідно: аналітичне представлення умов взаємного не перетину плоского геометричних об'єкту із границею зовнішнього контуру матеріалу зі складною конфігурацією зовнішнього контуру; ідентифікація будь-якого плоского геометричного об'єкту у схемі розкрою для вилучення його із схеми розкрою; контроль не перетину активного плоского геометричного об'єкту із уже розміщеними плоскими геометричними об'єктами.

Для вирішення цих задач, скористаємося алгоритмами, що запропоновані в роботі [14].

На основі структурних компонент поставленої задачі був запропонований алгоритм автоматизованого проектування раціональних

схем розкрою матеріалів зі складною конфігурацією зовнішнього контуру на плоскі геометричні об'єкти, що включає наступні пункти:

- апроксимація зовнішніх контурів матеріалу та плоских геометричних об'єктів, на які необхідно розкрити цей матеріал (рис. 4.а);
- визначення габаритних розмірів $DlSh$ та HSh матеріалу зі складною конфігурацією зовнішнього контуру (рис. 4.а);
- автоматичне проєктування раціональної схеми розкрою для матеріалу прямокутної форми з габаритними розмірами Dl та H , де $Dl > DlSh$ та $H > HSh$ (рис. 4.б);
- автоматичне перенесення зовнішнього контуру матеріалу на спроектовану схему розкрою (рис. 4.б);
- автоматичне вилучення із спроектованої схеми розкрою всіх деталей, які знаходяться зовні або перетинають зовнішній контур натурального матеріалу (рис. 4.в);
- інтерактивне коригування отриманої схеми розкрою для матеріалу зі складною конфігурацією зовнішнього контуру (рис. 4.г).

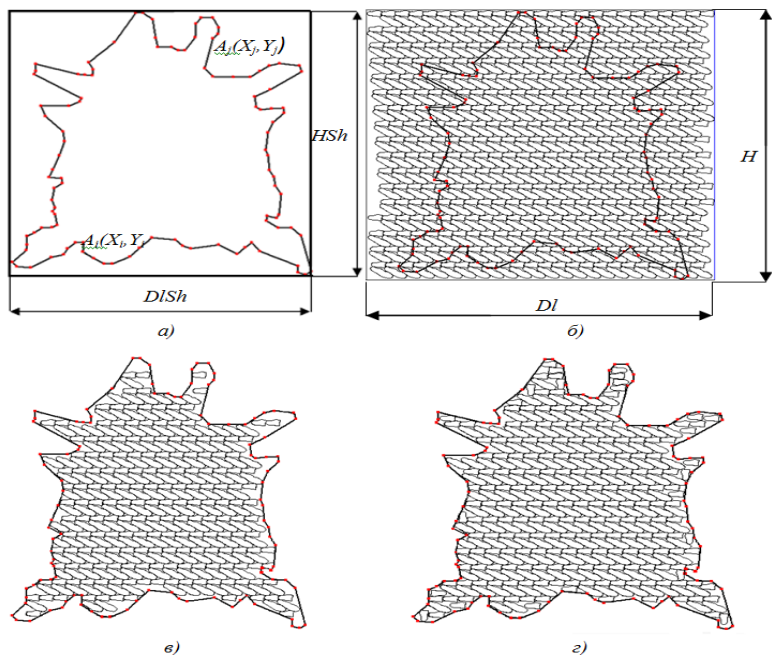


Рис. 4 – Етапи автоматизованого проєктування раціональних схем розкрою матеріалів зі складною конфігурацією зовнішніх контурів на плоскі геометричні об'єкти

ВИСНОВКИ

Запропоновані математична модель та алгоритми були реалізовані в програмне забезпечення автоматизованого проектування раціональних схем розкрою натуральних матеріалів на деталей низу взуття. Це програмне забезпечення підвищує показники економії матеріалу і продуктивність праці.

ЛІТЕРАТУРНІ ДЖЕРЕЛА

- [1] Guo P.N. Floor-planning using a tree representation/ P.N. Guo, T. Takahashi, C.K. Cheng, T. Yoshimura // IEEE Trans. on Computer Aided Design of Integrated Circuits and Systems, 2001. – 281 p.
- [2] Sakanushi K. The quarter-state-sequence floorplan representation / K. Sakanushi, Y/ Kajitani, D.P. Mehta// IEEE Trans. on Computer Aided Design of Integrated Circuits and Systems, 2003. – 376 p.
- [3] Бухвалова В.В. Задача прямоугольного раскроя: метод зон и другие алгоритмы СПб.: СПбГУ, 2001. – 96 с.
- [4] Колиско О.З. Модифікація генетичного алгоритму для генерації секцій розкрийних схем // Вісник КНУТД, 2009. – №1. – С.14 – 17.
- [5] Deb K. An efficient constraint handling method for genetic algorithms // Computer Methods in Applied Mechanics and Engineering, 2000. – Vol. 186, № 2–4. – P. 311–338.
- [6] Верхотуров М.А. Задача нерегулярного раскроя плоских геометрических бъектов: моделирование и расчет рационального раскроя // Информационные технологии, 2000. – №5. – С.37 – 42.
- [7] Скобцов Ю.А. К вопросу о применении метаэвристик в решении задач рационального раскроя и упаковки / Ю.А. Скобцов, В.Н. Балабанов // Вестник Хмельницкого национального университета, 2008. – Т. 1, № 4. – С. 205 – 217.
- [8] Валеева А.Ф. Применение конструктивной метаэвристики "муравьиная колония" к задаче гильотинного прямоугольного раскроя / А.Ф. Валеева, А.А. Петунии, Р.И. Файзрахманов// Вестник Башкирского университета. Раздел: Математика. Уфа, 2007. – Том 12. – N3. – С. 12 – 14.
- [9] Zhang D.F. An improved heuristic recursive strategy based on genetic algorithm for the strip rectangular packing problem/ D.F. Zhang, C.D.Chen, Y.I. Lin // Acta Automatica Sinica, 2007. – №33(9). – P. 911 – 916.
- [10] Воднев В.Т. Основные математические формулы / А.Ф. Наумович, Н.Ф. Наумович. Минск: Вышэйшая школа, 1988, – 270 с.
- [11] Стоян Ю.Г. Размещение геометрических объектов. / Ю.Г. Стоян – Киев: Наукова думка, 1975. – 175 с.

[12] Стоян Ю.Г. Методы и алгоритмы размещения плоских геометрических объектов / Ю.Г. Стоян, Н.И. Гиль. – Киев: Наукова думка, 1976, – 242 с.

[13] Чупринка В.І. Автоматична побудова решітчастих схем розкрою для двох видів плоских геометричних об'єктів / В.І. Чупринка, О.В. Чебанюк, А.В. Пінчук // Вісник Східноукраїнського національного університету ім. Даля, 2008. – №8. – С. 230 – 235.

[14] Чупринка В.І. Інтерактивна побудова та коригування схем розкрою натуральних матеріалів на деталі взуття / В.І. Чупринка, М.М. Шкоденко, В.П. Коновал // Вісник КНУТД, 2010. – №1. – С. 62 – 67.

AUTOMATED PREPARATION OF RATIONAL DISCLOSURES OF MATERIALS WITH COMPLEX CONFIGURATION OF EXTERNAL CONTOURS TO FLAT GEOMETRIC OBJECTS

Chuprinka V., Grikun D., Chuprinka N.

The paper presents a mathematical model and algorithms that have been implemented in the software of computer-aided design of rational schemes of cutting materials with complex configuration of the outer contour on flat geometric objects. This software enhances material savings and productivity. To develop a mathematical model of the problem of automated design of rational schemes of cutting materials with complex configuration of the outer contour on flat geometric objects, its structural components were formalized: analytical description of the outer contour of the material with complex configuration of the outer contour; Analytical description of flat geometric objects for which rational cutting schemes should be designed on material with complex external contour configuration; parameters that uniquely determine the position of a flat geometric object on a plane; construction of dense inserts for flat geometric objects; construction of the cutting scheme in the rectangular area of the given sizes according to the parameters of the designed dense laying; transfer of the design scheme to material with complex configuration of the outer contour; interactive adjustment of the cutting scheme, designed in the automatic mode, for material with complex configuration of the outer contour. Using the considered structural components, an algorithm and software were developed for the automated design of rational schemes of cutting materials with complex configuration of the outer contour on flat geometric objects

Keywords. Rational cutting, material, outer contour, system placement, plane, interactive adjustment

НЕКОТОРЫЕ ИНФОРМАЦИОННЫЕ ПРОБЛЕМЫ КОМБИНАТОРНОЙ ЛИНЕЙНОЙ ОПТИМИЗАЦИИ

Юхименко Б.

Дискретная оптимизация одна из наиболее затребованных дисциплин, с которыми связана процедура принятия решений. Ее использование связано со многими проблемами, которые процедурно объемны. Построение математической модели, подбор действительно существенных факторов, приведение к типам существующих математических задач и дальнейшая реализация на компьютерной технике довольно трудоемкий процесс. Существующие математические методы как аналитического типа и комбинаторного характера все еще относятся к классу NP сложности. Хотя современная вычислительная техника сверхмощная и хорошо оснащена математическим обеспечением решения задач особенно большой размерности практически не реализуемая в случае требования – получить оптимальное решение.

Кроме того, комбинаторные методы и алгоритмы решения задач дискретной оптимизации связаны с большим количеством промежуточной информации, необходимой для получения оптимального или хотя бы достаточно хорошего решения. В этой связи, разработка новых, модифицирующих блоков и модулей, включаемых в схемы существующих алгоритмов, в настоящее время затребовано, актуально и имеет как научное, так и прикладное значение. Данная разработка представляет в некотором смысле систематизацию проблем метода ветвей и границ для решения задачи линейного программирования с булевыми переменными в основном, и многомерной задачи о ранце в частности. Перечислены основные моменты метода, которые можно модифицировать. Показана возможность подключить моменты приближенных алгоритмов подготовительная часть точного алгоритма. Решен числовой пример, демонстрирующий положительное влияние двух алгоритмических процедур на эффективность работы метода. Использование рекордного значения целевой функции уменьшает на половину число пересматриваемых вершин, а введение приоритетной последовательности компонент для конкретизации еще отсеивает значительное количество вершин.

Ключевые слова. Дискретная оптимизация, рекорд, последовательное построение решения, приближенные алгоритмы, конкретизация, приоритетная последовательность.

ВВЕДЕНИЕ

Принятие оптимальных решений в практической деятельности человека это многогранная и многоэтапная проблема. Доведение процесса от постановки задачи до получения результата решения и его внедрения встречается много различных препятствий и задач, которые необходимо решать. Многие вычислительные алгоритмы оптимизации содержат моменты, связанные с решением задач информационного

характера. Эффективность работы алгоритмов комбинаторного характера не исключает проблему организации исходной, промежуточной и результирующей информации.

Метод ветвей и границ как основной комбинаторный метод дискретной оптимизации так же подвергается всеобщему обсуждению по всем его составным частям и процедурам. Ему посвящено множество публикаций, разработано немало модификаций, нацеленных на уменьшение количества перебираемых вариантов, однако, он все еще остается алгоритмом NP – сложности. Сложность вычислений увеличивается за счет обслуживания информационного обеспечения работы алгоритмов. В научной литературе этот вопрос обсуждается недостаточно хорошо как в теоретическом, так и в практическом плане. В работе [1] приведен перечень необходимых информационных массивов, целый список параметров, значения которых необходимо для формирования подзадач, используемых для получения подмножеств вариантов.

Использование идеи последовательного построения решения [2] как средство разбиения множества вариантов на подмножества, на наш взгляд, подает подсказку другого подхода к формированию и работы с массивами промежуточной информации. Практически последовательное построение решения предполагает конкретизацию значений компонент вектора решений.

Запишем этот процесс формально.

Пусть имеется задача целочисленного линейного программирования (ЦЛП) в постановке

$$Z = \max(C, X)$$

при ограничениях

$$\sum_{j=1}^n A_j x_j \leq B$$

$$X = (x_1, x_2, \dots, x_n), \quad x_j \in \{0, N\}, \quad j = \overline{1, n}$$

Обозначим через σ^k частичное решение, в котором конкретизировано (определено) k компонент вектора решений X . Множество I_1^k содержит номера конкретизированных компонент соответственно. Число k предопределяет ярус дерева решений. Действительно, любой алгоритм метода ветвей и границ имеет древовидную схему получения оптимального решения. При работе

алгоритма формируется само дерево, вершинами которого соответствуют подмножества вариантов, содержащих частичное решение σ_l^k . Поскольку конкретизируемая переменная может принимать не одно целочисленное значение, то на k -том уровне (ярусе) имеется S_k вершин ($l = \overline{1, S_k}$). Для каждой вершины формируется подзадача с целью получения оценки подмножества. Пусть G_l^k подмножество вариантов, содержащее частичное решение σ_l^k , а $\xi(G_l^k)$ - оценка подмножества. Оценка получается путем решения соответствующей подзадачи. Информационно каждая подзадача отличается одна от другой лишь компонентами двух векторов – вектора условий A_{l_0} и вектора ограничений B . Индекс l_0 показывает, какая компонента вектора решений конкретизирована при переходе от $k-1$ -го яруса к k -ому.

Таким образом, на k -ом ярусе дерева решений достаточно иметь информацию о одной подзадаче, для остальных безболезненно формируется программно.

Экспериментально было показано [3], что в алгоритмах, реализующих идею последовательного построения решения, скорость сходимости значительно зависит от выбора компоненты, подлежащей конкретизации на очередном уровне дерева решений. Выгодно ввести процедуру $\Psi(j)$, предопределяющую номер компоненты подлежащей конкретизации на конкретном шаге работы алгоритма. Под воздействием процедуры $\Psi(j)$ останутся только те компоненты, которые имеют потенциал на более одного не отрицательного значения. Обозначим множество индексов таких компонент через V_l^k

$$V_l^k = \{j / j \in \overline{1, S_k} / (b_i - \sum_{l \in I_{l_0}^k} a_{il} \sigma_{l_0}^k) \geq 0 \forall i\}$$

Если $V_l^k \equiv \emptyset$, то данную ветвь можно исключить из рассмотрения.

Еще один заслуживающий внимание момент. Поскольку в процессе работы алгоритма формируется дерево решений, согласно структуре, которого содержится информация о решаемых подзадачах, то чем меньше ветвей будет порождено (особенно при первом подходе), тем

больше подмножеств вариантов исключается из рассмотрения. Изначально наличие рекордного значения целевой функции R запретить порождать ветви, представляющие неперспективные подмножества. Они будут отсеены.

ОСНОВНАЯ ЧАСТЬ

Что касается решения задач с булевыми переменными, в частности, многомерной задачи о ранце, то само формирование частичного решения значительно упрощается. На каждом ярусе дерева решений имеется только два частичных решения. Конкретизируемая переменная может принимать только два значения либо «0» либо «1». Значительно уменьшается количество хранимой информации о подзадачах. Аналогично, как и в общем случае меняется только два вектора, компоненты вектора конкретизируемой переменной и вектора ограничений.

В алгоритмическом смысле увеличить количество отсеиваемых неперспективных подмножеств для алгоритма решений задач о ранце можно за счет следующих моментов:

- получение более точного рекордного значения R изначально до формирования дерева решений;
- установка приоритетной очереди номеров компонент вектора решений претендентов на значение «1»;
- формирование множества индексов компонент («хороших компонент»), имеющих потенциал на значение «1».
- проверка на наличие потенциала среди хороших компонент улучшить значение целевой функции;
- определение функции предпочтения при возврате к вершине дерева решений.

Алгоритмическая реализация некоторых из этих моментов основывается на идеях построения приближенных или вероятностно-приближенных алгоритмов. Разные приближенные алгоритмы предполагают различные подходы формирования приоритетной очереди компонент, претендентов на значение «1». Процедура $\Psi(j)$ может пошагово определять номер компоненты в зависимости от сложившейся ситуации, а может формировать всю приоритетную последовательность номеров компонент. В принципе эта идея не является новой. Алгоритм Балаша [4], представляющий алгоритм метода ветвей и границ с односторонним ветвлением, практически состоит из процедур поиска компонент, которым присваивается значение «1». Негативный момент алгоритма то, что признак оптимальности не строгий и может привести к полному перебору вариантов. Он тоже основан на поиске компонент, которые могут иметь значение «1». Сама процедура проверки на

оптимальность полученного решения более длительная, чем получение самого варианта. Он может считаться приближенным решением, хотя зачастую, он является оптимальным.

Современные приближенные алгоритмы предполагают формирование всей приоритетной очереди компонент-претендентов на значение «1». Обозначим эту последовательность номеров компонент через $P = (j_1, j_2, \dots, j_n)$. Место компоненты с номером j_i в разных алгоритмах определяется по-разному согласно своей методике алгоритма. Рассмотрим некоторые из них.

Пусть имеется математическая модель многомерной задачи о ранце

$$Z = \max \sum_{j=1}^n c_j x_j$$

при ограничениях

$$\sum_{j=1}^n a_{ij} x_j \leq b_i, \quad i = \overline{1, m};$$

$$x_j \in \{0, 1\}, \quad j = \overline{1, n}$$

- где $c_j, a_{ij}, b_i \forall_{ij}$ - натуральные числа.

Жадный алгоритм [5] предлагает упорядочить номера компонент j_i согласно ранжированной в порядке не возрастания компонент вектора стоимостей. Если $c_{j_1} \geq c_{j_2} \geq \dots \geq c_{j_n}$, то приоритетная очередь будет $P = (j_1, j_2, \dots, j_n)$.

С целью получения оптимального решения нецелочисленной одномерной задачи о ранце Дацинг [1] предлагал модифицированную

последовательность ранжировки. Если $\frac{c_{j_1}}{a_{j_1}} \geq \frac{c_{j_2}}{a_{j_2}} \geq \dots \geq \frac{c_{j_n}}{a_{j_n}}$, то

(j_1, j_2, \dots, j_n) составляет приоритетную очередь компонент вектора решений.

В некоторых комбинированных приближенных алгоритмах предполагается учитывать величину невязки в системе ограничений при подборе приоритетной компоненты. Каждая компонента оценивается величиной r_j , которая определяется так

$$r_j = c_j \sum_{i=1}^m (b_i - a_{ij}), \quad j = \overline{1, n}.$$

Величина r_j ранжируется в порядке не возрастания и определяет приоритетную очередь.

Автор проводила такой эксперимент [3]. Увеличение эффективности работы алгоритма по сравнению с обычным жадным алгоритмом практически не наблюдается.

Генетический алгоритм [7] какойто степени учитывает «наследственность». Если получено определенное количество вариантов решения, то приоритетность компонент оценивается частотой получения значений «1» в этих вариантах. На наш взгляд, можно использовать результаты той работы, которая продлевается, получая оценки подмножеств и решая для этого подзадачи. Решается m одномерных нецелочисленных задач о ранце, используя метод Дацига [6]. Имеем

$$Z_i = \sum_{j=1}^n c_j x_j$$

при ограничениях

$$\sum_{j=1}^n a_{ij} x_j \leq b_i \quad x_j \in [0, 1], \quad j = \overline{1, n}.$$

Имеем m оптимальных решений $X_i^* = (x_1^{*i}, x_2^{*i}, \dots, x_n^{*i})$ ($i = \overline{1, m}$).

Оценка компоненты вектора решений r_j определяется числом полученных значений соответствующей компоненты в векторах X_i^* .

$$r_j = \sum_{i=1}^m x_j^{*i}, \quad j = \overline{1, n}.$$

Величина r_j предопределяет приоритетность компонент вектора решений. В последнее время разрабатываются алгоритмы, основанные на поведении живых существ или каких-то природных явлений. Хорошо известный алгоритм муравьиной колонии [8,9] используется при решении многих оптимизационных задач. Использование такого типа алгоритмов позволяет получить достаточно хорошее приближенное, а иногда и оптимальное решение задач большой размерности [10].

Особи муравьиной колонии могут передавать информацию, действовать самостоятельно, достигая общей цели, менять направление движения в зависимости от создавшейся ситуации (обратная связь). Все это удастся формализовать и создавать эффективные вероятностно-приближенные алгоритмы. Приоритетность компонент вектора решений получить значение «1» определяется величиной вероятности p_j , количественное значение которой зависит от количества информации (феромонов) при массовом движении в одном направлении, изменение направления (испарение феромонов), а так же специфических особенностей решаемой задачи. Имеем

$$p_j = f^1(\alpha, \beta) \cdot f^2(c_j, b_j),$$

где α - предопределяет количество феромонов, как информацию о повторении некоторых алгоритмических действий, β - испарение феромонов. Величина β рассматривается как параметр, от которого зависит эффективность работы алгоритма. Функция $f^2()$ формируется индивидуально создателям алгоритма и зависит от его мнения [13]. Она каким то образом увеличивает количественно значение величины p_j и дает некое предпочтение компоненте j при получении значения «1». Процедурно сам процесс присвоения положительного значения осуществляется случайным образом. Определяется значение η , как выпад равномерно распределенной случайной величины из интервала (0,1). Попадание η в определенный интервал пропорционально вероятности p_j определяет претендента на значение «1». Естественно, необходимо учесть такую возможность при наличии ограничивающих условий.

Упростить процедуру поиска приоритетной компоненты (назовем ее хорошей компонентой) можно тем, что изначально выделяются компоненты, имеющие потенциал положительного значения [12]. Пусть на каком то k -ом уровне рассматривается компонента l_0 , тогда вершине соответствует множество $I_{l_0}^k$, определяющее номера конкретизированных компонент. Множество V_l^k предопределяет компоненты, которым можно присвоить значение «1». Они определяются так

$$V_l^k = \{j / j \in I_{l_0}^k / (b_i - \sum_{l \in I_{l_0}^k} a_{il}) > 0 \forall_i\}$$

В случае, если $V_l^k = \emptyset$, то продолжение по соответствующей ветви дерева решений нецелесообразно. Если $V_l^k \neq \emptyset$ т.е. существуют компоненты, которым можно присвоить значение «1», но в сумме не могут увеличить значение целевой функции чтобы превысить имеющийся рекорд, то продолжение рассмотрения подмножеств вариантов не имеет смысла. Формально можно записать так

$$\left(\sum_{j \in I^k} c_j x_j + \sum_{j \in V_l^k} c_j \right) \leq R$$

где I^k индексы конкретизированных компонент.

Если не выполняется признак оптимальности для полученного варианта решений, имеются висающие вершины с лучшими оценками, чем рекорд R , то вопрос – какой вершине отдать предпочтение. Возможны такие называемые функции предпочтения выбора подмножества вариантов:

- с максимальной оценкой;
- с частичным решением, где k максимально;
- случайный выбор пропорционально величинам оценок;
- случайный выбор среди трех (пяти, ...) вершин с максимальной оценкой;

Или еще какие-то способы по усмотрению разработчика алгоритма.

Какая функция предпочтения окажется более эффективной, уменьшить количество пересматриваемых подмножеств и тем самым уменьшить количество промежуточной информации, можно определить только экспериментально.

Использование первой функции предпочтения может привести к длительному процессу ветвления, поскольку на верхних ярусах дерева будут подмножества, содержащие частичное решение с маленьким числом конкретизированных переменных (k – минимальные числа). С этих позиций вторая функция предпочтения как бы эффективнее. Оценки подмножеств близки по значению к рекорду. Придется доводить до самого нижнего яруса. Оценить просто так невозможно. Необходимо экспериментировать.

Ниже приведен пример решения двухмерной задачи о ранце, с целью демонстрации влияния изначального наличия рекордного значения R . Приоритетная очередь конкретизации компонент вектора решений соответствует непосредственной их нумерации. Чтобы подчеркнуть влияние приоритетной очереди, определяемой согласно приближенному комбинированному алгоритму, повторно решается тот же пример.

ПРИМЕР

Приведем результаты решения многомерной задачи о ранце, используя выше сказанные соображения. Пусть дано

$$Z = \max(5x_1 + 3x_2 + 4x_3 + 5x_4 + 2x_5 + 4x_6 + 5x_7)$$

$$\begin{cases} 4x_1 + 3x_2 + 2x_3 + 4x_4 + 2x_5 + 3x_6 + 5x_7 \leq 10 \\ 2x_1 + 4x_2 + 2x_3 + x_4 + 4x_5 + 4x_6 + 3x_7 \leq 8 \\ x_j \in \{0,1\}, j = \overline{1,7} \end{cases}$$

Задача решалась с целью показать, как влияет наличие рекордного значения целевой функции R на процесс отсеивания неперспективных вариантов. И второй случай, влияние наличия приоритетной последовательности компонент вектора решений тоже на количество рассматриваемых вариантов.

Задача решается обычным методом ветвей и границ. Оценивание подмножеств вариантов ведется с использованием подхода расширения множества вариантов. Другими словами, доведения задачи до просто решаемой модели, содержащей исходную задачу. Решается две (два ограничения) нецелочисленных одномерных задачи о ранце. Следуя методу Дацинга, оптимальное решение достигается просто. Соотношение соответствующих коэффициентов целевой функции и ограничений (c_j/a_j) предопределяет порядок присвоения значений «1» при возможности или части единицы компонентам векторов решений для одномерных задач. Первая задача имеет вид

$$Z_1 = \max(5x_1 + 3x_2 + 4x_3 + 5x_4 + 2x_5 + 4x_6 + 5x_7)$$

при ограничениях

$$4x_1 + 3x_2 + 2x_3 + 4x_4 + 2x_5 + 3x_6 + 5x_7 \leq 10$$

$$x_j \in [0,1], j = \overline{1,7}.$$

$$\text{Определяется } \lambda_j = \frac{c_j}{a_j} ::= \frac{5}{4}, \frac{3}{3}, \frac{4}{2}, \frac{5}{4}, \frac{2}{2}, \frac{4}{3}, \frac{5}{5};$$

Упорядочив a_j в порядке не возрастания, имеем последовательность конкретизации переменных

$$P_1 = \{3, 6, 1, 4, 7, 2, 5\}.$$

Процедура получения решения следующая

$$x_3 = 1, \quad 10 - 2 = 8;$$

$$x_6 = 1, \quad 8 - 3 = 5;$$

$$x_1 = 1, \quad 5 - 4 = 1;$$

$$x_4 = 1/4.$$

Само решение $X^1 = (1, 0, 1, 1/4, 0, 1, 0)$;

Значение целевой функции $Z_1^* = 4 + 4 + 5 + 5 \cdot 1/4 = 14,25$.

Аналогично решается вторая одномерная задача в постановке

$$Z_2 = \max(5x_1 + 3x_2 + 4x_3 + 5x_4 + 2x_5 + 4x_6 + 5x_7)$$

при ограничениях

$$2x_1 + 4x_2 + 2x_3 + x_4 + 4x_5 + 4x_6 + 3x_7 \leq 8$$

$$x_j \in [0,1], \quad j = \overline{1,7}.$$

Последовательность присвоения положительных значений

$$P_2 = \{4, 1, 3, 7, 6, 2, 5\}.$$

Вектор решения $X^2 = (1, 0, 1, 1, 0, 0, 1)$;

Значение целевой функции

$$Z_2^* = 5 + 4 + 5 + 5 = 19,0.$$

Оценка множества вариантов решения

$$\xi = \min(Z_1^*, Z_2^*) = (14,25; 19,0) = 14,25.$$

Разбиение множества вариантов на подмножества-ветвление организовано последовательной конкретизацией компонент вектора решений согласно их непосредственной нумерацией.

Рекордное значение целевой функции R определяется согласно приоритетной очереди компонент вектора решений, используя

приближенный комбинированный алгоритм. Очередь на присвоение значения «1» следующая

$$P = \{1, 6, 3, 4, 7, 2, 5\}.$$

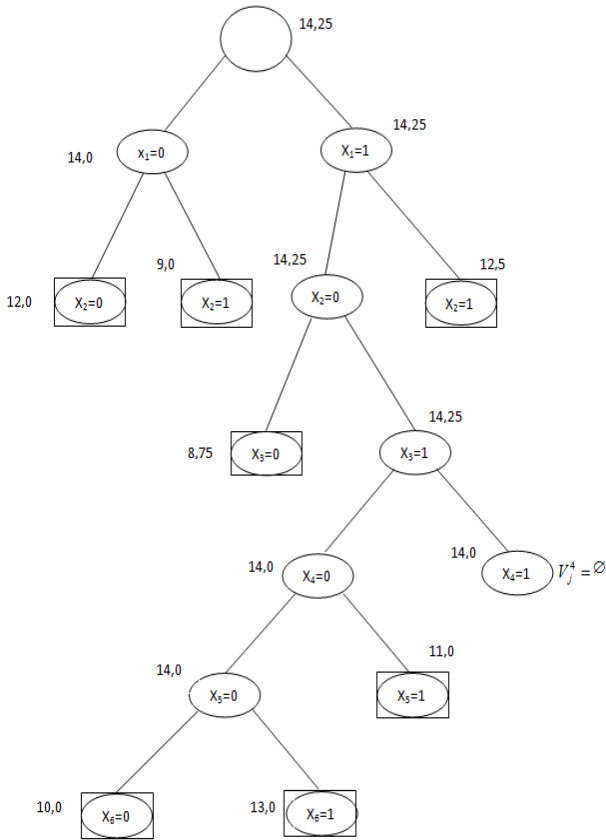


Рис.1 – Дерево решений 1

Компонентам x_1, x_6, x_3 присваивается значение «1». Больше не позволяет система ограничений. Значение целевой функции при векторе решений $X = (1, 0, 1, 0, 0, 1, 0)$ нетрудно определить $Z = 5 + 4 + 4 = 13$ оно воспринимается как изначальное рекордное значение R . Дерево решений, демонстрирующее влияние наличия изначального рекордного значения целевой функции приведено на рис.1. Использовано последовательное построение решение согласно нумерации компонент вектора решений. Рассмотренный пример дерева решений содержит 14

вершин. Реально рассматривалось только 6. Наличие рекордного значения целевой функции ($R=13,0$) позволило исключить из рассмотрения половину бесперспективных подмножеств. На рис.1 вершины дерева, подлежащие отсеиванию, отмечены. На рис.2. представлена работа алгоритма, содержащая больше модифицированных моментов.

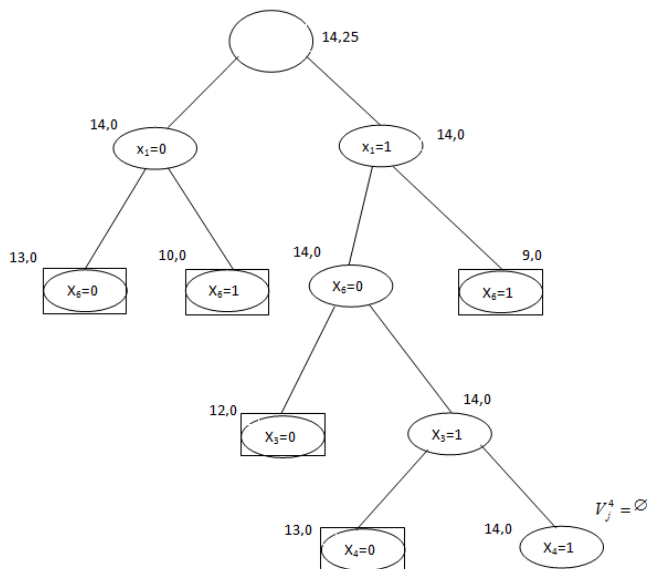


Рис. 2 – Дерево решений 2

Используется та же приоритетная последовательность компонент вектора решений при конкретизации значений, т.е. построение частичных решений ведется согласно этой очереди, а именно $P = (1, 6, 3, 4, 7, 2, 5)$. приоритетная очередь та же самая, полученная приближенным комбинированным алгоритмом, что и при определении рекордного значения. Позитивное влияние такого рода конкретизации компонент, очевидно. Количество отсеянных вершин дерева решений значительно увеличилось. Если учесть еще и момент проверки возможности определения значения «1» компоненте согласно очереди, то дерево решения имеет, лишь одну ветвь, определяющую оптимальное решение $X = (1, 0, 1, 0, 0, 1, 0)$. В первом случае, для достижения этого решения перебиралось значительно больше подмножеств, которые, в конечном счете, оказались бесперспективными. Размерность рассматриваемого примера не позволила продемонстрировать возможности функций предпочтения при поиске ветви с лучшей

оценкой, чем имеющийся рекорд. Очевидно, что рекордное значение, определенное изначально при построении дерева решений, изменилось и стало равной 14,0. В принципе ветвь с оценкой 14,0 и содержащей частичное решение $\sigma^1 \equiv (0)$ можно было и не продолжать, поскольку чем ниже по дереву решений передвигаемся, тем меньше вероятность хотя бы сохранить значение оценки.

ЗАКЛЮЧЕНИЕ

Комбинаторные методы, в том числе и метод ветвей и границ, являются открытыми методами. В них можно включать различные модифицирующие блоки, увеличивающие эффективность их работы. Однако следует отметить, что точные комбинаторные алгоритмы все еще не являются быстродействующими. Они с трудом используются для решения задач большой размерности. Кроме того, комбинаторные алгоритмы имеют такую организацию работы вычислительных процедур, что необходимо хранить большое количество информации. Это тоже усложняет работу при поиске оптимального варианта. Эффективное использование алгоритмов метода ветвей и границ в практической деятельности требует либо кардинального изменения вычислительных процедур, либо такую их организацию, чтобы создать возможность получить если не оптимальное, то хотя бы приемлемое решение.

Данная работа посвящена обобщению некоторых модифицирующих моментов, увеличивающих количество отсеиваемых подмножеств вариантов. Это делается, используя различного типа разработок при создании приближенных, а так же вероятностно-приближенных. Случайный поиск компоненты подлежащей конкретизации на очередном шаге работы алгоритма сохраняет детерминированность выполнения остальных процедур, оставляя возможность получить оптимальное решение. Естественно, если была такая установка получения решения. Приведенный пример демонстрирует влияние на эффективность работы алгоритма метода ветвей и границ с использованием рекордного значения. R можно получить оперативно, используя идею любого приближенного алгоритма. Второй процедурный прием решения приведенной задачи обращает внимание на то, насколько важен прием приоритетного упорядочения компонент вектора решений. Включается в работу по совершенствованию алгоритма и момент определения множества хороших компонент. Эффективность использования функций предпочтения при возврате на вершину с оценкой больше рекорда проверить не удалось из-за маленькой размерности примера.

На одном примере небольшой размерности выводы делать нет обоснования. Однако, это так же моменты на что действительно следует обратить внимание. В дальнейших разработках следовало бы проверить предлагаемое на большом количестве примеров. Необходимо провести числовой компьютерный эксперимент. На основе результатов эксперимента создать вычислительный алгоритм решения задач большой размерности, т.е. довести до практического использования.

СПИСОК ЛИТЕРАТУРЫ

- [1] Сигал И.Х., Иванова А.П. Введение в прикладное дискретное программирование. – М.: Физматгиз, 2007. – С. 304.
- [2] Михалевич В.С. Последовательные алгоритмы оптимизации и их применение. I, II // Кибернетика, 1965. – № 1,2.
- [3] Юхименко Б.И., Козина Ю.Ю. Сравнительная характеристика алгоритмов метода ветвей и границ для решения задач целочисленного линейного программирования. Труды Одесского политехнического университета, 2005. – Вып.2. – С. 199 – 204.
- [4] Balas E. An additive algorithm for solving linear programs with zero-one variables // Operat.Res, 1965. – V.13. – №4. – P. 517 – 546.
- [5] Дюбин Г.Н., Корбут А.А. Жадные алгоритмы для задач о ранце: поведение в среднем // Сибирский журнал индустриальной математики, – 1999. – Т.2. – №2(4). – С. 68 – 93.
- [6] Юхименко Б.И., Волкова В.П. Приближенные алгоритмы решения задачи о многомерном ранце // Дослідження в математиці і механіці, 2017. – Т.22. – Вип.2 (30). – С. 104 – 115.
- [7] Гладков Л.А. Куряйчик В.М. Генетические алгоритмы. – М.: ФИЗМАТЛИТ, 2006. – 230 с.
- [8] Штовба С.Д. Муравьиные алгоритмы: Теория и приложения // Программирование, 2005. – №4. – С.3 –18.
- [9] Леванова Т.В., Лореш М.А. Алгоритмы муравьиной колонии и имитации отжига для задачи о р-медиане. // Автоматика и телемеханика, 2004. – №3. – С. 80 – 88.
- [10] Сигал И.Х. Оценки параметров алгоритмов ветвей и границ для задач дискретной оптимизации большой размерности // Известия РАН. Теория и система управления, 2005. – №4. – С. 96 – 101.
- [11] Кочетов Ю.А. Вероятностные методы локального поиска для задач дискретной оптимизации. // Дискретная математика и приложения, 2001. – С. 84 –117.
- [12] Юхименко Б.И. Модификации метода ветвей и границ для решений задач целочисленного линейного программирования и их эффективность // Інформатика та математичні методи в моделюванні, 2015. – Т.5. – №.1. – С. 84 – 91.

[13] Юхименко Б.И., Ткаленко О.Ю. Алгоритм муравьиной колонии для многомерной задачи о ранце // Математичні методи обробки даних, 2019. – №2. – Т.21. – С. 3 – 12.

SOME INFORMATION PROBLEMS OF COMBINATORY LINEAR OPTIMIZATION

Yukhimenko B.

Discrete optimization is one of the most requested disciplines with which decision-making procedure is associated. Its use is associated with many problems that are procedurally voluminous. The construction of a mathematical model, the selection of really significant factors, the conversion to types of existing mathematical problems and the further computer realization is a rather laborious process. Existing mathematical methods of both analytical type and combinatorial type still belong to the complexity class NP. Although modern computing techniques is a heavy-duty and well equipped with mathematical software for solving problems of a particularly large dimension it is practically impossible to realize in the case of the requirement to get the optimal solution. In addition, combinatorial methods and algorithms for solving discrete optimization problems are associated with a large amount of intermediate information that necessary to obtain an optimal or at least sufficiently good solution. In this regard, the development of new, modifying blocks and modules to be included in the schemes, existing algorithms is currently requests, relevant and has both scientific and applied value.

This elaboration represents, in a sense, a systematization of the problems of the branch and bound method for solving the linear programming problem with Boolean variables in the main, and the multidimensional knapsack problem in particular. The main points of the method that can be modified are listed. The possibility to connect the moments of the approximate algorithms as the preparatory part of the exact algorithm is shown. A numerical example is solved that demonstrates the positive effect of two algorithmic procedures on the efficiency of the method. Using the record value of the objective function reduces by half the number of vertices being reviewed, and the introduction of a priority sequence of components for concretization still eliminates a significant number of vertices.

Keywords. Discrete optimization, the record, sequential construction of solutions, approximates algorithms, concretization, priority sequence.

DIMENSIONALITY REDUCTION AND EXTREMAL PARAMETER GROUPING FOR WEIGHTED GRAPH PROBLEMS

Vostrov G., Khrinenko A.

This work presents graph cut approach to dimensionality reduction and shows connections between the problem of extremal parameter grouping and the problem of maximization of positive definite quadratic form on the set of vertices of k -dimensional hypercube of side 1 by modulo. It is shown that the last problem is reduced to the minimum negative cut search for a weighted graph. A heuristic algorithm for solving this problem is given.

Keywords. Datasets, dimensionality reduction, graphs.

INTRODUCTION

In the process of studying objects and events, it becomes necessary to analyze and interpret a priori information obtained through an experimental research. The main form of presenting such information is a matrix, where columns correspond to various parameters, properties, etc. Matrix rows correspond to individual objects, which are described by one set of parameters. Advance in science has led to the development of new techniques for data analysis that are generating enormous amounts of data, which have implied growth of many databases and also aggravated visual representation of this data in a graph form, that could help to understand underlying nature and structure of the data. Since a vertex-edge diagram is the most popular method for graph visualization wide range of methods have been developed to lay out a vertex-edge diagrams [1]. Layout selection influences individual's understanding of the graph, thus it is important to find satisfactory method that can effectively depict the structure of the graph. Since real-world problems imply large-scale datasets it is not practical to compute different layouts and compare results for further usage. For a graph with million vertices it can take hours or days to calculate single representation and since there is no consensus on which layout is more preferable in general each dataset needs to be presented in its own way. Figure 1 shows that a graph in different layouts provides completely different understanding of dataset structure. One of the central problems in analysis of such information is selection of the most significant parameters in a given set or construction of generalized factors that sufficiently reflect the informational properties of the original description and allow a clear meaningful interpretation and representation in graph form. Thus, dimensionality reduction problem occurs.

In order to solve this problem, an alternative low-dimensional representation of given data should be obtained. There is a wide range of techniques for previously stated problem. Principal component analysis (PCA) and metric multidimensional scaling (MDS) represent most popular methods [2].

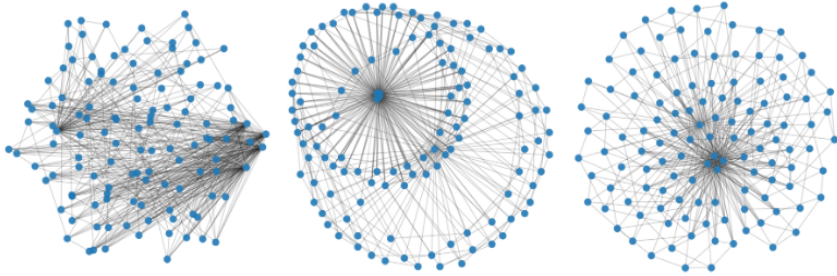


Fig.1 – Graph layouts

First method maximizes data variance captured by the low-dimensional projection, and has many extensions. First example is kernel PCA, which apply PCA on a feature space instead of initial data space. Second, probabilistic extension, that address a few weaknesses of PCA: partitioning of data into r -dimensional projected subspace (signal) and $(d-r)$ -dimensional subspace (noise), where d is initial dimensionality, and lack of explicit generative model. Third, robust extension that is driven by fact that a small number of highly corrupted data inputs can drastically influence standard PCA. The latter method also tries to preserve pairwise distances as closely as possible in the least squares sense. One of advantages of this method is that it generally requires only pairwise dissimilarities and not data itself. However, projections of data onto linear subspaces could be insufficient in many cases, when no linear projection is able to provide a satisfying representation. Therefore, this work introduces a graph cut approach to perform supervised dimensionality reduction.

FORMULATION OF PARAMETER GROUPING PROBLEM

Parameter and factor search is performed in order to simplify an a priori description and to study influence of some parameters on others. One of approaches to solve this problem is the approach based on the selection of closely correlated parameters in the original description of groups. A formal formulation of the problem of parameter grouping is given in [2], which is reduced to maximization of one of the two functionals.

$$J_1 = \sum_{j=1}^m \sum_{x_i \in A_j} (x_i, f_j)^2, \quad J_2 = \sum_{j=1}^m \sum_{x_i \in A_j} |(x_i, f_j)|, \quad (1)$$

where x_i, \dots, x_k – random variables (parameters), A_1, \dots, A_m – a partition

of the set $\{x_i, \dots, x_k\}$ on a given number of m groups, f_1, \dots, f_m – random functions (essential factors) defined on groups A_1, \dots, A_m , (x_i, f_i) – coefficient of correlation or covariance of the vectors x_i and f_i for parameter partition on groups as well as for random function selection f_1, \dots, f_m .

As it is shown in the [3], functional J_2 is similar to the functional J_1 . Maximization of the functional J_2 is reduced to maximization of next functional:

$$J_3 = \sum_{i=1}^m \left(\sum_{x_j, x_s \in A_i} \sigma_j \sigma_s (x_j, x_s) \right)^{1/2}, \quad (2)$$

both by splitting the set of parameters on groups A_1, \dots, A_m and by coefficient selection σ_j and σ_s , where each coefficient equals to 1 in absolute value. At the same time essential factors f_1, \dots, f_m are defined with next expression:

$$f_i = \sum_{x_j \in A_i} \sigma_j x_j \left/ \sum_{i=1}^m \left(\sum_{x_j, x_s \in A_i} \sigma_j \sigma_s (x_j, x_s) \right)^{1/2} \right., \quad (3)$$

where $i = 1, 2, \dots, m$. From the expression of the functional J_3 it follows that an important problem of functional maximization is to find the maximum of a positive definite quadratic form:

$$\sum_{i=1}^k \sum_{j=1}^k \sigma_i \sigma_j (x_i, x_j) \quad (4)$$

on the set of vertices of a k -dimensional hypercube of side 1 by modulo. This paper is devoted to present a heuristic algorithm for solving the problem Maximization of positive definite quadratic form on the set of vertices of k -dimensional hypercube. The formulated problem has a trivial solution obtained by complete enumeration method. However, this method of solving the problem is computationally inefficient. Previously formulated problem is reduced to the minimum negative cut problem for weighted graphs with an arbitrary weight function that is bounded from above and below.

Let $R = \|(x_i, x_j)\|$ – matrix of correlation coefficients of the parameters of the set $\{x_i, \dots, x_k\}$, that corresponds to an arbitrary group from the set of groups $\{A_1, \dots, A_m\}$. We define a weighted graph on this set

as follows. To each parameter x_i it is assigned vertex x_i . Each pair of vertices is connected by an edge. Let V – set of edges of the obtained complete graph. On a set of edges we construct a function according to the rule:

$$c_{ij} = c(v_{ij}) = (x_i, x_j), i, j = 1, \dots, k, i \neq j, \quad (5)$$

where v_{ij} – an edge connecting vertices x_i and x_j . We assume that there are no loops at the vertices. The matrix $C = \|c_{ij}\|$ of the weight function c_{ij} is obtained from the matrix R by replacing its diagonal elements with zeros. To denote the resulting weighted graph we will use the notation $L(X, V, C)$.

Let us consider the set of all cuts of the graph $L(X, V, C)$ into two subgraphs. Obviously, each of the cuts corresponds to a partition of the matrix of correlation coefficients of the parameters x_1, \dots, x_k and, accordingly, the matrix C into blocks:

$$R = \begin{array}{cccccc} (x_1, x_1) & \dots & (x_1, x_s) & (x_1, x_{s+1}) & \dots & (x_1, x_k) \\ \vdots & R_1 & \vdots & \vdots & R_2 & \vdots \\ (x_s, x_1) & \dots & (x_s, x_s) & (x_s, x_{s+1}) & \dots & (x_s, x_k) \\ \hline (x_{s+1}, x_1) & \dots & (x_{s+1}, x_s) & (x_{s+1}, x_{s+1}) & \dots & (x_{s+1}, x_k) \\ \vdots & R_3 & \vdots & \vdots & R_4 & \vdots \\ (x_k, x_1) & \dots & (x_k, x_s) & (x_k, x_{s+1}) & \dots & (x_k, x_k) \end{array} \quad (6)$$

The main minors R_1 and R_4 of orders s and $k-s$ determine, according to the rule described above, subgraphs of graph cut, which are also complete. A partition of the set X into two subsets A and \bar{A} corresponds to each cut of the graph $L(X, V, C)$. The magnitude of the cut is the sum of the correlation coefficients of the antidiagonal blocks R_2 and R_3 , which is determined from next expression:

$$W(A, \bar{A}) = \sum_{(x_i, x_j) \in R_2} (x_i, x_j) + \sum_{(x_i, x_j) \in R_3} (x_i, x_j) \quad (7)$$

If $W(A, \bar{A}) < 0$, then the cut will be called negative, otherwise – non-negative. The cut with the smallest value of $W(A, \bar{A})$ will be called minimal. Let construct a map of the set of all cuts of the graph $L(X, V, C)$ to the set of vertices of the k -dimensional hypercube of side 1 by modulo. To do this, we assign one-to-one correspondence to each i -th dimension ($i = 1, \dots, k$) of the hypercube some parameter x_i . To designate cuts of the graph $L(X, V, C)$ in terms of subsets of its vertices into which given graph is partitioned when all

edges of the cut are removed. The map φ of the set of all cuts of the graph to the set of vertices of the k -dimensional hypercube is defined as follows:

$$\begin{aligned}\varphi(A, \bar{A}) &= (1, \dots, 1), \text{ if } W(A, \bar{A}) \geq 0, \\ \varphi(A, \bar{A}) &= (\sigma_1, \dots, \sigma_k), \text{ if } W(A, \bar{A}) < 0,\end{aligned}\tag{8}$$

where $\sigma_i = -1$ under $x_i \in A$ and $\sigma_i = 1$ under $x_i \in \bar{A}$.

Theorem 1. Let matrix R defines a graph that contains at least 1 negative cut. Next, an arbitrary negative cut (A, \bar{A}) and its vertex $\sigma' = (\sigma'_1, \dots, \sigma'_k)$ after application of map φ . Suppose that hypercube dimensions are reassigned in a way that next is true:

$$\sigma'_1 = \dots = \sigma'_s = -1,\tag{9}$$

By multiplying rows and columns of matrix R on corresponding components of vector σ' matrix R' will be obtained. From definition of vertex σ' it is follows that elements (x_i, x_j) , $i = \overline{1, k}$, $j = \overline{s+1, k}$ of matrix R change their sign. Since negative cut (A, \bar{A}) was chosen arbitrarily for any such cut it is possible to write:

$$4W(A, \bar{A}) = \sum_{i,j=1}^k (x_i, x_j) - \sum_{i,j=1}^k \sigma_i \sigma_j (x_i, x_j) < 0,\tag{10}$$

i.e. expression (4) at vertex σ' obtains value higher that at vertex $\sigma = (1, \dots, 1)$ by exactly quadrupled value of negative cut for which σ' is an image. Since minuend at (E1) constant then $W(A, \bar{A})$ will be minimal only when subtrahend is maximal. It implies that if graph $L(X, V, C)$ does not contain negative cuts then maximum for expression (4) on hypercube vertices reach minimal negative graph cut at vertex $(1, \dots, 1)$ or otherwise at vertex that is an image of minimal negative graph cut. The quadratic form (4) attains maximal value at that vertex of the k -dimensional hypercube, which is the image of the minimal negative cut of the graph $L(X, V, C)$ under the map φ . This theorem allows us to construct an iterative algorithm to find the minimum negative cut of the graph $L(X, V, C)$. Suppose some procedure P for negative (not only minimum) cut search in weighted graph is given. As a result of its application to any graph some negative cut will be obtained or it will be determined that given graph does not consist negative cuts that can be found with this procedure. Weighted graph and solution vector, obtained on s -th step of the algorithm, are denoted as $L(X, V, C^s)$ and $\hat{\sigma}^s$ respectively. On

the first step of the algorithm $\hat{\sigma}^i = \sigma^1$. According to (5) map φ^s is defined on the set of graph cuts for $L(X, V, C^s)$:

1. Let on $(s+1)$ -th step application of procedure P to graph $L(X, V, C^s)$ gives negative cut (A^{s+1}, \bar{A}^{s+1}) . With map φ^s vector σ^{s+1} is obtained. If cut (A^{s+1}, \bar{A}^{s+1}) is not negative, then go to step 4.

2. Construct vector:

$$\hat{\sigma}^{s+1} = \bar{\sigma}^s \oplus \sigma^{s+1}, \quad (11)$$

where symbol \oplus denotes componentwise product of given vectors. For this vector next inequality occurs:

$$\sum_{i,j=1}^k \hat{\sigma}_i^{s+i} \hat{\sigma}_j^{s+i}(x_i, x_j) > \sum_{i,j=1}^k \hat{\sigma}_i^s \hat{\sigma}_j^s(x_i, x_j), \quad (12)$$

which follows from next relation:

$$4W(A^s, \bar{A}^s) = \sum_{i,j=1}^k \hat{\sigma}_i^s \hat{\sigma}_j^s(x_i, x_j) - \sum_{i,j=1}^k \hat{\sigma}_i^{s+1} \hat{\sigma}_j^{s+1}(x_i, x_j) < 0 \quad (13)$$

and definition of procedure P .

1. Multiply matrix R by rows and columns on corresponding coordinates of vector $\hat{\sigma}^{s+1}$. Matrix R^{s+1} , which defines graph $L(X, V, C^{s+1})$ is obtained. Go to item 1 of $(s+2)$ -th step.

2. If on $(s+1)$ -th step of procedure P application to $L(X, V, C^s)$ non-negative cut is obtained, then it means that iterative process is over. Vector $\hat{\sigma}^s$ is a solution.

Such algorithm will allow us to obtain a finite sequence of vertices of the k -dimensional hypercube $\hat{\sigma}^1, \dots, \hat{\sigma}^s$. From the construction of this sequence, it follows that at each subsequent vertex the quadratic form (4) receives greater value than at the previous one. At each iteration step the value of the quadratic form increases by a value, that is bounded from below. It follows that the algorithm converges in a finite number of steps.

For such iterative algorithm of minimum negative cut search of the graph, it is true that at the vertex $\hat{\sigma}^s$, the expression (4) receives a local maximum in regard to the procedure P . This obtained maximum will be global only when the procedure P has next property: if the graph $L(X, V, C)$ contains at least one negative cut, then the result of the procedure P will be some negative cut.

Algorithm for minimal negative cut problem of weighted graphs

The minimum negative cut search continues the search procedure for an arbitrary negative cut of the graph. At first, the orientation of the graph is formed and the initial flows are arranged. As a result, the graph turns into a network. The resulting network is then converted to a network with non-negative bandwidths. Then, using the Ford-Fulkerson algorithm [4], a minimal network cut is computed and its analysis is performed. Let a complete k -vertex weighted graph $L(X, V, C)$ be given. We assume that the capacity matrix of the edges of the network obtained by orientation of the graph coincides with the weight matrix C . If i -th edges enter the network vertex and j -th edges exit, then we will say that this vertex has (i, j) type. We assume that the set of all such pairs of vertices $\langle x_i, x_j \rangle$, that $i < j$, is lexicographically ordered. Then, cyclical sorting for vertices pairs sequencing from this set is started. At each stage of the algorithm, we will consider one pair of vertices. Suppose that to the $(l+1)$ stage of the algorithm, we obtain matrices C^l and a k -dimensional vector \hat{C}^l , where coordinates equal to 1 in absolute value. In addition, let a number of selected negative cuts for the next vertices pairs sorting equals r and the pair of vertices $\langle x_s, x_t \rangle$ is under study. In this case, each stage contains 12 steps:

1. We assume that x_s is a source, x_t is a sink.
2. The edges incident to the vertex x_s are replaced by edges that go out of this vertex, and the edges incident to the vertex x_t are replaced by edges that return to it.
3. Enumeration of all 4-vertex subgraphs, generated by a set of vertices $\{x_s, x_i, x_j, x_t\}$, $i, j = \overline{1, k}$ and $i \neq j \neq s \neq t$. In each such subgraph the edge v_{ij} is replaced by an edge that goes out of the vertex x_i if $c_{si} + c_{jt} > c_{sj} + c_{it}$ and an edge, that returns to this vertex, if $c_{si} + c_{jt} > c_{sj} + c_{it}$. In case of equality $c_{si} + c_{jt} > c_{si} + c_{it}$ the edge v_{ij} is not oriented at this step.
4. Arrange all the vertices of a partially oriented graph in decreasing order by the number of edges that return to them. Then, all the vertices are successively considered and replace incident to them edges that are not oriented by the moment of scanning, by the edges that go out of them.
5. Arrange the original streams $f(\bar{v}_{ij})$ according to rules:
 - a) $f(\bar{v}_{ij}) = -c$, if $i \neq j \neq s \neq t$,

$$b) f(\bar{v}_{si}) = -(q - p + 1)c, f(\bar{v}_{it}) = -c,$$

if vertex has type (p, q) and $q > p$,

$$c) f(\bar{v}_{si}) = f(\bar{v}_{it}) = -c, \text{ if vertex } x_i \text{ has type } (p, q) \text{ and } p = q,$$

$$d) f(\bar{v}_{si}) = -c, f(\bar{v}_{it}) = -(q - p + 1)c, \text{ if vertex } x_i \text{ has type } (q, p) \text{ and } q > p, \text{ where } c = \max_{v_{ij} \in V} |c_{ij}|.$$

6. Build a new network with non-negative capacities by using formula $\tilde{c}(\bar{v}_{ij}) = c(\bar{v}_{ij})$ for all $i, j = \overline{1, k}$ and $i \neq j$. Initial flows over all edges are set to zero.

7. Apply Ford-Fulkerson algorithm to find the minimum cut of the resulting network.

8. Check if the resulting network cut is a graph cut using the algorithm for collocation labeling [4]. Value of the obtained minimum cut of initial network is determined from the expression:

$$\overline{W}(A^{l+1}, \overline{A}^{l+1}) = \sum_{i=1, i \neq s} (f^*(\bar{v}_{si}) + f(\bar{v}_{si})), \quad (14)$$

where $A^{l+1}, \overline{A}^{l+1}$ – vertex sets into which the graph is cut. Go to step 11.

9. If the network cut is not a graph cut, then we divide the set of all vertices of the network, excluding source and sink, into non-overlapping subsets. In order to do so suppose that all these vertices are arranged in descending order according to number of its exit edges. If vertex is reachable from the source, then all vertices x_1, \dots, x_{s_1} that are reachable from x_s will be successively appended to subset A_1 until first unreachable vertex x_{s_1+1} . In case, when vertex x_1 is unreachable subset A_1 is appended by all vertices x_1, \dots, x_{s_1} that are not reachable from vertex x_s until first attainable vertex x_{s_1+1} . This process proceeds for all vertices. As a result, a sequence of non-overlapping subsets A_1, \dots, A_{k_1} will be obtained. In this sequence subsets of reachable and unreachable vertices will alternate. Next, complete set of these subsets is divided onto two non-overlapping groups. First group consist of those subsets, where vertices are reachable from x_s . The rest of subsets create second group.

10. Based on the initial network a new network is built by changing the orientation of some edges. Edge orientation that exit second group's vertices and enter first group's ones change to opposite. For all vertices that were

changed new flows for edges $\bar{v}_{si}, \bar{v}_{jt}$ are set according to step 6 and for the rest of edges source flow is unchanged. Go to step 7.

11. If there is inequality $\bar{W}(A^{l+1}, \bar{A}^{l+1})$, then we construct the vector σ^{l+1} , new matrix of weights C^{l+1} is created by using the matrix C^l and components σ^{l+1} and the vector $\hat{\sigma}^{l+1} = \hat{\sigma}^l \oplus \sigma^{l+1}$. Increase the counter of found negative cuts r by one. If $\bar{W}(A^{l+1}, \bar{A}^{l+1}) \geq 0$, then define matrix $C^{l+1} = C^l$ and proceed further.

12. Check if the pair $\langle x_s, x_t \rangle$ is the last in a lexicographically ordered list of all pairs of vertices. If true and $r = 0$, then the work is finished and $\hat{\sigma}^{l+1}$ is a solution. If $r \neq 0$, then define $r = 0$ and proceed to step 1. If the pair $\langle x_s, x_t \rangle$ is not the last one, then select the next pair and go to step 1 of $(l + 2)$ -th stage.

Next theorems were used to substantiate given algorithm.

Theorem 2. For every cut (A, \bar{A}) of canonic network to given arbitrary network there is next equality:

$$\begin{aligned}
 W^*(A, \bar{A}) &= \sum_{v_{ij} \in V(A, \bar{A})} c^*(v_{ij}) = \sum_{v_{ij} \in V(A, \bar{A})} (cv_{ij}) - c_0(\bar{v}_{ij}) = \sum_{v_{ij} \in V(A, \bar{A})} c(\bar{v}_{ij}) - \\
 &- \sum_{v_{ij} \in V(A, \bar{A})} c_0(\bar{v}_{ij}) + \sum_{v_{ij} \in V(A, A)} c_0(\bar{v}_{ij}) - \sum_{v_{ij} \in V(A, A)} c_0(\bar{v}_{ij}) = \sum_{v_{ij} \in V(A, \bar{A})} c(\bar{v}_{ij}) - \\
 &- \sum_{v_{ij} \in V(A, A)} c_0(\bar{v}_{ij}) - \left(\sum_{v_{ij} \in V(A, \bar{A})} c_0(\bar{v}_{ij}) - \sum_{v_{ij} \in V(A, A)} c_0(\bar{v}_{ij}) \right) = \bar{W}(A, \bar{A}) - v_0.
 \end{aligned} \tag{15}$$

Since initial flow value through any cut of an arbitrary network is constant then from above mentioned equality it follows that minimum cuts of an arbitrary network coincide with minimum cuts of its canonic network. Therefore, it is necessary to switch to canonic network and apply Ford-Fulkerson algorithm to it in order to find minimum cut of initial arbitrary network. Every graph with an arbitrary weight function can be transformed to an arbitrary network by selecting some vertex pairs as a source and a sink and certain orientation of its edges. Each such network is associated with number, that equals to minimum cut. Size of minimum cut for obtained network depends on graph orientation and choice of vertex pairs. Thus, some minimum arbitrary network is obtained.

Theorem 3. Let some weighted graph $L(X, V, C)$ and its minimal cut (A, \bar{A}) are given. Next, we choose vertices x_s, x_t from set A and \bar{A} in arbitrary way and take them as source and sink respectively.

All edges that are incident to vertex x_s substitute to arcs that exit this vertex and edges that are incident to vertex x_t substitute to entering arcs.

Edges that connect vertices from set A to vertices from set \bar{A} replace with arcs that exit set A . Orientation for the rest edges we choose arbitrary. For every arc we assign flow capacity that equals to weight of corresponding edge. Initial flow function is chosen according to 6th step of algorithm for minimal negative graph cut search. It is clear that for obtained network conditions (7) - (9) are true. Let show that cut (A, \bar{A}) of obtained network will be its minimal cut. Suppose that network cut (A^*, \bar{A}^*) is smaller then there is an inequality:

$$\begin{aligned} W(A, \bar{A}) &= \bar{W}(A, \bar{A}) = \sum_{\bar{v}_{ij} \in V(A, \bar{A})} c(\bar{v}_{ij}) > \bar{W}(A^*, \bar{A}^*) = \\ &= \sum_{\bar{v}_{ij} \in V(A^*, \bar{A}^*)} c(\bar{v}_{ij}) - \sum_{\bar{v}_{ij} \in V(A^*, \bar{A}^*)} c_0(\bar{v}_{ij}). \end{aligned} \quad (16)$$

With next condition:

$$c_0(\bar{v}_{ij}) \leq f(\bar{v}_{ij}) \leq c(\bar{v}_{ij}), \quad c_0(\bar{v}_{ij}) < 0, \quad |c_0(\bar{v}_{ij})| \geq c(\bar{v}_{ij}) \quad (17)$$

It is follows that:

$$\begin{aligned} \bar{W}(A^*, \bar{A}^*) &= \sum_{\bar{v}_{ij} \in V(A^*, \bar{A}^*)} c(\bar{v}_{ij}) - \sum_{\bar{v}_{ij} \in V(A^*, \bar{A}^*)} c_0(\bar{v}_{ij}) \geq \\ &\geq \sum_{\bar{v}_{ij} \in V(A^*, \bar{A}^*)} c(\bar{v}_{ij}) + \sum_{\bar{v}_{ij} \in V(A^*, \bar{A}^*)} c(\bar{v}_{ij}) = W(A^*, \bar{A}^*). \end{aligned} \quad (18)$$

Thereby, $\bar{W}(A, \bar{A}) > \bar{W}(A^*, \bar{A}^*)$ that contradict with minimum condition of graph cut (A, \bar{A}) . And it is follows that for any graph with an arbitrary weight function there is such orientation and vertex pair option as source and sink for which minimum cut of the obtained network coincide with minimum cut of the graph and also equals to it and such network is a minimum arbitrary network.

Theorem 4. Let minimal network cut (A, \bar{A}) is not the graph cut. Set of all vertices is divided on subsets in a way that is described in 9th step of the

algorithm. Suppose that $A_{i_1}, \dots, A_{i_{k_1}}$ and $A_{j_1}, \dots, A_{j_{k_2}}$ – reachable and

unreachable vertex subsets respectively. Suppose $B = \bigcup_{s=1}^{k_1} A_{i_s}$ and

$B' = \bigcup_{s=1}^{k_2} A_{j_s}$. In this regard minimal cut (A, \bar{A}) has next value:

$$\bar{W}(A, \bar{A}) = \sum_{x_i \in B} c(\bar{v}_{si}) + \sum_{\bar{v}_{ij} \in V^-(B, B')} c(\bar{v}_{ij}) + \sum_{x_i \in B'} c(\bar{v}_{it}) - \sum_{\bar{v}_{ij} \in V^-(B', B)} c_0(\bar{v}_{ij}). \quad (19)$$

After reorientation of arcs from set $V^-(B', B)$ we obtain network in which the same cut (A, \bar{A}) has value:

$$\bar{W}'(A, \bar{A}) = \sum_{x_i \in B'} c(\bar{v}_{si}) + \sum_{\bar{v}_{ij} \in V^-(B, B')} c(\bar{v}_{ij}) + \sum_{x_i \in B'} c(\bar{v}_{it}) + \sum_{\bar{v}_{ij} \in V^{**}(B, B')} c(\bar{v}_{ij}). \quad (20)$$

where $V^*(B, B')$ – arc set that link vertex sets B and B' where reorientation have not applied and $V^{**}(B, B')$ – arc set where reorientation was applied. According to condition (17) it is follows that $\bar{W}(A, \bar{A}) = \bar{W}'(A, \bar{A}) \leq \bar{W}(A, \bar{A})$ and then minimal cut of new network does not increase in accordance to theorems 2 and 3. As a result it can be stated that if minimal cut of arbitrary network is not a cut for corresponding graph then there is effective algorithm for reorientation that allows for minimal cut of new network not to increase and it is possible to achieve network with the cut that also belongs to the graph in finite number of steps. Hence, minimum negative cut problem for weighted graphs with an arbitrary weight function is reduced to search problem for graph orientation, that all edges of its minimum cut (A, \bar{A}) exit vertices of set A . Previously stated algorithm of minimum negative cut search of the graph is based on idea of minimum arbitrary cut search that will enable to find global minimum negative cut of given graph.

In this paper correlation coefficients were used for dependence estimation that imply some limitation on overall estimation of dependence between vectors since there is difference between dependence measurements of n random variables and pairwise dependence measurements. One of approaches to tackle this problem was presented in work [9] where new method that able to distinguish between pairwise independence and higher-order dependence of random vectors is shown. Since the main aim of this paper represent another

subject more details on above mentioned aspect will be examined in further works.

Results of given algorithm application to the real-world data will be presented in the following publications on this topic.

REFERENCES

- [1] O. Kwon and K. Ma. What Would a Graph Look Like in This Layout? A Machine Learning Approach to Large Graph Visualization. IEEE Transactions on Visualization and Computer Graphics, 2018. – vol. 24. – P. 478 – 488.
- [2] Айвазян, В.М. Бухштабер, И.С. Енюков, Л.Д. Мешалкин; под ред. С.А. Айвазяна. - М.: Финансы и статистика, 1989. – 607с.
- [3] Э. М. Браверман. Методы экстремальной группировки параметров и проблема распределения существенных факторов. Автоматика и телемеханика, 1970. – вып. 1. – С. 123 – 132.
- [4] L. R. Ford and D. R. Fulkerson. Flows in networks. Princeton University Press, 1962. – 332 p.
- [5] H. Rogers. Theory of Recursive Functions and Effective Computability. MIT Press, 1987. – 17 p.
- [6] M. R. Garey and D. S. Johnson. Computers and Intractability: A Guide to the Theory of NP-Completeness. W. H. Freeman, 1979. – 347 p.
- [7] J. E. Savage. The Complexity of Computing. Krieger Pub Co, 1987. - 333 p.
- [8] O. Kwon and K. Ma. What Would a Graph Look Like in This Layout? A Machine Learning Approach to Large Graph Visualization. IEEE Transactions on Visualization and Computer Graphics, 2018. – vol. 24. – P. 478 – 488.
- [9] B. Böttcher, M. Keller-Ressel, and R. L. Schilling. Distance Multivariate: New dependence measures for random vectors. The Annals of Statistics, 2019. – P. 2757 – 2789.
- [10] D. H. Bailey. Experimental Mathematics in Action. 2007. – 323 p.
- [11] A. Skorokhod. Random Perturbation Methods with Application in Science and Engineering. Springer – New York, 2002. – 489 p.

УМЕНЬШЕНИЕ РАЗМЕРНОСТИ И ГРУППИРОВКА ЭКСТРЕМАЛЬНЫХ ПАРАМЕТРОВ ДЛЯ ЗАДАЧ ВЗВЕШЕННОГО ГРАФА

Востров Г., Хриненко А.

В работе представлен подход графового разреза к уменьшению размерности и показаны связи между задачей группирования экстремальных параметров и задачей максимизации положительно определенной квадратичной формы на множестве вершин k -мерного гиперкуба стороны l по модулю. Показано, что последняя задача сводится к поиску минимального отрицательного разреза для взвешенного графа. Приведен эвристический алгоритм решения задачи.

Ключевые слова. Наборы данных, уменьшение размерности, графы.

ВЫВОДЫ

В выполненной коллективной монографии приведены результаты научных исследований в области информационных интеллектуальных технологий, моделирования в информационных управляющих системах, управления защитой информационных систем, совершенствования информационно-ресурсного обеспечения науки.

Приведенные материалы позволили решить ряд задач связанных с: совершенствованием информационно-ресурсного обеспечения науки, техники и социальной сферы; способы и методы защиты информации;

информационными интеллектуальными технологиями для автоматизированных систем обработки данных и управления;

математическим моделированием и оптимизацией в информационных управляющих системах;

информационными технологиями управления проектами.

Результаты выполнения работ по перечисленным разделам позволяют решить некоторые проблемы информационных управляющих систем и технологий.

ІНФОРМАЦІЯ О АВТОРАХ

ВАСЯНИН Володимир, доктор техн. наук, старший науковий співробітник, провідний науковий співробітник Інституту телекомунікацій і глобального інформаційного простору НАН України, Чоколівський бульвар, 13, м. Київ, Україна 03186, archukr@meta.ua, orcid.org/0000-0003-4046-5243

ВОСТРОВ Георгій, кандидат техн. наук, доцент кафедри прикладної математики і інформаційних технологій, Одеський національний політехнічний інститут, Шевченко, 1, Одеса, Україна, 65044, vostrov@gmail.com, orcid.org/0000-0003-3856-5392

ВЫЧУЖАНИН Владимир, доктор техн. наук, профессор, заведующий кафедрой "Информационные технологии", Одесский национальный политехнический университет, пр. Шевченко, 1, Одесса, Украина, 65044, 126.ist.onpu@gmail.com, orcid.org/0000-0002-5244-5808

БОБОК Іван, кандидат техн. наук, доцент кафедри інформатики та управління захистом інформаційних систем, Одеський національний політехнічний університет, пр. Шевченко, 1, Одеса, Україна, 65044, onu_metal@ukr.net, orcid.org/0000-0003-4548-0709

ГРИКУН Дмитрій, Київський національний університет технологій та дизайну, вул. Немировича-Данченка, 2, м. Київ, Україна, 01011, morph333@gmail.com

ДРОЗД Оксана, кандидат техн. наук, доцент, заступник директора по організаційно-виховній роботі Херсонської філії Національного університету кораблебудування імені адмірала Макарова, Ушакова, 44, м. Херсон, Україна, 73022, oksanadrozd183@gmail.com, orcid.org/0000-0002-01358659

ЄГОЛЬНИКОВ Олександр, викладач кафедри автоматизації та комп'ютерно-інтегрованих технологій Миколаївського коледжу транспортної інфраструктури Дніпровського національного університету залізничного транспорту імені академіка В. Лазаряна, вул. Пушкінська, 71, м. Миколаїв, Україна, 54020, alegolnikov1990@gmail.com, orcid.org/0000-0003-1033-3573

ЄГОШИНА Ганна, кандидат техн. наук, доцент кафедри «Інформаційні технології», Одеська національна академія зв'язку ім. О. С. Попова, вул. Пішоновська 5, м. Одеса, Україна, 65029, ann.kafsii@gmail.com, orcid.org/0000-0002-2381-1231

ЖУРИЛЕНКО Борис, кандидат фізико-математических наук, старший научный сотрудник, доцент кафедри автоматизации и энергоменеджмента аэрокосмического факультета Национального авиационного университета, ул. Рыбная 24а, г. Киев-138, Україна, 03138, zhurylenko@gmail.com, orcid.org/0000-0003-2980-5630

КАЗИМИРЕНКО Юлія, кандидат техн. наук, доцент, кафедра інформаційних управляючих систем та технологій, Національний університет кораблебудування імені адмірала Макарова, проспект Героїв Сталінграду, 9, м. Миколаїв, Україна, 54025, u.a.kazimirenko@gmail.com, orcid.org/0000-0002-7097-5335

КАЛБАЗОВ Дем'ян, директор департаменту професійних послуг і аутсорсинга ДП "ЭС ЭНД ТИ УКРАИНА", пр. Академика Палладина, г. Киев, Україна, 4403142, Demyan.Kalbazov@snt.ua

КІРЬЯЗОВ Ю., заступник директора, «Адес-авто», Україна

КОБОЗЄВА Алла, доктор техн. наук., професор, завідувач кафедри інформатики та управління захистом інформаційних систем, Одеський національний політехнічний університет, пр. Шевченко, 1, Одеса, Україна, 65044, alla_kobozeva@ukr.net, orcid.org/0000-0001-7888-0499

КОНОВАЛОВ Сергій, аспірант кафедри технічної кібернетики й інформаційних технологій, Одеський національний морський університет, вул. Мечникова 34, м. Одеса, Україна, 65029, konserge-serge@rambler.ru, orcid.org/0000-0002-2533-8660

КОРАБЛЕВ Вячеслав, ст. преподаватель, Южноукраинский национальный педагогический университет имени К.Д. Ушинского, Старопортофранківська, 26, Одеса, Україна, 65000, k.viacheslav@hotmail.com, orcid.org/0000-0002-7544-1252

КОРАБЛЕВ Николай, доктор техн. наук, кафедри комп'ютерних інтелектуальних технологій і систем, Харківський національний університет радіоелектроніки, пр. Науки, 14, г. Харків, Україна, 61166, mykola.korablyov@nure.ua, orcid.org/0000-0002-8931-4350

ЛИСЕЦКИЙ Юрий, доктор техн. наук, Генеральный директор ДП "ЭС ЭНД ТИ УКРАИНА", пр. Академика Палладина, г. Киев, Україна, 4403142, Yurii.Lysetskyi@snt.ua, orcid.org/0000-0002-5080-1856.

МАЗУРОК Татьяна, доктор техн. наук, професор, зав.кафедрой прикладной математики и информатики, Южноукраинский национальный педагогический университет имени К.Д. Ушинского, Старопортофранківська, 26, Одеса, Україна, 65000, tl@mazurok.com, orcid.org/0000-0002-7829-4446

МИХАЙЛОВ Сергей, доктор техн. наук, професор, Национальный университет «Одесская морская академия», ул. Дидрихсона, 8, Одеса, Україна, 65000, smikhailov@i.ua, orcid.org/0000-0002-3218-2238

НИКОЛАЕВА Надежда, инженер 1-ой категории кафедры автоматизации и энергменеджмента Национального авиационного университета, ул. Рыбная 24а, г. Киев-138, Україна, 03138, nnikolaeva1717@gmail.com

ОПИАТА Роман, аспирант кафедры прикладной математики и информационных технологий, Одесский национальный политехнический университет, пр. Шевченко, 1, Одесса, Украина, 65044, roma.opyata@gmail.com, orcid.org/0000-0001-5806-9615

ОТРАДСКАЯ Татьяна, кандидат техн. наук, директор частного учебного заведения «Одесский колледж компьютерных технологий «Сервер», Багрицкого, 126, г. Одеса, Украина, 65026, tv_61@ukr.net, orcid.org/0000-0002-5808-5647

ПЕТРОВ Игорь, кандидат техн. наук, профессор кафедры морских перевозок, Национальный университет «Одесская морская академия», Дидрихсона, 8, г. Одеса, Украина, 65029, firmness@list.ru, orcid.org/0000-0002-8740-6198

РУДНИЧЕНКО Николай, кандидат техн. наук, доцент кафедры информационных технологий, Одесский национальный политехнический университет, пр. Шевченко, 1, Одесса, Украина, 65044, nickolay.rud@gmail.com, orcid.org/0000-0002-7343-8076

СИТНИКОВ Валерій, доктор техн. наук, профессор, завідуючий кафедрою «Комп'ютерні системи», Одесский национальный политехнический университет, пр. Шевченко, 1, Одесса, Украина, 65044, sitnvs@gmail.com, orcid.org/0000-0003-3229-5096

СОЛОВЬЕВ Дмитрий, аспирант кафедры компьютерных интеллектуальных технологий и систем, Харьковский национальный университет радиоэлектроники, пр. Науки, 14, г. Харьков, Украина, 61166, solovyov.dima@gmail.com

ТЕПЛЕЧУК А., старший викладач кафедри «Автомобільний транспорт», Одесский национальный политехнический университет, пр. Шевченко, 1, Одесса, Украина, 65044

ТРОФИМЧУК Олександр, член-корреспондент НАН України, професор, доктор техн. наук, директор Інституту телекомунікацій і глобального інформаційного простору НАН України, Чоколівський бульвар, 13, м. Київ, Україна 03186, itelua@kv.ukrtel.net, orcid.org/0000-0003-3358-6274

УХІНА Г., аспірант кафедри «Комп'ютерні системи», Одесский национальный политехнический университет, пр. Шевченко, 1, Одесса, Украина, 65044, anyuta.uhina@gmail.com, orcid.org/0000-0003-3797-1460

УШАКОВА Людмила, здобувач наукового ступеня кандидата технічних наук, провідний інженер Інституту телекомунікацій і глобального інформаційного простору НАН України, Чоколівський бульвар, 13, м. Київ, Україна 03186, archukr@meta.ua, orcid.org/0000-0002-9020-1329

ФОМИЧЕВ Александр, кандидат техн. наук, старший преподаватель кафедры электронных вычислительных машин, Харьковский национальный университет радиоэлектроники, пр. Науки, 14, г. Харьков, Украина, 61166, alexandros_1985@mail.ru

ХОРОШКО Владимир, доктор техн. наук, профессор, Національний авіаційний університет, просп. Космонавта Комарова, 1, м Київ, Україна, 03058, professor_va@ukr.net, orcid.org/0000-0001-6213-7086

ХРИНЕНКО А., Одесский национальный политехнический институт, Шевченко, 1, Одесса, Украина, 65044, khrynenko@stud.opu.ua, orcid.org/0000-0001-6000-2102

ЧУПРИНА Андрей, аспирант кафедры компьютерных интеллектуальных технологий и систем, Харьковский национальный университет радиоэлектроники, пр. Науки, 14, г. Харьков, Украина, 61166, chuprina2610@gmail.com

ЧУПРИНКА Віктор, доктор техн. наук, профессор кафедры компьютерных наук і технологій, Київський національний університет технологій та дизайну, вул. Немировича-Данченка, 2, м. Київ, Україна, 01011, Chuprinka_V_I@ukr.net, orcid.org/0000-0001-6869-3091

ЧУПРИНКА Наталія, кандидат техн. наук, Київський національний університет технологій та дизайну, вул. Немировича-Данченка, 2, м. Київ, Україна, 01011, Natasha-chup@ukr.net, orcid.org/0000-0003-1507-489X

ШЕВЦОВ Юрий, кандидат техн. наук, Национальный университет «Одесская морская академия», ул. Дидрихсона, 8, Одесса, Украина, 65000, yurishevcev1976@gmail.com, orcid.org/0000-0001-5507-3942

ШИБАЕВ Денис, Одесский национальный морской университет, аспирант кафедры технической кибернетики и информационных технологий, Мечникова 34, г. Одеса, Украина, 65044, denscreamer@gmail.com, orcid.org/0000-0002-3260-5843

ШИБАЕВА Наталья, кандидат техн. наук, доцент кафедры информационных технологий, Одесский национальный политехнический университет, пр. Шевченко, 1, Одесса, Украина, 65044, nati.shibaeva@gmail.com, orcid.org/0000-0002-7869-9953

ШПНАРЕВА Ірина, кандидат фізико математичних наук, доцент кафедри математичного забезпечення комп'ютерних систем, Одеський національний університет ім. І.І.Мечникова, вул. Дворянська, 2, м. Одеса, Україна, 65000, ishpinareva@gmail.com, orcid.org/0000-0001-9208-4923

ЮХИМЕНКО Бируте, кандидат економічних наук, доцент, зав. кафедрой «Прикладной математики и информационных технологий», Одесский национальный политехнический университет, пр. Шевченко, 1, Одесса, Украина, 65044, pm1987pm@gmail.com, orcid.org/0000-0002-2607-4687

У монографії відображені результати наукових досліджень в області інформаційних інтелектуальних технологій для автоматизованих систем обробки даних і управління, моделювання та оптимізації в інформаційних управляючих системах, управління захистом інформаційних систем.

Матеріали монографії будуть корисними для аспірантів, магістрантів, викладачів вищих навчальних закладів, що спеціалізуються в області ІТ-технологій.

The monograph reflects the results of scientific researches in the field of information intelligent technologies for automated data processing and control systems, modeling and optimization in information control systems, management of information systems protection.

The materials of the monograph will be useful for graduate students, undergraduates, teachers of higher educational establishments specializing in the field of IT technologies.

Наукове видання

ВАСЯНИН Володимир Олександрович
ВОСТРОВ Георгій Миколайович
ВИЧУЖАНИН Володимир Вікторович
та інші

ІНФОРМАЦІЙНІ УПРАВЛЯЮЧІ СИСТЕМИ ТА ТЕХНОЛОГІЇ. ПРОБЛЕМИ ТА РІШЕННЯ

Монографія

Українською, російською, англійською мовами

Формат 60×84/16. Ум. друк. арк. 14,18. Тираж 200 прим. Зам. 133 (22).

Надруковано з готового оригінал-макета

Видавництво «Екологія»

Свідоцтво суб'єкта видавничої справи ДК № 1873 від 20.07.2004

65091, м. Одеса, вул. Разумовська, 23 / 1

тел. (0482) 33-07-17, 37-14-25, (048) 7-855-855

email: astro_print@ukr.net; www.astroprint.ua; www.stranichka.in.ua

ФОП Побута М.І.

65044, м. Одеса, пр-т Шевченка, 1-а

Тел.: 770-24-74, email: pd5@ukr.net