

НАУЧНО • ТЕХНИЧЕСКАЯ ИНФОРМАЦИЯ

Серия 2. ИНФОРМАЦИОННЫЕ ПРОЦЕССЫ И СИСТЕМЫ
ЕЖЕМЕСЯЧНЫЙ НАУЧНО-ТЕХНИЧЕСКИЙ СБОРНИК

Издается с 1961 г.

№ 3

Москва 2020

ОБЩИЙ РАЗДЕЛ

УДК [002:004]–049.5

А.А. Грушо, М.И. Забежайло, В.О. Писковский, Е.Е. Тимонина

Индустрия 4.0: возможности и риски в контексте проблем информационной безопасности*

Представлен обзор проблематики обеспечения информационной безопасности информационно-телекоммуникационной инфраструктуры решений, развиваемых в рамках направления Индустрия 4.0. Обсуждаются текущее состояние и перспективы эффективной разрешимости ряда проблем, идентифицированных в данной области (в том числе – некоторые возможности использования моделей и методов искусственного интеллекта при решении задач обеспечения информационной безопасности соответствующих систем и продуктов)

Ключевые слова: *Индустрия 4.0, Интернет вещей (IoT), информационная безопасность, интеллектуальный анализ данных, математические модели и методы*

DOI: 10.36535/0548-0027-2020-03-1

* Работа выполнена при частичной поддержке РФФИ (проект 18-29-03081 «Разработка архитектурных решений по обеспечению информационной безопасности в гетерогенных информационных системах цифровой экономики»)

ВВЕДЕНИЕ

Термин *цифровая экономика* сегодня вошел в нашу повседневную жизнь. Объем этого понятия весьма широк: от *Больших Данных (Big Data)* и *искусственного интеллекта* до так называемых *смарт-контрактов (smart – умный, смысловой)* и *блокчейнов* (распределенных реестров), от *цифровизации* (перевода на цифровые технологии производства и управления) различных «традиционных» областей деятельности человека («умный» дом, беспилотный транспорт, ...) до преобразования к новому качеству целых отраслей индустрии (*умное сельское хозяйство, цифровое здравоохранение* и др.). Полезно будет напомнить, что в основании *цифровой экономики* лежит (за последние 3-5 лет уже несколько подзабытое) понятие так называемой *Индустрии 4.0*.

В середине 2000-х гг. в рамках регулярно проходящих в швейцарском Давосе экономических форумов была сформулирована и стала активно обсуждаться идея так называемого нового технологического уклада (получившего наименование *Индустрия 4.0*), в основе которого – возможности персонализированной «настройки» производств на индивидуальные потребности большого количества потребителей. Добиться этого, по мнению адептов данного подхода, позволяет максимальная интеграция перспективных информационно-коммуникационных технологий в технологические цепочки производства и распределения продукции: глобальные информационные инфраструктуры – это «платформа» и для учета индивидуальных потребностей заказчика, и для доставки ему персонально-ориентированной продукции. Достигший своего пика к 2009 г. (см., например, [1]) эмоциональный подъем как бизнес-сообщества, так и сферы высоких технологий, поддавшийся влиянию базирующейся на промышленной стратегии так называемой *Индустрии 4.0* идее очередной – четвертой – технологической революции, в последующие 5–7 лет преобразовался в серию национальных программ развития цифровой экономики (см., например, [2–7]) и в национальные стратегии (см., например, [8–12]) исследований и разработок для наиболее перспективных направлений информационно-телекоммуникационных технологий¹. Ключевая идея подхода, предлагаемого в рамках *Индустрии 4.0*, – «способствовать повсеместному развитию инфраструктуры компьютерных технологий таким образом, чтобы реальная мировая экономика стала частью, своего рода приложением соответствующей технологической экосистемы» [1] представлялась привлекательной и для бизнеса в целом (как основа будущего устойчивого экономического роста), и для лидеров компьютерной индустрии (которым, в отличие от ряда других высокотехнологичных направлений эта идея гарантировала достаточно короткие инновационные циклы смены «поколений» актуальных ИТ-решений с возвратом инвестиций в исследования и разработки, а также ук-

репление их доминирующего положения на глобальном рынке).

Однако подробный технический анализ ассоциируемых с *Индустрией 4.0* технологических проблем и вызовов, последовательно проводимый экспертным сообществом в прошедшее десятилетие, позволил идентифицировать ряд фундаментальных рисков, без эффективного управления которыми, по-видимому, придется следовать уже не хорошо известной формуле «хотелось – как лучше, а получилось – как всегда!», а предпринимать значительные усилия, «чтобы хуже не стало». Специфичные именно для технологии и решений *Индустрии 4.0* риски, «интерферируя» с рядом угроз, наследуемых из предыдущего уклада компьютерных технологий, требуют целенаправленной разработки адекватных инструментов управления и противодействия их влиянию². Обзору укрупненной «карты» таких рисков, ассоциируемых, в первую очередь, с проблематикой информационной безопасности (ИБ) рассматриваемой нами предметной области, и посвящена настоящая работа.

КАК ВСЁ НАЧИНАЛОСЬ

Технологическим основанием программ развития *Индустрии 4.0* является разработка и эксплуатация программно-аппаратных платформ, обеспечивающих функционирование в распределенной сетевой среде (интегрирующей локальные сетевые структуры и Интернет) гибко перестраиваемых комплексов бизнес-процессов, охватывающих весь цикл производства соответствующих продуктов и услуг (от сбора первичной информации датчиками и организации глобального обмена данными, до выполняемого специальными средствами поддержки принятия решений целенаправленного анализа эффективности процесса производства и управления им, а также удаленного использования различных средств поддержки непрерывности процесса производства³ – выявлению сбоев в работе оборудования, их устранению, удаленной консультационной поддержке эксплуатационного персонала и т.п.). Национальные программы развития *Индустрии 4.0* имеют определенные архитектурные и организационные особенности. Так, например, в отличие от США, где ключевым элементом соответствующей сферы деятельности определен Интернет вещей (*Internet of Things – IoT*) [2], в Германии системообразующими составными частями формируемого комплекса решений стали унификация и стандартизация компонентов, разработка единой технологической платформы, а также отработка механизмов интеграции бизнес-процессов (от уровня данных, через технологические уровни – вплоть до средств управления и обеспечения\поддержки работоспособности) интегрируемых «умных» производств [4]. Великобритания нацелилась на 8 «великих технологий» [3], а во Франции в фокусе

¹ Далее, если это не потребует специальных уточнений, мы будем употреблять термин *компьютерные технологии*, имея в виду *информационные и телекоммуникационные технологии*.

² В том числе (по ряду направлений) начиная с уровня разработки адекватных математических моделей и алгоритмов компьютерного анализа данных и поддержки эффективных управленческих решений.

³ В том числе за счет обеспечения необходимого уровня информационной безопасности.

внимания – 10 перспективных технологических направлений будущего [5]. Между тем, если обратиться к высокотехнологичным основаниям этих примеров, то несложно убедиться, что в общем виде их перечень охватывается следующими четырьмя классами функций:

1) поддержка **сетевой доступности** (*connectivity*) для экономических **агентов**, вовлеченных в соответствующий комплекс бизнес-процессов;

2) обеспечение требуемого **уровня** (характеристик производительности, защищенности, ...) **доступа** к информационным **ресурсам**, необходимых для исполнения соответствующего комплекса бизнес-процессов;

3) эффективное управление **производительностью** используемой информационно-коммуникационной инфраструктуры (в том числе – ее архитектурными конфигурациями, актуальными емкостями⁴, пропускными способностями, уровнями защиты и др.), задействованной в соответствующем комплексе бизнес-процессов;

4) эффективная **интеграция** как данных, так и приложений, используемых экономическими агентами в соответствующем комплексе бизнес-процессов.

Опыт показал, что проблематика защиты данных и обеспечения информационной безопасности представляет собою важный аспект успешного развития этой области науки и технологий, так как в ней идентифицирован ряд принципиально важных вопросов, формирование надежных ответов на которые требует изучения соответствующих **открытых**, не имеющих на текущий момент удовлетворительного решения, **исследовательских** проблем. В качестве примеров можно привести тематику

надежной **аутентификации** (персон, источников сообщений, виртуальных машин, ...);

обеспечения **процессно-реального времени** анализа данных и поддержки принятия решений (с учетом оценок вычислительной сложности в контексте Больших Данных, необходимости поиска быстрых алгоритмов, выделения быстро разрешимых подклассов вычислительно ресурсоемких переборных задач и др.);

компьютерного анализа данных при работе с **незамкнутыми предметными областями** (в том числе с учетом особенностей не только *Vig-*, но и *Open*-эффектов, характерных для обработки Больших Данных);

пригодного для эффективной компьютерной обработки **представления знаний** о незамкнутых – пополняемых новыми сведениями – предметных областях,

а также ряд других областей, требующих применения релевантных их специфике математических моделей, методов и алгоритмов.

В самом общем виде единая для национальных версий программ **Индустрии 4.0** инфраструктурная схема может быть уложена в «формулу»

унификация «*составных элементов*» +
+ **единая технологическая/коммуникационная платформа** +
+ **интеграция компонентов и приложений**,

⁴ Например, для ресурсов (в том числе облачных) хранения данных и др.

которая хорошо известна специалистам уже почти четверть века по опыту разработки и эксплуатации предыдущих «поколений» компьютерных систем для крупных (банковских, индустриальных, ...) бизнес-приложений. Действительно, унификация функциональных компонентов (реализуемых соответствующими проблемно-ориентированными программными сервисами – см., например [13–15]), проблематика организации эффективного взаимодействия прикладных подсистем (переход от технологий *EAI*⁵ к *ESB*⁶-решениям – см., например, [15, 16]), крупных технологических решений и, наконец, потребность в интеграции компонентов сводного решения не только по функциональности, но и по задействованным в общих бизнес-процессах данным (формирования в каждой предметной области «базовой» типологии используемых сущностей – объектов, отношений между ними и т.п., позволяющей унифицировать как обмен данными, так и способы их «нормализации»⁷), а также хранения в крупных прикладных системах – см., например, [15, 17–22]) – вот те же три составные части представленной выше «формулы», выделенные уже как базовые элементы так называемого компонентно-интеграционного подхода [15, 23], обеспечившего успех многим крупным ИТ-проектам первого десятилетия XXI в. Уже тогда стала понятна критическая значимость проблем обеспечения информационной безопасности архитектурных решений в области информационной безопасности для крупных банковских информационных систем в [15].

Опыт разработки и эксплуатации больших банковских систем показал, что ключевую роль в обеспечении их операционной надежности играет проблематика защиты «доверенного» характера среды интеграции приложений (причем на всех её уровнях: от хранения и транспортировки данных до уровня собственно взаимодействия и синхронизации приложений, исполняющих соответствующие бизнес-процессы). Именно удачными архитектурными решениями, обеспечившими эффективную защиту компонентно-интеграционных архитектур (в частности конфигурациями типа *интеграционная шина + проблемно-ориентированные адапторы*⁸ как средство контроля доступа), обусловлены высокий уровень эксплуатационной надежности и, как следствие, широкое распространение подобных решений в арсеналах лидеров глобального рынка информационных технологий – IBM, SAP AG, Oracle и др.

Обсудим более подробно выделенные нами три базовые составляющие подхода, характерного для **Индустрии 4.0**: (1) стандартизацию компонентов решения, (2) коммуникационную инфраструктуру (по факту – *IoT*) и (3) интеграцию (по данным) с позиций

⁵ *Enterprise Application Integration* – интеграция приложений уровня предприятия

⁶ *Enterprise Service Bus* – интеграционная (сервисная) шина предприятия

⁷ Представления в форме, удобной для их последующей обработки разными программными инструментами

⁸ В то числе содержащие существенную аппаратную составляющую (способную обеспечить высокую производительность и отказоустойчивость при больших нагрузках)

оценки их устойчивости (способности обеспечивать *непрерывность* бизнеса в рамках решений *Индустрии 4.0*) по отношению к внешним вредоносным влияниям. Очевидно, что в части унификации (стандартизации) используемых компонентов и подходов текущая ситуация может быть оценена как вполне благоприятная: понятно, что необходимо делать (при этом многое уже сделано и надежно работает).

С интеграцией (по данным) ситуация менее обнадеживающая, хотя и позволяющая формировать по крайней мере локальные решения, которые способны функционировать вполне надежно. Так в «замкнутых» (ограниченных как по «периметру», так и по типам «составных элементов») технологических ландшафтах всё, что требуется, можно контролировать (например, за счет «погружения всех используемых типов данных в единую «объемлющую» их «среду» формирования компонентно-интеграционных архитектур, в которых задействованы сервис-конверторы, адапторы контроля доступа и т.п.). Примером комплексных технологических решений такого типа могут служить глобальные платёжные системы (VISA, Mastercard, ...). Здесь (внутри соответствующего «периметра») всё, что не идентифицируется как авторизованная сущность, немедленно удаляется из пределов защищенного ландшафта прикладных информационных систем. При необходимости расширить «замкнутый» ландшафт новый тип данных\сообщений снабжается соответствующими «обработчиками» (конверторами, адапторами и т.п.), которые в рамках стандартной процедуры обновления предыдущей версии «замкнутого» ландшафта «накатываются» на соответствующие компоненты общего текущего ландшафта. В этих условиях значимой оказывается проблема авторизации – обеспечение доверия к каждому подобному «обновлению». К соответствующей интеграционной технологии возникает вопрос: как обеспечить защиту от несанкционированных рассылок неавторизованных «обновлений»? Дополнительные трудности связаны с проблемой ненадежности – «дырявости»\компрометируемости – используемой коммуникационной архитектуры *IoT*.

К сожалению, совсем уже печально сегодня обстоят дела с надежностью коммуникационной инфраструктуры – *IoT*, используемой в рамках задаваемого требованиями *Индустрии 4.0* подхода. Огромное число точек генерации новых данных\сообщений плюс слабые вычислительные возможности в подавляющем большинстве таких точек выводят в разряд критически значимых проблему анализа данных и поддержки принятия таких решений (локально – в отдельных точках *IoT*-инфраструктуры), которые обеспечивали бы надёжную защиту от внешних вредоносных активностей. При оценках, в целом пессимистичных, текущей ситуации (см., например, [24]) некоторые надежды эксперты связывают с *SDN/NFV*⁹-подходом [25, 26], в рамках которого необходимый анализ данных и принятие

решений можно попытаться перенести на общий *SDN*-контроллер защищаемой сети. Дополнительные «дебютные соображения» по обеспечению непрерывности (бесперебойного режима) функционирования коммуникационной *IoT*-инфраструктуры аналитики связывают с архитектурными инструментами и возможностями «резервирования» надёжности коммуникационной инфраструктуры, примером которых являются в частности, предлагаемые компанией *CISCO Intent-Based Network Architectures* [27].

Чтобы проиллюстрировать характерные для обсуждаемой предметной области – коммуникационной инфраструктуры *IoT* – трудности, приведем два примера эффектов, с возникновением которых ассоциируются существенные риски надежности и непрерывности бизнес-систем *Индустрии 4.0*. Так, в частности, ещё в распределенных архитектурах предыдущего поколения (компонентно-интеграционных архитектурах) стали проявляться негативные эффекты нового типа – так называемые рассинхронизации и «разладки», «мерцающие» ошибки и т.п. (см., например, [28–30]), которые могут быть инициированы высококритичными внешними вредоносными воздействиями на интегрированные («собранные» из различных компонентов) ИТ-комплексы. Не менее серьезного внимания заслуживают и эффекты, возникшие совсем недавно при анализе свойств *5G*-коммуникационных решений. При целенаправленном формировании в сетях этого типа образуются так называемые «слепые» зоны, в которых злоумышленниками могут быть размещены вредоносные «закладки», которые переводят весь используемый соответствующим бизнес-решением коммуникационный ландшафт в некорректный режим функционирования [31, 32].

Переход к облачным технологиям и решениям вместе с новыми возможностями *SDN*- и *NFV*-подходов к выстраиванию интеграционных компьютерно-технологических сред принес также новые угрозы и вызовы в части информационной безопасности и защиты данных. Так, характерные именно для облачных ИТ-решений новые вредоносные эффекты оказались связанными с использованием собственно виртуализированных компонентов – виртуальных машин, обеспечивающих не только сами вычисления, но и хранение данных или же коммуникацию. В работах [30, 33, 34] мы уже рассматривали характерные для таких компьютерных архитектур угрозы (вместе с некоторыми вариантами организации противодействия их влиянию). Уместно упомянуть и о существенном вкладе аппаратных компонентов и решений в «реестр» угроз и рисков нештатного функционирования облачных технологических решений. Не менее важна и распределенность архитектурных комплексов обсуждаемого типа: проблема контроля их огромной размерности стимулировала развитие и совершенствование различных приемов разбиения общего для конкретного бизнес-решения информационно-технологического ландшафта на «самостоятельные» подобласти, относящиеся к зоне ответственности соответствующих средств обеспечения информационной безопасности. Так, в обзоре [24] представлен целый комплекс согласованных подхо-

⁹ Software-Defines Networks – программно-конфигурируемые сети, Network Functions Virtualization – виртуализация сетевых функций.

дов и решений, охватывающих меры выявления вредоносного программного обеспечения (так называемые технологии *malware protection*), противодействия несанкционированным проникновениям внутрь «периметра» системы (*intrusion protection*) и активностям инсайдеров, использования легкой криптографии, защиты маршрутизации (так называемых *secure routing and forwarding*).

Своего рода «вишенкой на торте» оказалась в рамках *Индустрии 4.0* проблема компрометации всей интегрированной среды конкретного бизнес-решения за счет первоначального вредоносного проникновения в один из её «слабых» сегментов. Значимым примером может служить подмена\компрометация цифрового «слежка» участника крупной платежной системы (см., в частности, текущую дискуссию о рисках некорректного использования биометрических данных в рамках национальной банковской инфраструктуры Российской Федерации [35]). Эксперты компании *Trend Micro* обращают особое внимание на этот класс уязвимостей ИТ-ландшафтов *Индустрии 4.0*, прогнозируя [32], что сотрудники компаний, работающие удаленно (в первую очередь – в сетях с низким уровнем защиты от взломов), станут новой целью злоумышленников. Перехват управления (подмена контакта с авторизованным сотрудником на контакт со злоумышленником в результате взлома системы аутентификации сотрудников\пользователей), а также инфицирование удаленных компонентов информационно-телекоммуникационного ландшафта корпоративной сети вредоносным программным обеспечением могут вести к утрате существенной информации и нештатному режиму функционирования соответствующей компьютерной инфраструктуры.

Отдельного обсуждения заслуживает специфика *Интернета вещей (IoT)* как своего рода «среды» содержательной интеграции объектов, активов и т.п., взаимодействие между которыми существенно в различных (значимых для *Индустрии 4.0*) смыслах: *everything is connected via IoT technologies*¹⁰. Всеобъемлющий характер такой интеграции порождает специальные требования к продуктам, сервисам и компонентам, задействованным в каждом вертикально интегрированном бизнес-решении. Чтобы обеспечить хотя бы мониторинг аномалий (возникновения «нетиповых» ситуаций) в нормальном режиме функционирования соответствующих интегрированных бизнес-решений необходимо, как минимум, погрузить все задействованные объекты, отношения между ними и т.п. в единую формализованную среду описания (например, в стиле [17–19, 36, 37] или же [38]), где «дифференциальную диагностику» состояний *НОРМА* и *ПАТОЛОГИЯ* функционирования соответствующих бизнес-процессов можно будет описать синтаксически и, как следствие, перенести в работающую в автоматическом режиме систему контроля и управления соответствующей ИТ-инфраструктурой.

Проиллюстрируем написанное на примере управленческих подсистем интегрированных решений *Индустрии 4.0*. Центральное место здесь занимаю так

называемые *ERP*¹¹-системы, в рамках которых различные *IoT*-компоненты (датчики, измерители и т.п.) в режиме реального времени автоматически собирают и передают текущие показатели в подсистему анализа данных и принятия решений. Автоматический (исключающий в общем случае участие человека-оператора) режим генерации и анализа таких данных проактивно влияет на клиенто-ориентированные *ERP*-процессы и, фактически, формирует новую (и при этом в существенной степени «завязанную» на *IoT*-инфраструктуру крупных бизнес-решений *Индустрии 4.0*) реальность в использовании *ERP*-систем. Важными, с точки зрения защиты от угроз внешнего вредоносного влияния, здесь оказываются:

- территориально распределенная, однако, целостная с точки зрения бизнес-логики информационно-технологическая, архитектура производства;
- вертикально интегрированные бизнес-цепочки, производства и даже отрасли;
- принятие решений средствами искусственного интеллекта (интеллектуальных систем анализа данных и поддержки принятия решений) на основе автоматически собираемых, так называемых *machine-generated* машинно-генерируемых, данных и др.

С «технической» точки зрения критичные последствия внешнего вредоносного воздействия связаны с нарушением целостности или же рассинхронизации «боевой» сети – коммуникационно-интеграционной *IoT*-инфраструктуры соответствующего бизнес-решения в рамках *Индустрии 4.0* (см., например, [24–26]).

Несколько замечаний о типологии угроз

Если взглянуть на обсуждаемую предметную область – устойчивость коммуникационно-интеграционной *IoT*-инфраструктуры бизнес-решений *Индустрии 4.0*, – так сказать, «с высоты птичьего полета», то подавляющую часть существенных вредоносных воздействий на объекты защиты здесь можно отнести к сравнительно небольшому набору классов угроз и вызовов, наиболее значимыми из которых являются:

- «прорыв» (нарушение целостности) защищаемого «периметра»;
- нарушение доверенности содержания сообщений, пересылаемых внутри и за пределы защищаемого «периметра»;
- компрометация доверенности (штатного режима функционирования) системы маршрутизации и адресации пересылаемых сообщений, а также нарушение сбалансированности (например, рассогласование, разладка и т.п.) функционирования системы маршрутизации и адресации пересылаемых сообщений;
- компрометация доверенности (штатного режима функционирования) системы «перевода» содержания сообщений, пересылаемых между компонентами сводного ИТ-ландшафта бизнес-решения *Индустрии 4.0* (нештатный режим интеграции по данным, которыми обмениваются компоненты соответствующей информационной системы);

¹⁰ Все соединено со всем через *Интернет вещей* как коммуникационную среду.

¹¹ *Enterprise Resource Planning systems* – системы планирования и управления ресурсами предприятия

- компрометация (утрата доверия, корректной авторизуемости) участников информационного обмена внутри и за пределы защищаемого «периметра», в частности – нештатный характер аутентификации участников информационного обмена.

Используемые для борьбы с такого рода «эффектами» подходы можно в самом общем виде разделить на две группы. В первую естественно отнести все те «инструменты», которые применяются в очевидных ситуациях типа ликвидации «пролома в крепостной стене». Во вторую попадает все то, что оказывается полезным в ситуации, когда необходимо иметь возможности идентифицировать и анализировать «АНОМАЛЬНЫЕ» события в функционировании объекта мониторинга и защиты с целью разделения «СЛУЧАЙНЫХ ФЛЮКТУАЦИЙ» и «ЦЕЛЕНАПРАВЛЕННЫХ ВРЕДНОСНЫХ ВОЗДЕЙСТВИЙ» (в том числе на основе выборок, содержащих расширяющееся, однако, ограниченное множество ранее уже проанализированных прецедентов). Ключевую роль здесь играют интеллектуальный анализ данных и поддержка принятия решений в решаемых задачах диагностического типа с учетом эффектов причинности. При этом необходимо помнить (принимая во внимание размеры обсуждаемых ИТ-ландшафтов), что все требуемые процедуры анализа данных, принятия и реализации соответствующих управленческих решений (цель которых – обеспечить восстановление (причем за приемлемое для соответствующего бизнеса время) штатного режима работы защищаемого объекта – *disaster recovery*) должны исполняться в автоматическом режиме, т.е. с использованием соответствующих решений на базе искусственного интеллекта [24, 39].

Основные подходы для противодействия угрозам и вызовам

Если обобщить приведенные нами описания угроз и вызовов, то перечень проблем, с которыми приходится иметь дело в подавляющем большинстве инцидентов [40-43] можно свести к четырём группам:

- целенаправленные нарушения штатного режима функционала сетевой доступности (*connectivity*): целенаправленно вызываемые нарушения достижимости – узел–узел, группа узлов–группа узлов и т.п., сбой в маршрутизации, рассинхронизации и т.п. (см. предыдущий Раздел);
- идентификация и анализ аномалий, выполняемые локально и нацеленные, в первую очередь, на восстановление после идентифицированных вредоносных воздействий (которые необходимо осуществить в условиях существенных ограничений на ресурсы процессора и устройства памяти на локальных устройствах);
- аутентификация (персонажей, пакетов сообщений, используемых в облачной среде вычислительных объектов – виртуальных машин) и подтверждение статуса авторизованности участников информационного обмена в рамках соответствующей ИТ-инфраструктуры;
- анализ новых (ранее не описанных при конфигурировании ИТ-инфраструктуры и настройке соответствующей системы информационной безопасности

как допустимые) состояний объекта управления (с целью идентификации новых допустимых состояний и организации противодействия вредоносным воздействиям в нештатных ситуациях) – реакция на характерный для Больших Данных эффект *Open* (открытости и пополняемости имеющихся данных принципиально новой информацией). В этой группе особенно очевидна необходимость в использовании новейших математических моделей и методов интеллектуального анализа данных [33, 34, 39, 44].

Перечень подходов и мер противодействия по названным четырем группам сведем к пяти направлениям:

- 1) стандартизация конфигураций и архитектур вместе со стандартизацией действий, обеспечивающих функционирование соответствующих ИТ-решений (см., в частности, комплекс мер, предлагаемых Национальным институтом стандартов США [45, 46]);

- 2) использование принципиально нового качества коммуникации за счет высокоскоростных беспроводных 5G-решений в рамках стандарта 3GPP Релиз 16 (в том числе новые возможности мониторинга инфраструктуры за счёт использования видео-контроля; существенное расширение количества датчиков «на единицу площади», которые можно использовать в режиме процессно-реального времени, замену кабельных решений на мобильные; переход к новым, учитывающим перечисленные нами преимущества, операционным и бизнес-моделям и т.п. [47]);

- 3) особые архитектурные подходы, ориентированные на использование тех или иных проблемно-ориентированных свойств соответствующего варианта ИТ-инфраструктуры (см., в частности, возможности отделения функций маршрутизации и управления трафиком от собственно пересылки сообщений в рамках *SDN/NFV*-подхода [25, 26], а также предлагаемые компанией *Cisco* решения в рамках так называемых *Intent-Based Network Architectures* [27] и т.п.);

- 4) частные¹² 5G-сети, в которых значительный объем управленческих данных (данных мониторинга ИТ-инфраструктуры) можно обрабатывать локально;

- 5) организация и сопровождение специализированных центров анализа инцидентов и поддержки информационной безопасности, предлагаемые лидерами рынка (*IBM, Cisco, Symantec* и др. – см. [48-50]).

Обобщение результатов эксплуатационной практики показывает (см., например, [24]), что наиболее часто используемые подходы в обеспечении безопасности *IoT*-решений – это:

- поддержка приватности коммуникаций (в том числе в части профилирования, позиционирования и трекинга, безопасности передачи данных и др.);
- использование облегченных (*lightweight*) средств криптографической защиты, ориентированных на применение в ограниченных по производительности ресурсов хранения и обработки данных, в узлах *IoT*;
- защита сетевой доступности и собственно пересылки сообщений (защита маршрутизации, изоляция зараженных узлов, стабильность протоколов

¹² Например, в пределах одного предприятия или же его крупного подразделения.

безопасности в части самовосстановления без участия оператора);

- устойчивость по отношению к внешним вредоносным воздействиям (способность устойчиво противостоять попыткам несанкционированного проникновения в сеть, способность идентифицировать атаки на ранних стадиях их развертывания, быстрое восстановление после сбоев и др.), в том числе устойчивость к DoS¹³-атакам и воздействию инсайдеров (ресурсоэффективные средства идентификации и противодействия DoS\DDoS¹⁴-атакам, ресурсоэффективное противодействие активностям инсайдеров и др.).

К сожалению, все эти очевидные технологические достижения не обеспечивают всеобъемлющих решений по управлению рисками, перечисленными в п.п. i-iv. Так, например, обзоры [31, 32] компании *TrendMicro* демонстрируют неустраняемые на текущий момент проблемы в части уязвимости 5G-сетей. Не менее существенные проблемы идентифицированы в рамках проблематики аутентификации (см. п. iii), о которых мы поговорим подробнее в следующем Разделе. Особого анализа (в том числе на уровне разработки эффективных математических моделей и алгоритмов анализа данных и поддержки принятия решений) требует проблематика формирования адекватного «процедурного» ответа на вызовы характерного для Больших Данных эффекта *Open* (см. п. iv), причем такого, который был бы ориентирован на построение надёжных систем мониторинга и управления информационной безопасностью в индустриальных бизнес-решениях в рамках *Индустрии 4.0*.

Подводя промежуточный итог, можно заключить, что на текущий момент средства нападения (в рассматриваемой области *Индустрии 4.0*) имеют очевидные и пока что ничем не компенсируемые преимущества (в части эффективности – результативности при использовании идентичных ресурсов) по отношению к средствам защиты.

Два примера критически значимых вызовов информационной безопасности бизнес-решений *Индустрии 4.0*

Обсудим более подробно две проблемы, важных для успеха всего «проекта» *Индустрии 4.0*, которые остаются на текущий момент без удовлетворительных средств защиты от актуальных угроз информационной безопасности.

1. *Проблема аутентификации участников информационного взаимодействия*. Одним из наиболее уязвимых в настоящий момент элементов в промышленных бизнес-средах *Индустрии 4.0* оказывается процесс аутентификации участников информационного обмена. Фактически речь здесь идет об авторизации статуса (доверенности) и полномочий (акцептованных видов действий) для «акторов» двух типов: (1) персоналий (физических лиц, наделенных определенными ролями и полномочиями, от корректности действий которых зависит нормальность режима функционирования соответствующей бизнес-

системы) и (2) программных модулей\компонентов, генерирующих соответствующие информационные и\или управленческие сообщения.

Своевременное выявление нештатных сообщений и\или способных генерировать нештатные действия объектов – персоналий, программных модулей, виртуальных машин, ... – внутри защищаемого периметра представляет собой хорошо известную на сегодня, но все еще не получившую абсолютно бесспорного решения и техническую, и бизнес-проблему. Иллюстрацией этого утверждения может служить текущая ситуация с так называемыми биометрическими аутентификационными решениями (см., например, использование «цифровых слепков» лица и голоса аутентифицируемого персонажа, задействованные в ряде стран в рамках национальных банковских структур). Доказанная (см., например, [32, 51]) компрометируемость подобных «синтаксических» аутентификационных подходов – сигнал для специалистов по информационной безопасности о необходимости поиска альтернативных решений, способных обеспечить надежную защиту соответствующих ИТ- и бизнес-ландшафтов [35].

2. *Проблема мониторинга и обеспечения непрерывности бизнеса крупных интегрированных структур *Индустрии 4.0**. Речь идет о мониторинге «штатности» текущего режима функционирования информационно-технологической инфраструктуры. Цель такого мониторинга – выявлять АНОМАЛИИ в текущей работе объекта мониторинга (в том числе и заранее неклассифицированные¹⁵!). На «техническом» уровне это ведет к потребности реализовать обучение на заранее найденных и «отработанных» примерах, которые приспособлены к возможностям компьютерного обучения на прецедентах и к использованию проблемно-ориентированных компьютерных «инструментов». Особо важным требованием оказывается возможность порождать (формировать компьютерными средствами) неформальные «объяснения»¹⁶ решений, предлагаемых системой интеллектуального анализа данных. Эти решения характеризуются «бесшовностью» интеграции рассуждений, реализованных в математических моделях и компьютерных системах обучения на прецедентах, с теми способами анализа данных и принятия решений, которые задействованы использующим их экспертом. Такая «бесшовная» интеграция позволяет эксперту не только понять и объяснить рекомендации, предлагаемые ему компьютерной системой интеллектуального анализа данных, но и вполне обоснованно принять на себя ответственность за последствия принятых решений. Представленная совокупность обстоятельств – существенный аргумент в пользу особого статуса ИИ-решений как критически важного инструмента поддержки работоспособности (устойчивости по отношению к внешним вредоносным воздействиям) бизнес-решений, обсуждаемых в рамках *Индустрии 4.0*.

¹³ Denial of Service – Отказ от Обслуживания

¹⁴ Distributed Denial of Service – Распределенный Отказ от Обслуживания

¹⁵ Проявления –*Open*-эффекта, характерного для Больших Данных.

¹⁶ См. Explainable Artificial Intelligence (объясняемый искусственный интеллект) – [44] и др.

Два традиционных вопроса: Кто виноват? и Что делать?

Анализ текущей ситуации в рассматриваемой нами здесь проблематике (с учетом уже накопленного, в первую очередь – европейцами, опыта) позволяет выделить по крайней мере два поучительных обстоятельства:

1) успехи европейских стран в продвижении подходов и решений *Индустрии 4.0* оказались в наибольшей степени заметны в рамках англо-американской модели ее развития. Косвенным подтверждением этому тезису может служить, в частности, оценка по развитости европейских национальных сетей *start-up*'ов в области искусственного интеллекта (как, пожалуй, наиболее обсуждаемого компонента бизнес-решений *Индустрии 4.0*) [44];

2) углубленного анализа заслуживает особенный характер британского позиционирования на европейской «карте» *start-up*'ов в области искусственного интеллекта: это – и «чемпионство» по численности (треть от общего количества), и наличие развитой (причем постоянно поддерживаемой организационно и финансово) бизнес-экосистемы, обеспечивающей анализ НИР\ОКР-рисков, и поиск надежных средств управления этими рисками. Существенным элементом этой инновационной среды стала экосистема так называемых *Eagle-Labs*, которую поддерживает и модерирует в её успешном функционировании банк *Barclays* [44]. (Здесь было бы уместно провести сравнение с позиционированием Сбербанка РФ в аналогичной ситуации – в развитии исследований и разработок в области искусственного интеллекта в Российской Федерации).

В стремлении предложить свою версию ответа на традиционный для нашей страны вопрос «*Что делать?*», обратим внимание на следующие три (достаточно очевидных) обстоятельства.

1. Представляется важным целенаправленно поддерживать поисковые и прикладные НИРы по обсуждаемой тематике¹⁷ (поиск и апробацию того, что способно «*взлететь*»). В том числе – например, по теперь уже успешно зарекомендовавшей себя организационной схеме «*Barclays & Eagle-Labs*» (см. выше в данном Разделе).

2. Следует не забывать о формировании проблемно-ориентированных *тестовых сред* (возможно – в специализированных Центрах Коллективного Пользования) для экспериментальной оценки возможностей перспективных НИР\ОКР-разработок в обсуждаемой нами области науки и технологий.

3. Бесспорным приоритетом должна стать целенаправленная подготовка кадров¹⁸ в данной области (в том числе не только новых студентов и аспирантов, но и переподготовка действующих сотрудников служб информационной безопасности предприятий промышленного сектора, а также переподготовка преподавателей в данной области).

¹⁷ В том числе на стыке проблематики собственно информационной безопасности и активно развивающихся исследований и разработок в области искусственного интеллекта.

¹⁸ В том числе с акцентом на взаимодействие с интенсивно развивающимися смежными областями, в частности – искусственным интеллектом и др.

ЗАКЛЮЧЕНИЕ

Как организовать работу, чтобы потом не было «мучительно больно за бесцельно» растроченные ресурсы – время и деньги? Чему учит (однако, – к сожалению, не всех!) опыт «больших информационно-технологических проектов» последних полутора десятилетий (в частности, на примере крупных банковских модернизационных ИТ-проектов [15])? Ответ достаточно очевиден:

сперва – постановка задачи, далее – четко ориентированные прикладные НИР и поисковые ОКР (до стадии разработки решений-прототипов, обеспечивающих *proof_of_the_concept*¹⁹), и только затем – детальное планирование (ресурсов, сроков, компетенций, ...) и запуск в реализацию «внедренческих» проектов.

А в части перспектив дальнейшего результативного развития планов формирования *Индустрии 4.0*, сформулированных еще 10 лет назад, можно на данный момент (принимая во внимание представленные выше доводы о значительном и ничем не компенсированном превосходстве в результативности средств нападения над средствами защиты бизнес-решений *Индустрии 4.0*), к сожалению, аргументированно констатировать: **все это²⁰ слишком дорого и недостаточно эффективно, чтобы иметь возможность быть надёжно реализованным!**

СПИСОК ЛИТЕРАТУРЫ

1. World Economic Forum (2009). ICT for Economic Growth: A Dynamic Ecosystem Driving the Global Recovery. Annual Meeting Report – World Economic Forum Publishing. – URL: <https://www.ifap.ru/pr/2009/n090910b.pdf>
2. A National Strategic Plan for Advanced Manufacturing. Executive Office of the President, US National Science and Technology Council, 2012. – URL: <https://www.federalregister.gov/documents/2018/02/05/2018-02160/national-strategic-plan-for-advanced-manufacturing>
3. Eight Great Technologies. UK Department for Business, Innovation & Skills, 2012. – URL: <https://www.gov.uk/government/speeches/eight-great-technologies>
4. The New High-Tech Strategy. Innovations for Germany. Federal Ministry of Education and Research of Germany (BMBF), 2014. – URL: <https://www.bmbf.de/en/the-new-high-tech-strategy-2322.html>
5. La Nouvelle France Industrielle. Ministère de l'Économie et des Finances, 2015. – URL: <https://www.economie.gouv.fr/entreprises/nouvelle-france-industrielle>
6. Made in China 2025. The State Council of China, 2015. – URL: <http://isdpc.eu/content/uploads/2018/06/Made-in-China-Backgrounder.pdf>
7. Программа «Цифровая экономика Российской Федерации», утвержденная Распоряжением Пра-

¹⁹ Когда, в частности, все риски разработки идентифицированы, а средства управления ими апробированы.

²⁰ Надежная ИТ-инфраструктура для бизнес-решений в рамках *Индустрии 4.0*.

- вительства РФ от 28 июля 2017 г. № 1632. – URL: <http://static.government.ru/media/files/9gFM4FHj4PsB7915v7yLVuPgu4bvR7M0.pdf>
8. G20 Leaders' Communique (2015). Antalya Summit, 2015-11 – URL: <http://www.g20.utoronto.ca/2015/151116-communicue.pdf>
 9. The National AI R&D Strategic Plan. - National Science and Technology Council, Networking and Information Technology Research and Development Subcommittee. - October 2016. – URL: https://obamawhitehouse.archives.gov/sites/default/files/whitehouse_files/microsites/ostp/NSTC/national_ai_rd_strategic_plan.pdf
 10. A Next Generation Artificial Intelligence Development Plan. – URL: <https://chinacopyrightandmedia.wordpress.com/2017/07/20/a-next-generation-artificial-intelligence-development-plan/>
 11. Materials Genome Initiative for Global Competitiveness. Executive Office of the President, US National Science and Technology Council, 2011. – URL: https://mgi.gov/sites/default/files/documents/materials_genome_initiative-final.pdf
 12. National Robotics Initiative (NRI). US National Science Foundation, 2011; National Robotics Initiative 2.0: Ubiquitous Collaborative Robots (NRI-2.0). US National Science Foundation, 2016. – URL: http://robotunion.ru/files/USA_National_Robotics_Initiative_2.0.pdf
 13. Bell M. Service-Oriented Modeling: Service Analysis, Design, and Architecture. – город: Wiley & Sons, 2008. – 366 p.
 14. Биберштейн Н., Боуз С., Фиамманти М., Джонс К., Ша Р. Компас в мире сервис-ориентированной архитектуры. Ценность для бизнеса, планирование и план развития предприятия. – М: КУДИЦ-ПРЕСС, 2007. – 256 с.
 15. Забежайло М.И. Банковский бизнес в России: индустрия или искусство? – Волгоград: Волгоградское научное издательство, 2009. – 331 с.
 16. Chappell D. A. Enterprise Service Bus. – O'Reilly Media, 2004. – 247 p.
 17. IBM Industry Models for Financial Services. The Information FrameWork (IFW) Overview. – Dublin, Ireland: IBM Financial Services Solution Centre, 2004.
 18. IFW Critical Business Process Models. General Information Manual. – Dublin, Ireland: IBM FSSC, 2004.
 19. IFW Object Models. General Information Manual. – Dublin, Ireland: IBM FSSC, 2004.
 20. Inmon W. H. Building the Data Warehouse. – New York (NY, USA): John Wiley, 1994.
 21. Inmon W. H. Using the Data Warehouse. – New York (NY, USA): John Wiley, 1995.
 22. Катлип Р., Медик Д. DB2: решения по интеграции. – М: Кудлиц-Образ, 2005. – 304 с.
 23. Norris N., Yarow B., Byrne B. Building Service-Oriented Banking Solutions with IBM Banking Industry Models and Rational SDP. (IBM Redpaper). – Armonk, NY: IBM Corporation, 2007. – XYI+236 p.
 24. Hameed S., Khan F. I., Hameed B. Understanding Security Requirements and Challenges in Internet of Things (IoT): A Review // Journal of Computer Networks and Communications. – 2019. – P. 1-14. – URL: <https://www.hindawi.com/journals/jcnc/2019/9629381/>
 25. Limoncelli T. A. OpenFlow: A Radical New Idea in Networking // Communications of the ACM. – 2012. – Vol. 55, № 8. – P. 42–47.
 26. Network function virtualization – ETSI. – URL: <https://www.etsi.org/images/files/ETSITechnologyLeaflets/NetworkFunctionsVirtualization.pdf>
 27. What is intent-based networking? – Cisco. – 2019. – URL: <https://www.cisco.com/c/en/us/solutions/intent-based-networking.html>
 28. Berenson H. et.al. A Critique of ANSI SQL Isolation Levels. Proc. ACM SIGMOD 95, pp. 1-10, San Jose CA, June 1995
 29. Revilak S., O'Neil P., O'Neil E. Precisely Serializable Snapshot Isolation (PSSI) // International Conference on Data Engineering. – Boston (MA 02125, USA): University of Massachusetts at Boston, 2011.
 30. Грушо А.А., Забежайло М.И., Зацаринный А.А., Николаев А.В., Писковский В.О., Сенчило В.В., Судариков И.В., Тимонина Е.Е. Об анализе ошибочных состояний в распределенных вычислительных системах // Системы и средства информатики. – 2018. – №1. – С. 99-109.
 31. From SIMjacking to Bad Decisions. 5G Security Threats to Non-Public Networks. – Trend Micro. – 19 Nov. 2019. – URL: <https://www.trendmicro.com/vinfo/ru/security/news/internet-of-things/from-esim-jacking-to-fake-news-threats-to-5g-and-security-recommendations>
 32. The New Norm. Trend Micro Security Predictions for 2020. – Trend Micro Inc., 2019. – 22 p. – URL: <https://documents.trendmicro.com/assets/rpt/rpt-the-new-norm-trend-micro-security-predictions-for-2020.pdf>
 33. Грушо А.А., Забежайло М.И., Зацаринный А.А., Писковский В.О., Борохов С.В. О возможностях приложений интеллектуального анализа данных в задачах обеспечения информационной безопасности облачных сред // Научно-техническая информация. Сер. 2. – 2015. – № 11. – С. 1-11.
 34. Грушо А.А., Забежайло М.И., Зацаринный А.А., Писковский В.О. О некоторых методах и технологиях искусственного интеллекта, используемых при защите облачных вычислений // Научно-техническая информация. Сер. 2. – 2017. – № 3. – С. 1-15.
 35. О внесении изменений в статью 7 Федерального закона "О противодействии легализации (отмыванию) доходов, полученных преступным путем, и финансированию терроризма. (в части совершенствования регулирования деятельности кредитных организаций по сбору биометрических персональных данных и проведению удаленной биометрической идентификации) – Законопроект №613239-7. – URL: <https://sozd.duma.gov.ru/bill/613239-7>
 36. Norris N., Yarow B., Byrne B. Building Service-Oriented Banking Solutions with IBM

- Banking Industry Models and Rational SDP. (IBM Redpaper). – Armonk, NY: IBM Corporation, 2007. – XYI+236 p.
37. Banking Data Warehouse General Information Manual. The Comprehensive Data Warehousing Offering for Financial Institutions. – Dublin, Ireland: IBM FSSC, 2006.
 38. Huvar M., Falter T., Fiedler T., Zubev A. Developing Applications with Enterprise SOA. Developing Applications with Enterprise SOA. – Galileo Press, Boston et al. – 329 p.
 39. Ghosh A., Chakraborty D., Law A. Artificial Intelligence in Internet of Things // CAAI Trans. Intell. Technol. – 2018. – Vol. 3, Iss. 4. – P. 208–218.
 40. Chen K., Zhang S., Li Z., Zhang Y., Deng Q., Ray S., Jin Y. Internet-of-Things Security and Vulnerabilities: Taxonomy, Challenges, and Practice // Journal of Hardware and Systems Security. – 2018. – Vol. 2. – P. 97–110. – URL: <https://doi.org/10.1007/s41635-017-0029-7>
 41. Chhetri S.R., Faezi S., Rashid N., Al Faruque M.A. Manufacturing Supply Chain and Product Lifecycle Security in the Era of Industry 4.0 // Journal of Hardware and Systems Security. – 2018. – Vol. 2. – P. 51–68. – URL: <https://doi.org/10.1007/s41635-017-0031-0>
 42. Smarter Security for Manufacturing in The Industry 4.0 Era. Industry 4.0 Cyber Resilience for the Manufacturing of the Future. (White paper). – Mountain View (CA, USA): Symantec, 2017. – 11 p. – URL: <https://www.symantec.com/content/dam/symantec/docs/solution-briefs/industry-4.0-en.pdf>
 43. Made in China 2025 (Backgrounder) - Institute for Security & Development Policy. – June 2018 – 9 p. – URL: www.isdp.eu
 44. Kelnar D. The State of the AI 2019: Divergence. – London, UK: MMC Ventures & Barclays. – 2019. – 148 p. – URL: <https://www.stateofai2019.com/>
 45. Stouffer K., Zimmerman T., Tang C.Y., Lubell J., Cichonski J., McCarthy J. Cybersecurity Framework Manufacturing Profile // NISTIR 8183. – Sept. 2017. – 56 p. – URL: <https://doi.org/10.6028/NIST.IR.8183>
 46. Framework for Improving Critical Infrastructure Cybersecurity. Version 1.0. – NIST, Feb. 12, 2014. – 39 p. – URL: <https://www.nist.gov/system/files/documents/cyberframework/cybersecurity-framework-021214.pdf>
 47. Technology, Media, and Telecommunications Predictions 2020. – Deloitte Insights. – Deloitte Dev LLC, 2019. – 133 p. – URL: https://www2.deloitte.com/global/en/insights/industry/technology/technology-media-and-telecom-predictions.html?icid=dcom_promo_featured|global;en
 48. Internet Security Threat Report Internet Report. – Vol. 21 (April 2016). – Mountain View (CA, USA): Symantec, 2016. – 80 p. – URL: <https://www.symantec.com/content/dam/symantec/docs/reorts/istr-21-2016-en.pdf>
 49. Brem W. Communication Structures for Industry 4.0. – Cisco, 2016. – 5 p. – URL: https://www.cisco.com/c/dam/en_us/solutions/trends/iot/docs/communication-structures-so.pdf
 50. IBM X-Force IRIS Cyberattack Preparation and Execution Frameworks. – Armonk (N.Y., USA): IBM, July 2018. – 20 P. – URL: <https://security-intelligence.com/media/ibm-x-force-iris-cyberattack-preparation-and-execution-frameworks/>
 51. Peter's J. Researchers fooled Chinese facial recognition terminals with just a mask // The Verge. – Dec.13, 2019. – URL: <https://www.theverge.com/2019/12/13/21020575/china-facial-recognition-terminals-fooled-3d-mask-kneron-research-fallibility>

Материал поступил в редакцию 21.01.20.

Сведения об авторах

ГРУШО Александр Александрович – доктор физико-математических наук, профессор, заведующий лабораторией Института проблем информатики, Федеральный исследовательский центр «Информатика и управление» Российской академии наук (ФИЦ ИУ РАН), Москва
e-mail: grusho@yandex.ru

ЗАБЕЖАЙЛО Михаил Иванович – доктор физико-математических наук, доцент, заведующий лабораторией Института проблем информатики, ФИЦ ИУ РАН
e-mail: m.zabezhailo@yandex.ru

ПИСКОВСКИЙ Виктор Олегович – кандидат физико-математических наук, старший научный сотрудник Института проблем информатики, ФИЦ ИУ РАН
e-mail: vpvp80@yandex.ru

ТИМОНИНА Елена Евгеньевна – доктор технических наук, профессор, ведущий научный сотрудник Института проблем информатики, ФИЦ ИУ РАН
e-mail: elimon@yandex.ru

Элементы информационной технологии оптических систем идентификации необитаемых подводных аппаратов

Существующие информационные технологии для решения задач идентификации оптических изображений имеют ограничения применимости, связанные с высокой чувствительностью к внешним возмущениям. Рассмотренная модель системы оптимальной фильтрации обеспечивает обнаружение отклонения от эталонного состояния объектов сложной структуры.

Ключевые слова: информационные технологии, идентификация, подводный аппарат, алгоритм

DOI: 10.36535/0548-0027-2020-03-2

ВВЕДЕНИЕ

Существующие информационные технологии для решения задач идентификации оптических изображений имеют ограничения применимости, связанные с высокой чувствительностью к внешним возмущениям. Цель настоящего исследования – реализация основных элементов оптических информационных систем идентификации необитаемых подводных аппаратов.

Структурно системы идентификации и распознавания оптических образов делятся на три типа: 1) разомкнутые системы, корректирующие результат распознавания по внешним условиям; 2) замкнутые системы, системы с обратной связью, корректирующие результат идентификации по ошибке распознавания; 3) комбинированные системы, содержащие разомкнутый и замкнутый контуры обработки информации.

Для реализации поставленной цели использовался алгоритм метода компенсации информационных потоков при оценке информации отклонения J_e по затратам управления в системе стабилизации координат точки корреляционного максимума, имеющий достаточно простую реализацию. Модель системы оптимальной фильтрации обеспечивает обнаружение отклонения от эталонного состояния объектов сложной структуры. При применении рассмотренного метода реально построение системы, инвариантной к возмущениям в пространстве объекта. Подобная система кроме преимуществ имеет существенный недостаток – слабые корреляционные свойства изображения значительно ухудшают качество ее работы [1-3].

УСЛОВИЯ ИНВАРИАНТНОСТИ АЛГОРИТМА ИДЕНТИФИКАЦИИ ОТНОСИТЕЛЬНО ВОЗМУЩЕНИЙ В ПРОСТРАНСТВЕ ОБЪЕКТОВ

Для образов ω , принадлежащих конечному счётному множеству Ω , включающему n элементов $\omega_i \in \Omega$, $i=1, \dots, n$, сформируем на основе знаний об этих образах множество эталонов ω^* , принадлежащих конечному счётному множеству эталонов Ω^* , включающему n элементов $\omega_i^* \in \Omega^*$, $i=1, \dots, n$. Вопрос формирования эталонов решается либо описанием свойств объекта, либо формированием признаков или созданием набора тестовых изображений – эталонов.

Определим внешние возмущения как объекты и искажения в пространстве образов. Предполагая, что эти возмущения известны, рассмотрим множество возмущений $g_j \in G$, $j=1, \dots, m$. Причём если число элементов в Ω и Ω^* известно, то число элементов в множестве возмущений известно не всегда.

Задача идентификации оптических образов определяет операцию соотнесения входного элемента элементу из множества эталонов и формирование логического высказывания о принадлежности эталона и объекта. Определим для каждого из элементов множества объектов логическую функцию $y_i \in Y$. Естественно $\cup y_i=1$, $i=1, \dots, n$.

Рассмотрим введенное ранее информационное пространство с нормой

$$\|\omega\| = I_\omega \quad (1)$$

и метрикой, определяющей расстояние между образами объекта и эталона

$$a(\omega, \omega^*) = I_{\omega/\omega^*}. \quad (2)$$

Используем понятие оператора идентификации как отображение из информационного пространства I на множество логических функций Y

$$A: I_{\omega} \rightarrow y_{\omega}. \quad (3)$$

Основные свойства оператора определяются сутью задачи идентификации и перечисленными далее свойствами информационного пространства:

1) нулевая информация отображается в пустое множество – логический ноль

$$A: (I_{\omega} = 0) \rightarrow \emptyset. \quad (4)$$

2) оператору идентификации A соответствует обратный оператор генерации образа A^{-1}

$$\begin{aligned} A^{-1}: y_{\omega} &\rightarrow I_{\omega}; \\ A^{-1}(A: I_{\omega}) &= I_{\omega}. \end{aligned} \quad (5)$$

3) при формировании возмущения g как элемента пространства возмущений, объединяющегося с образом, оператор идентификации формирует высказывание, содержащее информацию об образе и возмущении

$$A: I_{g\omega} \rightarrow y_{g\omega}, \quad (6)$$

при этом высказывания y_{ω} и y_g связаны, в общем случае, дизъюнкцией

$$A: I_{g\omega} \rightarrow y_{g\omega} = y_{\omega} \cup y_g. \quad (7)$$

4) для числа элементов в пространствах образов n , эталонов n и возмущений m количество переменных k функции решения для парного сочетания объекта и возмущения равно

$$k = \log_2(mn) \quad (8)$$

и, как следствие, для неопределённого m не определён порядок функции решения и возможно только неточное описание решения, следовательно, обратный оператор не обеспечивает точного восстановления информации об образе [4–6].

На этом основании формируется важное свойство: ошибка идентификации нарушает взаимную однозначность оператора идентификации. Так, если

$$\begin{aligned} A: I_{\omega} &\rightarrow y_{\omega}; \\ A: I_{\omega^*} &\rightarrow y_{\omega^*} \end{aligned} \quad (9)$$

и расстояние между образом и эталоном в информационном пространстве не равно нулю

$$I_{\omega/\omega^*} \neq 0, \quad (10)$$

то обратное преобразование восстанавливает информацию с ошибкой

$$A^{-1}: y_{\omega} \neq A^{-1}: y_{\omega^*}. \quad (11)$$

Собственно при такой постановке математического ансамбля, оператор идентификации оптических изображений описывает любую процедуру идентификации объекта по отношению к эталонам как знаниям об объектах.

Естественно возникает задача построения оптимального оператора идентификации

$$\begin{aligned} A^*: I_{g\omega_i} &\rightarrow y_{\omega_i}; \\ A^*: I_{\omega_i} &\rightarrow y_{\omega_i^*}. \end{aligned} \quad (12)$$

При этом выполняется условие отсутствия ошибки идентификации для всех объектов

$$\begin{aligned} A^{-1*}: y_{\omega_i} &\rightarrow I_{\omega_i}; \\ A^{-1*}: y_{\omega_i^*} &\rightarrow I_{\omega_i}; \\ I_{\omega_i/\omega_i^*} &= 0; \\ i &= \overline{1, n}, \end{aligned} \quad (13)$$

которое возможно только при определённости размерности логического пространства, а это, к сожалению, возможно исключительно для стационарного случая; при условии неопределённости k оптимальный оператор не идентифицируется и приводит к локализации области решения задачи.

Поэтому для построения оптимального оператора идентификации, обеспечивающего отсутствие ошибок, необходимо полное знание об объектах и возмущениях. Если знания не полны или система не стационарна, найти строго оптимальный алгоритм идентификации невозможно [7]. Это условие известно в теории управления как условие инвариантности или полной независимости от возмущений.

Из свойства оператора отображать результаты в пространство логических функций k переменных следует возможность коррекции результатов классификации $(y_{\omega} \cup y_g) \cap y_{\omega} = y_{\omega}$ [8–10].

Рассмотрим структуру системы идентификации с контуром коррекции (рис. 1).

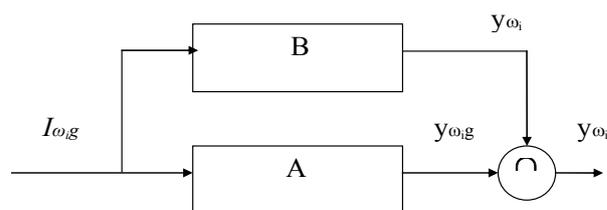


Рис. 1. Коррекция операции принятия решения

В этом случае B является оператором идентификации, генерирующим гипотезу

$$B: I_{g\omega} \rightarrow y_{\omega}. \quad (14)$$

Если гипотеза верна, то и результат верен, если гипотеза ошибочна, то результат ошибочен. Такой подход дает возможность представить процедуру перебора гипотез, но при этом позволяет реализовать строго оптимальную процедуру только для стационарной задачи.

Возможно построение корректирующей процедуры по входной информации. Приведенная на рис. 2 схема включает датчик D , чувствительный к возмущениям. В этом случае появляется необходимость изучать возмущения, и задача суживается до стационарного случая. Таким образом, схема коррекции в пространстве решения более перспективна, так как не требует знания возмущений. Достоинством процедур коррекции является их принципиальная сходимость и простота реализации, свойственная управлению по возмущению. Однако эти достоинства реализуются только в стационарных случаях.

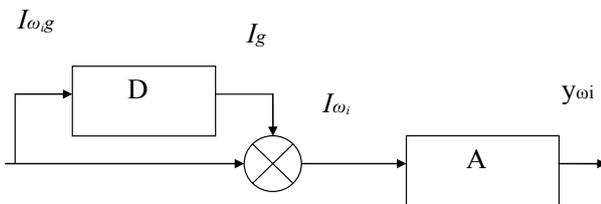


Рис. 2. Коррекция операции принятия решения по входной информации

Дополнив блок корректировки, отраженный на рис. 2, операцией проверки правильности идентификации, получаем структуру с двумя контурами: разомкнутый контур прогноза и замкнутый контур проверки гипотезы (рис. 3).

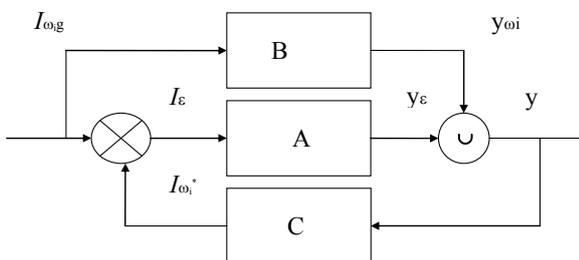


Рис. 3. Структура системы с контуром прогноза и контуром проверки гипотезы

Полученный вариант удовлетворяет условию структурной инвариантности – критерий двухканальности Попова [5-7]. Исходя из условия инвариантности к возмущениям, как условия возможности

построения строго оптимальной системы, определим операторы системы.

Так как оператор B должен удовлетворять условию (14) то это оптимальный оператор идентификации, построенный на априорных данных системы $B=A^*$. При этом к оператору A условия оптимальности не предъявляются, достаточно выполнения условия (4), т. е. это исходная процедура, построенная без оптимизации к пространству образов и возмущений.

На основе условия инвариантности в момент $t=T$ при условии $y_{\epsilon}=0$, $y=y_{\omega} \cup y_{\epsilon}$ оператор в обратной связи определяется как обратный оператору в контуре прогноза [3, 9]:

$$\begin{aligned} C &= A^{*-1}; \\ A^{*-1}: y_{\omega i} &\rightarrow I_{g\omega i}; \\ I_{\omega^*} &= I_{g\omega i}. \end{aligned} \quad (15)$$

Таким образом, входная информация равна информации эталона $I_{\omega^*}=I_{g\omega i}$ и, следовательно, расстояние между эталоном и объектом равно нулю: $I_{\omega}=0$. При неполной информации о возмущениях оператор A^* не оптимален и, естественно, эталон будет восстанавливаться с ошибкой.

Так для оператора A имеем

$$\begin{aligned} I_{\epsilon} &= I_{\omega g/\omega^*}; \\ A: I_{\omega g/\omega^*} &\rightarrow y_{\epsilon}. \end{aligned} \quad (16)$$

Тогда:

$$\begin{aligned} y &= y_{\omega i} \cup y_{\epsilon}; \\ A^{*-1}: y &\rightarrow I_{\omega i \epsilon} = I_{\omega^*}. \end{aligned} \quad (17)$$

Следовательно, отклонение входной информации от информации в контуре обратной связи:

$$I_{\epsilon} = I_{\omega_i g/\omega^*} = I_{\omega_i g/\omega_i \epsilon}. \quad (18)$$

Так как образ ошибки, в общем случае, не соответствует образу возмущения, то и выходная компонента решения y_{ϵ} отлична от нуля. Действительно, представим, что образ ошибки сходится к образу возмущения $\omega_{\epsilon} \rightarrow \omega_g$, но формируется ω_{ϵ} при ненулевом расстоянии между входным образом и образом эталона и, естественно, ошибка решения y_{ϵ} может быть сколь угодно мала, но принципиально существует. Она известна как ошибка статизма и легко устраняется за счёт накопления информации в контуре обратной связи [11].

Предположив сходимость процедуры идентификации, что естественно для оптимального оператора, получаем последовательность:

$$I_{\epsilon} = I_{\omega_i g/\omega_i \epsilon 1} > I_{\omega_i g/\omega_i \epsilon 2} > \dots > I_{\omega_i g/\omega_i \epsilon k}. \quad (19)$$

Введём в контур обратной связи процедуру накопления информации. Оценив его операцией усреднения, или в предположении не смещённости – операцией определения математического ожидания, получаем:

$$I_{\omega^*} = I_{\omega_i g / \omega_i \varepsilon r} + M \{ I_{\omega_i g / \omega_i \varepsilon} \}. \quad (20)$$

Учитывая, что математическое ожидание информации является энтропией, можно записать:

$$\begin{aligned} I_{\omega^*} &= I_{\omega_i g / \omega_i \varepsilon r} + H_{\omega_i g / \omega_i \varepsilon}; \\ I_{\omega_i g / \omega_i \varepsilon r} &\rightarrow I_{\omega_i}; \\ H_{\omega_i g / \omega_i \varepsilon} &\rightarrow I_g. \end{aligned} \quad (21)$$

Получается, что контур обратной связи инвариантной системы идентификации должен строиться как ассоциативная память с накоплением информации о возмущениях. Следовательно, для того, чтобы система идентификации оптических образов могла эффективно работать при возмущениях в пространстве объектов, необходимо выполнение определенных условий.

1. Структурное условие инвариантности, заключающееся в том, что структура системы идентификации должна включать:

- выбранный оператор идентификации A ;
- канал прогноза, реализованный как оптимальный оператор $B=A^*$;
- контур обратной связи, реализованный как ассоциативная память $C=A^{*-1}$ с накоплением информации о возмущениях.

2. Параметрическое условие инвариантности, заключающееся в том, что для принятия решения за конечное время информационная мощность контура обратной связи должна превышать информационную мощность источника:

$$H_{\omega^*} > H_{\omega_i g}. \quad (22)$$

При выполнении перечисленных условий существуют операторы, и возможно построить систему, работающую без ошибок на известном алфавите объектов, не зависимо от возмущений. Если эти условия нарушены, то возмущения будут снижать качество работы системы.

АЛГОРИТМ УПРАВЛЕНИЯ ФОРМИРОВАНИЕМ ЭТАЛОННОГО ИЗОБРАЖЕНИЯ

При реализации алгоритма основное значение имеет динамика системы управления эталоном. Так как процедура компенсации предполагает оптимальность совмещения, рассмотрим систему как динамический объект, эволюция которого во времени определяется свойствами алгоритма совмещения. В таком случае получаем схему (рис. 4), в которой сигнал источника входного изображения компенсируется в процессе свободного движения обобщенного объекта управления, содержащего генератор эталонного изображения и алгоритм совмещения изображений.

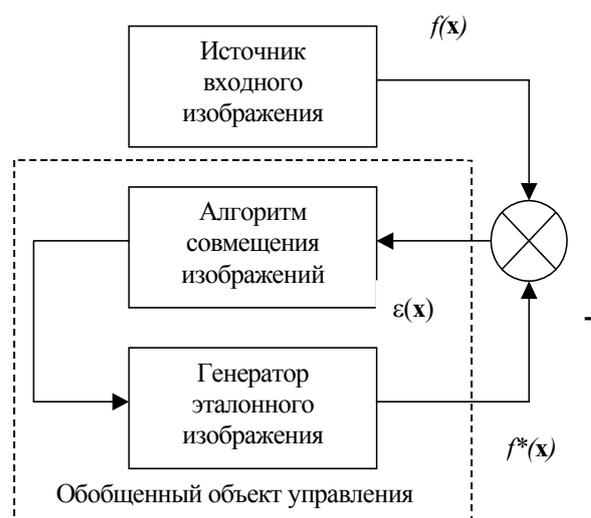


Рис. 4. Схема построения обобщенного объекта управления

При таком подходе критерий качества формируется как функционал цели [12]. Предполагая сильный оптимум, выбираем для измерения расстояния равномерную метрику:

$$\rho(f, f^*) = \sup |f(\mathbf{x}) - f^*(\mathbf{x})|, \quad (23)$$

где \mathbf{x} – базовый набор показателей модели.

Так как в основе процедуры совмещения находится градиентный алгоритм, рассматриваем движение градиентного алгоритма оптимизации для целевой функции ϕ при единичном шаге по времени при движении по антиградиенту:

$$\mathbf{x}_{m+1} - \mathbf{x}_n = -\alpha \nabla \phi(\mathbf{x}), \quad (24)$$

однако при шаге, отличном от единичного, алгоритм имеет вид:

$$\mathbf{x}_{m+1} - \mathbf{x}_n = -\alpha \nabla \phi(\mathbf{x}) \Delta t. \quad (25)$$

Легко получить исходное уравнение градиентной процедуры:

$$\begin{aligned} \frac{\mathbf{x}_{m+1} - \mathbf{x}_n}{\Delta t} &\approx -\alpha \nabla \phi(\mathbf{x}); \\ \lim_{\Delta t \rightarrow 0} \frac{\mathbf{x}_{m+1} - \mathbf{x}_n}{\Delta t} &= -\alpha \nabla \phi(\mathbf{x}). \end{aligned} \quad (26)$$

Таким образом, исходное уравнение метода имеет вид:

$$\frac{d\mathbf{x}}{dt} = -\alpha \frac{\partial \phi}{\partial \mathbf{x}}. \quad (27)$$

Сопоставим данное уравнение с уравнением динамической системы в форме Коши

$$\frac{d\mathbf{x}}{dt} = \Psi(\mathbf{x}, \mathbf{u}). \quad (28)$$

где \mathbf{u} – фактор оптимальности.

Для того чтобы обе системы двигались с одинаковой скоростью, необходимо выполнение условия:

$$\Psi(\mathbf{x}, \mathbf{u}) = -\alpha \frac{\partial \phi}{\partial \mathbf{x}}. \quad (29)$$

В этом случае исследование сходимости алгоритма можно свести к анализу устойчивости соответствующей динамической системы.

Рассматривая динамическую систему с гладкой правой частью:

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{u}, t), \quad (30)$$

определим векторное поле правой части как потенциальное:

$$\mathbf{f}(\mathbf{x}, \mathbf{u}, t) = \nabla \phi(\mathbf{x}, \mathbf{u}, t). \quad (31)$$

Переходя к оценке производной:

$$\frac{dx}{dt} \approx \frac{x(t + \Delta t) - x(t)}{\Delta t}, \quad (32)$$

обозначим $x(t + \Delta t) = x_{i+1}$, $x(t) = x_i$ и для $k = \text{const}$ обозначим произведения шага квантования по времени на константу как $\Delta t * k = \alpha$. Тогда получаем:

$$\mathbf{x}_{i+1} = \mathbf{x}_i + \alpha \nabla \phi(\mathbf{x}_i, \mathbf{u}_i, t). \quad (33)$$

Такое описание градиентной процедуры как движения динамической системы позволяет перейти от анализа сходимости к анализу устойчивости динамической системы (12).

При линейризации правой части уравнения получаем:

$$\dot{\mathbf{x}} = \frac{\partial^2 \phi}{\partial^2 \mathbf{x}} \mathbf{x} + \frac{\partial^2 \phi}{\partial \mathbf{x} \partial \mathbf{u}} \mathbf{u}. \quad (34)$$

Таким образом, в начальной точке линейризация имеет вид:

$$\begin{aligned} \dot{\mathbf{x}} &= A\mathbf{x} + B\mathbf{u}; \\ A &= \frac{\partial^2 \phi}{\partial^2 \mathbf{x}} \Big|_{\substack{\mathbf{x}=\mathbf{x}_0 \\ \mathbf{u}=\mathbf{u}_0}}; \\ B &= \frac{\partial^2 \phi}{\partial \mathbf{x} \partial \mathbf{u}} \Big|_{\substack{\mathbf{x}=\mathbf{x}_0 \\ \mathbf{u}=\mathbf{u}_0}}. \end{aligned} \quad (35)$$

Выбор функции ϕ определяется видом целевой функции в задаче компенсации поля входного изображения.

Для построения критерия совмещения использовано проективное преобразование или модифицированное преобразование Радона [7-9], что приводит к упрощению процесса оптимизации вычисления

$$\begin{aligned} F_x(y) &= \int_{x_{\min}}^{x_{\max}} f(x, y) dx; \\ F_y(x) &= \int_{y_{\min}}^{y_{\max}} f(x, y) dy. \end{aligned} \quad (36)$$

При этом, для сохранения знакоположительности компенсации позитива изображения F ведется негативом эталона F^* , и суммарное изображение стремится к уровню максимальной яркости F_{\max} исходного изображения. Необходимо помнить о недопустимости перекомпенсации, что отображается введением ограничения факторных значений критерийности. Таким образом, сравнение с эталоном сводится к оптимизационной задаче:

$$\begin{aligned} \mathbf{u}^* &\xrightarrow{a=a^*} \min \{ \phi(\mathbf{u}, a) \}; \\ a^* &\rightarrow \min \sum_1^n |\mathbf{u}|; \\ F_x(y) &\leq F_{\max}; \quad F_y(x) \leq F_{\max}. \end{aligned} \quad (37)$$

Целевая функция определена как суммарное отклонение поля эталона от поля изображения при $A = A(\mathbf{u})$.

$$\begin{aligned} \phi(\mathbf{x}, \mathbf{x}_0, A) &= \\ &= \left[\int_{y_{\min}}^{y_{\max}} \left(F_{\max} - \left| \int_{x_{\min}}^{x_{\max}} f(\mathbf{x}) - f^*(A\mathbf{x} + \mathbf{x}_0) dx \right| dy + \right. \right. \\ &\quad \left. \left. + \int_{x_{\min}}^{x_{\max}} \left(F_{\max} - \left| \int_{y_{\min}}^{y_{\max}} f(\mathbf{x}) - f^*(A\mathbf{x} + \mathbf{x}_0) dy \right| dx \right) \right]. \end{aligned} \quad (38)$$

При этом правая часть уравнения динамики имеет вид:

$$\begin{aligned} \frac{\partial \phi}{\partial x} &= \frac{\partial \int_{x_{\min}}^{x_{\max}} (F_{\max} - |\Delta F_y(x)|) dx}{\partial x}; \\ \frac{\partial \phi}{\partial y} &= \frac{\partial \int_{y_{\min}}^{y_{\max}} (F_{\max} - |\Delta F_x(y)|) dy}{\partial y}. \end{aligned} \quad (39)$$

Следовательно, если учитывать неотрицательность ΔF , оценку динамики можно выполнить по уравнению

$$\begin{aligned}\frac{dx}{dt} &= \Delta F_y(x); \\ \frac{dy}{dt} &= \Delta F_x(y).\end{aligned}\quad (40)$$

Анализ условий сходимости процедуры компенсации входного сигнала связан с анализом устойчивости дифференциального уравнения (18). Причем простая градиентная процедура описывается свободным движением системы, а модификации градиентной процедуры будут содержать управляемую компоненту.

Однако уравнение (40) в общем случае – нелинейное дифференциальное уравнение и вопрос о его устойчивости требует дополнительного исследования.

ЗАКЛЮЧЕНИЕ

Таким образом, можно сделать следующие выводы: для достижения инвариантности информационной технологии идентификации оптических образов при возмущениях в пространстве объектов необходимо выполнение следующих условий. Структурное условие инвариантности, заключающееся в том, что информационная технология идентификации должна включать: выбранный оператор идентификации A ; канал прогноза, реализованный как оптимальный оператор $B=A^*$; контур обратной связи, реализованный как ассоциативная память $C=A^{*-1}$ с накоплением информации о возмущениях. Параметрическое условие инвариантности, заключающееся в том, что для принятия решения за конечное время информационная мощность контура обратной связи должна превышать информационную мощность источника.

СПИСОК ЛИТЕРАТУРЫ

1. Катыс Г.П. Автоматический обзор и поиск в оптическом диапазоне. – М.: Наука, 1966. – 159 с.
2. Шибанов Г.П. Распознавание в системах автоконтроля. – М.: Машиностроение, 1973. – 424 с.
3. Васильев И.В. Распознающие системы: справочник. – изд. 2-е, перераб. и доп. – Киев: Наукова думка, 1983. – 423 с.

4. Вапник В.Н., Червоненкис А.Я. Теория распознавания образов. – М.: Наука, 1974. – 416 с.
5. Дуда Р., Харт П. Распознавание образов и анализ сцен. – М.: Мир, 1976. – 511 с.
6. Автоматический анализ сложных изображений: сборник / пер. под ред. Э.М. Бравермана. – М.: Наука, 1969. – 310 с.
7. Файн В.С. Оpozнание изображений. – М.: Наука, 1970. – 299 с.
8. Катыс Г.П. Автоматическое сканирование. – М.: Машиностроение, 1969. – 520 с.
9. Горелик А.Л., Скрипник В.А. Методы распознавания: учеб. пособие для вузов. – М.: Высшая школа, 1977. – 222 с.
10. Ту Дж., Гонсалес Р. Принципы распознавания образов / пер. с англ. И.Б. Гуревича; под ред. Ю.И. Журавлёва. – М.: Мир, 1978. – 416 с.
11. Шестов Н.С. Выделение оптических сигналов на фоне случайных помех / под ред. ак. А.А. Лебедева. – М.: Советское радио, 1967. – 348 с.
12. Zhilenkov A., Nyrkov A., Chernyi S., Sokolov S. Simulation of in-sensor processes in the sensor – Object system type when scanning the elements of underwater communication lines with a probe beam // International Review on Modelling and Simulations. –2017. – Vol. 10(5). – P. 363-370. DOI: 10.15866/iremos.v10i5.11691.

Материал поступил в редакцию 11.12.19.

Сведения об авторах

ЧЕРНЫЙ Сергей Григорьевич – кандидат технических наук, доцент, заведующий кафедрой электрооборудования судов и автоматизации производства Керченского государственного морского технологического университета, г. Керчь; доцент кафедры комплексного обеспечения информационной безопасности, Государственный университет морского и речного флота имени адмирала С.О. Макарова, Санкт-Петербург
e-mail: sergiiblack@gmail.com

ДОРОВСКОЙ Владимир Алексеевич – доктор технических наук, профессор кафедры электрооборудования судов и автоматизации производства Керченского государственного морского технологического университета, г. Керчь
e-mail: dora1943@mail.ru

АВТОМАТИЗАЦИЯ ОБРАБОТКИ ТЕКСТА

УДК 81'322.4:81'37

А.А. Хорошилов, А.В. Кан, Ю.В. Никитин, Ал-др А. Хорошилов

Машинный фразеологический перевод научно-технических текстов на основе модели обобщенных синтагм*

Предлагается новое решение проблемы фразеологического машинного перевода текстов на основе технологии «семантического переводческого конструктора», базирующейся на концепции обобщенных синтагм. В основу этой технологии положено утверждение, что синтаксическую структуру предложения или его фрагментов (синтаксических конструкций) можно представить в виде последовательности обобщенных синтагм, отражающих грамматические свойства конкретных текстовых слов. Концепция обобщенных синтагм позволяет трансформировать исходное предложение в совокупность шаблонов синтаксических конструкций, сопровождаемых их «стыковочными узлами», и тогда процесс перевода будет выполняться путем «разборки» исходного предложения на шаблоны синтаксических конструкций, соотнесения этих исходных шаблонов с шаблонами их переводных эквивалентов и окончательной «сборки» переводного предложения в соответствии с «технологическими инструкциями», заложенными в «правилах сборки». Предложены механизмы автоматического формирования шаблонного представления иерархии синтаксических конструкций двуязычных предложений.

Ключевые слова: фразеологический машинный перевод текстов, семантический переводческий конструктор, концепция обобщенных синтагм, синтаксическая структура предложения, шаблоны синтаксических конструкций, автоматическое формирование шаблонов переводных соответствий

DOI: 10.36535/0548-0027-2020-03-3

ОСНОВНЫЕ ПОДХОДЫ К РЕШЕНИЮ ЗАДАЧИ МАШИННОГО ПЕРЕВОДА

В настоящее время современные промышленные системы машинного перевода используют, в основном, два подхода: традиционный (семантико-синтаксический, основанный на правилах, *rule-based machine translation*) и статистический (основанный на статистической обработке словарных баз, *statistical based machine translation*) [1]. Традиционный метод машинного перевода используется большинством разработчиков систем перевода (примером сервиса, который основан на этой технологии, является <http://translate.promt.ru>). Работа программы, в основе

которой лежит этот метод, состоит из нескольких этапов и, по сути, заключается в использовании лингвистических правил (алгоритмов). Соответственно, создание такого электронного переводчика включает разработку грамматических правил для входного и выходного языков и автоматизированное формирование словарных баз системы. От разработки необходимых лингвистических алгоритмов зависит качество перевода на выходе. Богатый словарь системы также позволяет справиться с переводом самых разнообразных по тематике текстов.

В основе статистического подхода (например, этот подход применяется сервисами <https://www.google.com>, <https://translate.yandex.ru/translator/English-Russian>) лежат математические методы получения перевода. Точнее сказать, весь принцип работы подобной системы основан на статистическом вычислении вероятности совпадений фраз исходного текста с фразами, которые хранятся в базе системы перевода. На сегодняшний

* Статья подготовлена при поддержке гранта РФФИ (проект 18-37-00110 мол_a – «Разработка метода быстрого создания "новых" направлений машинного перевода в условиях невозможности использования корпусов параллельных текстов»)

ний день самыми продуктивными разработками в этой области занимается компания *Google*, которая предлагает онлайн-сервис перевода на своем портале. *Google* заявляет о преимуществе именно статистического метода перевода и видит причину успеха в максимально обширных словарных базах, имеющих необходимое количество сегментов текста для их правильного перевода (при этом представители *Google* умалчивают о роли купленных технологий старшей традиционной системы перевода *Systran*).

При статистическом подходе для работы системы необходимы огромные базы параллельных текстов, где попарно хранятся словосочетания (фразы из 2–3 слов) и их переводы, так называемые *N*-граммы. В процессе перевода также используется механизм анализа, но не лингвистический, а статистический. Система подбирает вариант перевода, основываясь на частоте совпадений, т.е. в конечном итоге будет подставлен вариант, имеющий наиболее высокий процент совпадений. Существенная проблема этого метода состоит в том, что для корректной работы такой системы необходимо иметь в базе не просто очень большое, а невероятно большое количество параллельных *N*-грамм. А так как тексты в различных предметных областях отличаются лексическим составом и синтаксической структурой, необходимо располагать огромными массивами тематических параллельных текстов, что существенно ограничивает сферу применения этого метода.

Существует еще ряд подходов к решению проблемы машинного перевода, базирующихся на использовании параллельных двуязычных текстов. К ним можно отнести технологии *Translation Memory* [2] и широко разрекламированные нейросетевые технологии [3, 4], предполагающие предварительное машинное обучение систем. При этом возникает проблема создания огромных объемов размеченных корпусов текстов, необходимых для обучения систем.

Между тем, для решения задачи разработки высококачественного машинного перевода необходимо, прежде всего, глубокое понимание сложности процессов автоматического перевода текстов с одних языков на другие. Такие процессы базируются на анализе смысловой структуры исходного текста и передаче его содержания средствами выходного языка. При этом переводятся не столько значения слов и их последовательностей, как это делается в статистических и нейросетевых системах машинного перевода, сколько мыслительные образы, порождаемые в сознании человека под их воздействием.

Именно понимание всей сложности процесса перевода текстов с одного естественного языка на другой положено в основу метода фразеологического машинного перевода (ФМП) [5–7]. Этот метод использует преимущества вышеупомянутых подходов – традиционного подхода с его методами семантико-синтаксического анализа и синтеза исходных и выходных текстов и статистического подхода, использующего закономерности построения исходных и переводных текстов при создании декларативных средств. При этом ФМП базируется на современной концепции фразеологического машинного перевода, основная идея которого заключается в использова-

нии в качестве основных единиц смысла устойчивых наименований понятий. Эта концепция наилучшим образом обеспечивает основную идею машинного перевода – адекватную передачу смыслового содержания текста на входном языке средствами выходного языка. Процесс фразеологического машинного перевода включает следующие основные этапы: 1) формализацию смысловой структуры исходного текста; 2) соотнесение слов и словосочетаний исходного текста с их переводными соответствиями; 3) генерацию выходного текста средствами синтаксического и морфологического синтеза.

Необходимо отметить, что при этом подходе перевод текстов выполняется по достаточно длинной технологической схеме, включающей такие сложные операции, как морфологический и синтаксический анализ и синтез исходного и выходного текстов, а также ресурсозатратную операцию трансфера – соотнесение наименований понятий исходного и выходного языков, выполняемую путем поиска переводных соответствий в комплексе многомиллионных двуязычных словарей.

Современные промышленные системы ФМП реализованы в соответствии с описанным выше технологическим процессом перевода текстов. В настоящей статье предлагается новая модель функционирования систем фразеологического машинного перевода, использующая концепцию обобщенных синтагм. В рамках этой модели представлен механизм автоматического формирования шаблонного представления иерархии синтаксических конструкций двуязычных предложений. Кратко рассмотрим процедуры семантико-синтаксического анализа и синтеза русского и английского языков, основанных на концепции обобщенных синтагм¹.

СИНТАКСИЧЕСКИЙ АНАЛИЗ ТЕКСТОВ НА ОСНОВЕ ОБОБЩЕННЫХ СИНТАГМ

Как ранее было показано в работах [7–9] концепция фразеологического концептуального анализа текстов базируется на широком применении принципов лингвистической аналогии. На их основе была разработана система флективных классов слов русского языка, в рамках которой были созданы новые классы слов, *базирующиеся на одинаковых наборах грамматических признаков, соответствующие их формам представления в сходных контекстных окружениях*. Двухбайтовое представление этих классов слов мы назвали обобщенными синтагмами [9]. Такое представление структур слов позволило решить ряд задач автоматической обработки текстовой информации. Например, выявление в текстах наименований понятий, выполняющих сходные синтаксические функции, разрешение грамматической и лексической омонимии слов в русских и английских

¹ Предлагаемое новое решение проблемы машинного перевода выполнено на материальной базе и при поддержке руководства отдела №62 Федерального исследовательского центра «Информатика и управление» РАН и кафедры №805 «Математическая кибернетика» НИУ Московского авиационного института (национального исследовательского университета).

текстах, а также обеспечение возможности представления синтаксической структуры тестов

При разработке синтаксической модели текстов мы исходили из следующей гипотезы: *одинаковым последовательностям обобщенных символов классов слов (обобщенным синтагмам) должны соответствовать одинаковые синтаксические структуры*. При этом предполагалось, что такая гипотеза верна для любых синтаксических моделей и может быть полезной при решении как глобальных, так и частных задач синтаксического анализа и синтеза разноязычных текстов.

В разработанной синтаксической модели текстов на основе использования механизма обобщенных синтагм была разработана иерархия синтаксических конструкций предложений, обеспечивающая при синтаксическом анализе возможность адекватного построения синтаксических структур предложений любой сложности. Эта модель решает задачу построения синтаксической структуры предложения в рамках следующего утверждения: *представление синтаксической структуры текстов в виде последовательности контактно расположенных двухбайтовых индексов обобщенных синтагм, обладающих грамматическими свойствами конкретных слов-эталонов, позволяет фиксировать грамматические и синтаксические свойства различных отрезков реальных текстов, а также дает возможность в ряде задач распознавать аналогичные по заданным свойствам отрезки текстов* [7, 9].

Для решения задачи синтаксического анализа русских и английских текстов в качестве исходных данных использовалась последовательность слов предложения с назначенной грамматической информацией. Результатом анализа являлась иерархия формализованных структур глобальных и локальных конструкций с указанием их местоположения в предложении и «стыковочными узлами», обеспечивающими их смысловую связь с контекстом предложения [9].

В табл. 1 приведены результаты поэтапного преобразования предложения в формализованное модифицированное (поисковое) представление его синтаксической структуры, построенной на основе обобщенных синтагм. В разделе 1 табл. 4 показано текстовое представление с нумерацией каждого слова предложения. В разделе 2 – результаты преобразования предложения в его формальное представление в виде символов грамматических и символов обобщенных классов слов. В разделе 3 отражено формализованное представление глобальных и локальных конструкций предложения. В разделе 4 – преобразование текстового представления локальных конструкций (слов и словосочетаний) предложения в их формальные синтаксические структуры. В разделе 5 – формальная семантико-синтаксическая структура предложения, полученная на основе обобщенных синтагм, в которой вся информация позиционно разделена на блоки грамматической, синтаксической и семантической информации. Блоки информации в словарной статье представлены в позиционной форме и разделяются идентификаторами определенного вида. Каждому блоку предшествует знак «=» (равно), информация внутри блоков разделяется знаком «<»

(тире) или знаком «#» (решетка), однородная информация следует через запятую. Блоки информации расположены в следующем порядке:

1) поисковая синтагма предложения (модифицированный «скелет» предложения);

2) формализованное текстовое представление «скелета» предложения;

3) поисковая синтагма предложения с указанием позиции главного слова словосочетания в предложении и его идентификатора;

4) дерево зависимостей «скелета», представленного в символах обобщенных синтагм;

5) модифицированный «скелет» предложения в символах классов слов с указанием дерева зависимостей (в терминах «хозяин-слуга») в словосочетании, набора однозначной грамматической информации (род, число, падеж, лицо) для каждого слова словосочетания и количество слов «скелета»;

6) число словосочетаний в предложении;

7) информация о каждом словосочетании предложения: структура словосочетаний в символах классов слов и символах обобщенных синтагм с указанием позиции первого слова в предложении (для разрывных конструкций через запятую указан номер позиции второй части словосочетания), позиции главного слова в предложении и в словосочетании и его длины;

8) информация о конструкции предикатно-актантной структуры (ПАС) в виде символов обобщенных синтагм словосочетаний и их позиций в предложении (позиция первого слова словосочетания в предложении и его длина);

9) текстовое представление ПАС;

10) структура эталонного предложения в символах обобщенных синтагм, символах классов слов и длины предложения;

11) эталонное предложение в текстовом представлении.

На основе результатов преобразования предложения в его формализованное модифицированное (поисковое) представление можно автоматически построить словари эталонных семантико-синтаксических конструкций для анализа структурного сходства предложений английских текстов. Как и в случае построения таких словарей для русских текстов эти словари должны в своем составе иметь эталонные словари конструкций семантико-синтаксических структур английских предложений – (*English Sentences – SentEng*) и структур английских словосочетаний (*English Word Combination – WCEng*).

В соответствии с рассмотренным процессом получения синтаксической структуры конструкций английских предложений на основе упрощенного синтаксического анализа были разработаны алгоритмы формирования словарей: эталонного словаря предикатно-активных структур предложений (*PSOtEng*), эталонного словаря структур «скелетов» предложений (*SkEng*), эталонного словаря структур слов и словосочетаний (*WCEng*), эталонного словаря структур предложений (*SentEng*). Все они отражают структуры различных взаимосвязанных глобальных и локальных конструкций предложений, и потому можно утверждать, что эти словари являются отражением иерархии смысловых единиц, представленных в

предложениях. Исходными данными для формирования такой совокупности словарей может быть любой корпус текстов. При этом для формирования словарей, предназначенных для использования в отраслевых системах перевода, это должен быть репрезентативный корпус отраслевых научно-технических текстов.

Для иллюстрации возможности автоматического формирования эталонных словарей структур конструкций исходных предложений в работе [9] приведены два алгоритма: автоматического формирования словаря структур эталонных словосочетаний (СЭС) и автоматического формирования словаря структур эталонных предложений (СЭП).

По аналогии с этими алгоритмами были созданы алгоритмы формирования структур двуязычных словарей конструкций предложений: словаря предикатно-актантных структур предложений – *PSOEng* (табл. 2), словаря структур «скелетов» предложений – *SkEng* (табл. 3), словаря структур слов и словосочетаний – *WCEng* (табл. 4) и словаря структур предложений – *SentEng* (табл. 5).

В табл. 2 приведен фрагмент словаря предикатно-актантных структур (ПАС) (словарь *PSOEng*). Входом в эту статью служит усеченная поисковая син-

тагма ПАС, далее следует идентификационный номер словарной статьи, синтагма ПАС, ее текстовое представление и поисковая синтагма «скелета» предложения, соответствующая структуре и семантическому наполнению синтагмы ПАС.

В табл. 3 отражен фрагмент словаря структур «скелетов» предложения (словарь *SkEng*). Входом в статью – усеченная поисковая синтагма «скелета» предложения, далее следует идентификационный номер словарной статьи, синтагма «скелета» предложения, его текстовое представление, поисковая синтагма предложения-эталона и поисковые синтагмы локальных конструкций, главные слова, семантическое наполнение которых соответствуют структуре и семантическому наполнению «скелета» предложения и предложения-эталона.

В табл. 4 продемонстрирован фрагмент словаря структур слов и словосочетаний (словарь *WCEng*). Входом в эту статью служит усеченная поисковая синтагма слова или словосочетание с его контекстом, показанные в левой колонке таблицы. В средней колонке отражены идентификационные номера словарных статей, в правой колонке – текстовое представление слова или словосочетания.

Таблица 1

Результаты преобразования предложения в формализованное модифицированное (поисковое) представление его синтаксической структуры

1. Перенумерованное исходное предложение		
#0 The #1 purpose #2 of #3 these #4 tests #5 is #6 to #7 provide #8 a #9 laboratory #10 means #11 of #12 determining #13 the #14 performance #15 characteristics #16 of #17 airborne #18 equipment #19 in #20 environmental #21 conditions #22 representative #23 of #24 those #25 which #26 may #27 be #28 encountered #29 in #30 airborne #31 operation #32 of #33 the #34 equipment #35 .		
2. Результаты синтаксического анализа исходного предложения		
Мнемоническое обозначение	Формальная структура предложения	Пояснения к элементам формальной структуры предложения
NumSen	000000000011111111112222222222333333 012345678901234567890123456789012345	Нумерация слов в предложении (символы нумерации расположены вертикально).
FrmSen	TNfeNvtITANfNTNNfANFANAfekzvEFANfTN.	Предложение в виде символов грамматических классов (обобщенных синтагм) слов.
SntSen	3ТЪЙТz5Ÿ4ДТЬQ3АДЬДА€vТДЪййbzн€ДАЪЗА.	Предложение в виде символов обобщенных синтагм слов.
BndWC	() () () () () () ()	Границы слов и словосочетаний в предложении.
SkISen	Sf N P Nf Nf NF N feS PF Nf N.	«Скелет» предложения в символах классов слов.
BndSen	B eb E	Границы простых предложений.
LnkSen	000000000111111111122222 2222233333 012345789012345678900123 567890145	Связь слов типа «хозяин-слуга» (символы нумерации расположены вертикально).
Sk1MS_1	ТЬТŸТЬДЬА€ТЬŸ= purpose-of-tests-is-to-provide-means-of-determining-characteristics-of-equipment -of-conditions-representative-of-those-which-may-be-encountered-in-operation-of-equipment	ПАС предложения в виде символов грамматических классов слов. Позиции первых слов ПАС в предложении и их длина.
Sk1MS_2	Ÿн€АЪА =which-may-be-encountered-in-operation-of-equipment	

3. Модифицированное представление «скелетов» ПАС предложения		
SPOMSI	<i>z5Ÿ-TĤT-TĤÿ=«is to provide - purpose of test - means of determining characteristics of equipment» z5Ÿ-TĤÿT-vTDĤÿ=«is to provide- the purpose of these tests - a laboratory means of determining the performance characteristics of airborne equipment»</i>	Модифицированный «скелет» предложения в символах классов слов с их текстовым представлением, а также с указанием позиций главных слов и их длины.
SPOMSI	<i>z5Ÿ€-TĤT-TĤÿ=«is to provide in- purpose of test - operation of equipment» z5Ÿ€-TĤÿT-vTDĤÿ=«is to provide in- the purpose of these tests - airborne operation of the equipment »</i>	
SPOMSI	<i>bzn€-k-AĤA=«may be encountered in - which - operation of equipment» bzn€-k-DAĤ3A=«may be encountered in - which - airborne operation of the equipment»</i>	
4. Текстовое и формальное представление локальных конструкций (слов и словосочетаний) предложения		
WC01	<i>T/zT=T=#01=#1 purpose</i>	Структура словосочетаний в символах классов слов с указанием позиции первого слова в предложении и длины словосочетания. Нумерация словосочетаний указана в порядке их следования в предложении.
WC02	<i>T/ĤT=ÿT=#03=#3 these #4 tests</i>	
WC03	<i>z5Ÿ/Tz=z5Ÿ=#05=#5 is #6 to #7 provide</i>	
WC04	<i>A/ŸA=DA=#09=#9 laboratory #10 means</i>	
WC05	<i>Q/ĤQ=Q3AD=#12=#12 determining #13 the #14 performance #15 characteristics</i>	
WC06	<i>T/ĤT=vT=#20=#20 environmental #21 conditions</i>	
WC07	<i>DĤÿÿ/TD=DĤÿÿ=#22=#22 representative #23 of #24 those #25 which</i>	
WC08	<i>bzn/ÿb=bzn=#22=#26 may #27 be #28 encountered</i>	
WC09	<i>A/€A=DA=#30=#30 airborne #31 operation</i>	
WC10	<i>A/ĤA=A=#34=34 equipment</i>	
5. Текстовое и формализованное представление глобальных и локальных конструкций предложения		
<i>z5ŸTĤTTĤÿ=z5Ÿ-TĤT-TĤÿ=«is to provide - purpose of test - means of determining characteristics of equipment»=z5Ÿ-TĤÿT-vTDĤÿ=«is to provide- the purpose of these tests - a laboratory means of determining the performance characteristics of airborne equipment»=TĤTŸTĤDĤA€TĤÿ= «purpose-of-tests-is-to-provide- means-of-determining- characteristics-of - equipment -of-conditions-representative-of- those- which - may - be - encountered - in - operation - of - equipment»=T/zT-T-01- purpose##T/ĤT-ÿT-03-these tests##z5Ÿ/Tz-z5Ÿ-05-is to provide##A/ŸA-DA-09- laboratory means##Q/ĤQ-Q3AD-12-determining the performance characteristics##T/ĤT -vT-20-environmental conditions##DĤÿÿ/TD-DĤÿÿ-22-representative of those which##bzn/ÿb-bzn-26-may be encountered##A/€A-DA-30- airborne operation##A/ĤA -A-34- equipment=3TĤÿTz5Ÿ4DTĤQ3ADĤDA€vTDĤÿÿbzn€DAĤ3A=«The purpose of these tests is to provide a laboratory means of determining the performance characteristics of airborne equipment in environmental conditions representative of those which may be encountered in airborne operation of the equipment.»</i>		

Таблица 2

Фрагмент словаря предикатно-актантных структур предложений – PSOEng

<i>z5ŸTĤTTĤÿ=03528654= z5ŸTĤÿTvTDĤÿ=«is to provide- the purpose of these tests - a laboratory means of determining the performance characteristics of airborne equipment»= TĤTŸTĤDĤA€TĤÿ</i>
<i>z5Ÿ€TĤTTĤÿ= 03528654=z5ŸTĤÿTvTDĤÿ=«is to provide- the purpose of these tests - a laboratory means of determining the performance characteristics of airborne equipment»= TĤTŸTĤDĤA€TĤÿ</i>
<i>bzn€kDAĤ3A=0087335 = bzn€kDAĤ3A =«may be encountered in - which -airborne operation of the equipment»=ÿbzn€ĤA</i>

Таблица 3

Фрагмент словаря структур «скелетов» предложений – SkiEng

<i>TĤTŸTĤDĤA€TĤÿ=00364876= «purpose-of-tests-is-to-provide-means-of-determining-characteristics-of-equipment-of-conditions- representative-of-those»= T/zT-ÿT /ĤT-z5Ÿ/Tz-DA/ŸA-Q3AD /ĤQ-DA/ĤA-vT/ĤT-DĤÿÿ/TD =3TĤÿTz5Ÿ4DTĤQ3ADĤDA€vTDĤÿÿbzn€DAĤ3A</i>
<i>ÿn€ĤA=00784564=«which-may-be-encountered-in-airborne-operation-of-equipment»=bzn/ÿb-A/ĤA=3TĤÿTz5Ÿ4DTĤQ3ADĤDA€vTDĤÿÿÿbzn€DAĤ3A</i>

Фрагмент словаря структур слов и словосочетаний – *WCEng*

Поисковая синтагма	№ словарной статьи	Содержание словарной статьи
<i>T/zT</i>	<i>03528654</i>	<i>purpose</i>
<i>йт/ьт</i>	<i>03845286</i>	<i>these tests</i>
<i>z5Y/Tz</i>	<i>02753528</i>	<i>is to provide</i>
<i>ДА/ЎА</i>	<i>01743524</i>	<i>laboratory means</i>
<i>Q3AD/ЬQ</i>	<i>02783352</i>	<i>determining the performance characteristics</i>
<i>ДА/ЬА</i>	<i>03248651</i>	<i>airborne equipment</i>
	<i>01874639</i>	<i>in</i>
<i>vT/ЬT</i>	<i>01583635</i>	<i>environmental conditions</i>
<i>ДЬй/ТD</i>	<i>02385437</i>	<i>representative of those which</i>
<i>йbzн/йb</i>	<i>00837845</i>	<i>which may be encountered</i>
	<i>01874743</i>	<i>in</i>
<i>А/ЬА</i>	<i>03043836</i>	<i>airborne operation of the equipment</i>

Таблица 5

Словарная статья структур словаря предложений-эталонов – *SentEng*

3ТЬйТz5Y4DТЬQ3ADЬДА€vTDЬййbzн€ДАЬ3А= 0564783=

«*The purpose of these tests is to provide a laboratory means of determining the performance characteristics of airborne equipment in environmental conditions representative of those which may be encountered in airborne operation of the equipment.*»

В табл. 5 предложен пример словарной статьи структур словаря предложений-эталонов (словарь *SentEng*). Входом в статью здесь является поисковая синтагма исходного предложения, далее следуют идентификационный номер словарной статьи в словаре, завершает содержание статьи текстовое представление исходного предложения.

Процесс формализации синтаксической структуры конструкций исходного предложения и преобразования их в поисковые представления конструкций предложений эталонных словарей достаточно подробно был изложен выше. Теперь необходимо кратко пояснить, как будет осуществляться поиск сходных синтаксических конструкций в иерархии рассмотренных выше эталонных словарях.

Этап поиска осуществляется после того как была выполнена формализация синтаксической структуры конструкций исходного предложения и осуществлено их преобразование в поисковые представления. Далее производится поиск сходных структур локальных конструкций предложений по схеме: вначале по сформированной на этапе синтаксического анализа поисковой синтагме предикатно-актантной структуры (ПАС) предложения в словаре *PSOEng* устанавливается соответствие с эталонной структурой ПАС и соответствующей этой ПАС поисковой синтагмой «скелета» предложения, затем по этой синтагме в словаре *SntEng* находится формализованное представление «скелета», и, наконец, по поисковым синтагмам слов и словосочетаний и их окружению в словаре *WCEng* ищутся эталонные структуры слов и

словосочетаний, соответствующие структуре «скелета» анализируемого предложения. При необходимости всю недостающую информацию для полноценного анализа исходного предложения можно получить из словарной статьи эталонного словаря *SentEng*, входом в которую служит поисковая синтагма «скелета» предложения.

При этом, в случае полного совпадения поисковых синтагм конструкций исходного предложения с формализованными представлениями синтаксических структур эталонных конструкций можно весь набор содержащейся в них информации приписать анализируемому предложению – как результат его семантико-синтаксического анализа. В случае частичного совпадения синтаксических структур конструкций нужно выполнить дополнительный анализ тех частей формализованных представлений конструкций анализируемого предложения, которые не совпали с частями словарных статей структур эталонных словарей с целью внесения необходимых корректив в построение синтаксической структуры анализируемого предложения.

Рассмотрим процесс внесения корректив более подробно. Но прежде еще раз остановимся на содержащихся в эталонных словарях исходных и результирующих данных, к которым относится: 1) информация о формализованной предикатно-актантной структуре (ПАС) исходного предложения; 2) информация о формализованной структуре «скелета» исходного предложения; 3) информация о формализованных представлениях исходных слов и словосочетаний с указанием их

местоположения и «стыковочных узлах», обеспечивающих их синтаксическую связь с контекстным окружением в предложении.

При установлении структурного сходства предложений, прежде всего, определяется тождественность предикатно-актантной структуры исходного предложения с одной из структур словарной статьи словаря *PSOEng*. Если такая статья отсутствует, синтаксический анализ дальше следует выполнять по традиционной схеме. В случае обнаружения тождественной эталонной структуры ПАС необходимо в массиве поисковых синтагм «скелетов», приведенном в этой словарной статье, найти наиболее близкую по структуре «скелета» его поисковую синтагму. Далее по ней в словаре *SkEng* находится словарная статья, анализируются структуры эталонного скелета со скелетом исходного предложения и устанавливается степень их структурного сходства. Исключаются как избыточные элементы эталонной структуры, так и соответствующие им формализованные структуры слов и словосочетаний вместе с их «стыковочными узлами». В случае недостатка таких конструкций они включаются в состав «новой» структуры предложения.

Таким образом, выполнив упрощенный синтаксический анализ предложения и получив усеченную информацию о его структуре в виде поисковых синтагм, можно менее затратно добиться результатов более высокого качества, чем, если бы этот анализ был произведен традиционной процедурой семантико-синтаксического анализа.

ПРИНЦИПЫ ПОСТРОЕНИЯ ПРОЦЕДУРЫ СЕМАНТИКО-СИНТАКСИЧЕСКОГО СИНТЕЗА ТЕКСТОВ

Основная задача семантико-синтаксического синтеза при переводе текстов заключается в построении формализованной структуры переводного предложения и его текстового представления на основе информации о синтаксической структуре исходного предложения, информации о входящих в его состав локальных синтаксических конструкциях (об отдельных словах, словосочетаниях, причастных и деепричастных оборотах др.) и их переводных соответствий. Исходными данными для синтаксического синтеза текстов при переводе являются результаты морфологического и синтаксического анализа исходного (английского) текста, а также последовательность переводных соответствий, полученных на этапе трансфера [7]. На основе этой информации строится синтаксическая структура переводного предложения, производятся необходимые перестановки слов и словосочетаний и выполняется их морфологический синтез. Такая последовательность процедур синтаксического синтеза характерна для автоматического перевода текстов с одного естественного языка на другой, в частности при переводе текстов с английского языка на русский.

На этапе трансфера (этот этап предшествует этапу семантико-синтаксического синтеза [7]) все слова и словосочетания исходного предложения соотносятся с их переводными соответствиями. Зафиксированная между ними связь обеспечивает возможность реализации локального синтаксического синтеза именных

и глагольных словосочетаний, выполняющегося путем их локальных синтаксических трансформаций и перестановками слов внутри них.

На этапе синтаксического синтеза переводного предложения осуществляются глобальные синтаксические трансформации и перестановки локальных синтаксических конструкций. Установленная формальная связь позволяет более точно определить синтаксическое соответствие между главными (подлежащим и сказуемым) и второстепенными (дополнением и обстоятельством места, времени и т.д.) членами исходного предложения и их переводных соответствий. По сути, переводное предложение строится путем преобразования последовательности переводов слов и словосочетаний исходного предложения при воспроизведении его синтаксической структуры в соответствии с грамматическими правилами выходного текста.

Например, по правилам русского языка сказуемое должно грамматически согласовываться с подлежащим, а внутри групп сказуемого и подлежащего, не содержащих словарных фразеологических единиц, согласовываться должны слова «слуги» со словами «хозяевами». Например, «слуга» – прилагательное (А) должно согласовываться с «хозяином» – существительным (N) в роде, числе и падеже.

Таким образом, синтаксическая структура выходного текста в значительной мере определяется синтаксической структурой фразеологических словосочетаний, выбранных из словаря. А та часть текста, которая не покрывается словарными фразеологическими словосочетаниями, представляется переводными эквивалентами отдельных слов, которые согласуются друг с другом по правилам грамматики и в соответствие с их контекстным окружением. При этом иногда требуется выполнить локальные перестановки слов. Например, при переводе на английский язык русского именного словосочетания «поиск информации» переводной эквивалент слова «информация» (*information*) располагается впереди переводного эквивалента слова «поиск» (*retrieval*) и в результате формируется словосочетание *information retrieval*.

В выходных английских текстах могут также выполняться и более масштабные перестановки, например, группы подлежащего и группы сказуемого, если во входном русском тексте группа сказуемого предшествовала группе подлежащего.

Важной проблемой, возникающей при синтезе русских текстов, является правильный учет таких характеристик слов, как их модели управления. Обычно в системах машинного перевода эти характеристики указываются в словаре основ слов. Но в рассматриваемом варианте реализации концепции фразеологического машинного перевода не предусматривается хранение в памяти ЭВМ сводного словаря всех основ слов, встречающихся в наименованиях понятий. Поэтому здесь выбран другой путь решения проблемы: управляемым словам текстов «по умолчанию» придаются такие грамматические формы, которые являются наиболее вероятными в типовых контекстных окружениях, а «аномальные» случаи отражаются в специальном словаре моделей управления. В нем хранится перечень основ только тех слов, которые способны выступать в роли управ-

ляющих и у которых подчиненные слова могут иметь “аномальную” форму.

Например, для существительного, стоящего после другого существительного, наиболее вероятной падежной формой является форма родительного падежа, значительно реже встречаются формы дательного или творительного падежа; у существительного, стоящего после глагола, чаще всего бывает форма винительного падежа, реже – дательного или творительного. В случае предлогов, допускающих два варианта управления (например, предлогов “в”, “на”, “о” и “с”), существительные, стоящие после предлогов “в”, “на”, или “о”, чаще всего принимают форму предложного падежа, а стоящие после предлога “с” – форму творительного падежа. Другие падежные формы встречаются значительно реже.

“Аномальные” случаи управления “подчиненными” словами выявлялись путем автоматизированной обработки русских текстов с использованием процедур их морфологического и синтаксического анализа. При этом находились те управляющие слова, которые были причиной появления в текстах “аномальных” падежных форм.

Основная идея синтаксического синтеза переводных предложений заключается в последовательном преобразовании локальных конструкций предложений – именных и глагольных словосочетаний, причастных и деепричастных оборотов и др. в переводные синтаксические конструкции, эквивалентные синтаксическим конструкциям исходного предложения, но построенные в соответствии с правилами выходного языка. Такие преобразования выполняются путем локальных перестановок слов и словосочетаний в пределах локальных синтаксических конструкций и назначения им соответствующих морфологических и синтаксических признаков.

На следующем этапе синтаксического синтеза производятся глобальные перестановки уже самих синтаксических конструкций в пределах предложения и выполняется грамматическое согласование опорных слов этих конструкций.

МАШИННЫЙ ФРАЗЕОЛОГИЧЕСКИЙ ПЕРЕВОД НА ОСНОВЕ МОДЕЛИ ОБОБЩЕННЫХ СИНТАГМ

Выполненные авторами исследования [9] показали, что на основе модели обобщенных синтагм возможно однозначно представить исходное предложение и его перевод в виде иерархии структур глобальных и локальных конструкций исходных предложений и их переводов с соответствующими текстовыми представлениями. Имея представительные словари таких конструкций предложений, можно реализовать процесс фразеологического машинного перевода по более короткой и простой технологической схеме, которая сводится этапам: 1) этапу упрощенного синтаксического анализа исходного текста и формирования на основе его результатов поисковых синтагм глобальных и локальных конструкций предложений; 2) этапу поиска в эталонных словарях структур исходных глобальных и локальных конструкций их переводных соответствий и формирования на их основе текстового представления переводного предложения.

Реализация основного требования – правильной передачи смыслового содержания исходного текста средствами выходного языка – обеспечивается: а) качественным переводом локальных синтаксических конструкций исходного предложения; б) порядком их следования в переводном предложении; в) порядком следования слов внутри этих конструкций; г) генерацией правильных грамматических форм слов переводных соответствий. Если предположить, что перевод локальных конструкций выполнен на качественном уровне, то на этапе синтаксического синтеза необходимо корректно построить синтаксическую структуру переводного предложения. Эту задачу возможно решить путем последовательного построения структур локальных и глобальных конструкций переводного предложения по аналогии с уже имеющимися конструкциями в иерархии словарей.

Проведенные исследования [9] по решению задачи однозначного соотнесения синтаксической структуры исходного предложения, представленной в виде последовательности символов обобщенных синтагм, и его семантического наполнения, представленного в виде последовательности слов и словосочетаний исходного предложения, с синтаксической структурой (в виде обобщенных синтагм) и семантического наполнения (в виде переводов слов и словосочетаний исходного предложения) переводного инварианта показали что, основной проблемой является значительная вариативность представления одного и того же смыслового содержания исходных предложений. Поэтому возникла необходимость разработки методов формализации представления смысловой структуры предложений в виде совокупности их глобальных и локальных конструкций. Основным требованием, предъявляемым к таким конструкциям, являлась их повторяемость в различных текстах. Как известно, такими устойчивыми единицами смысла, содержащимися в составе предложений и обладающими большей повторяемостью в текстах, являются фразеологические и терминологические наименования понятий, представленные словами и словосочетаниями. Значительно меньшей повторяемостью характеризуются конструкции предикатно-актантной структуры предложений. Еще меньше повторяемость у конструкций формализованных «скелетов» предложений, отражающих различные текстовые представления одного смысла. Исходя из этого, при формализации смысловой структуры предложений было решено опираться на иерархию этих синтаксических конструкций, обеспечивающих возможность распознавания смысловой структуры эталонных предложений и возможность соотнесения локальных и глобальных конструкций (исходных и переводных) в рамках контекста предложения.

Поэтому при разработке методов и технологий реализации модернизированной системы фразеологического машинного перевода на основе обобщенных синтагм было необходимо разработать иерархию словарей конструкций эталонных предложений и их переводов. Эта иерархия должна включать:

- 1) словарь эталонных конструкций предикатно-активных структур;
- 2) словарь эталонных конструкций «скелетов» предложений;

3) словари структур исходных эталонных словосочетаний и их переводов;

4) словари структур исходных и переводных эталонных предложений, построенные на основе обобщенных синтагм.

Таким образом, процесс машинного перевода в реализации предлагаемого технологического процесса будет состоять из двух этапов: (1) формализации смысловой структуры конструкций исходного предложения и преобразования их в поисковые представления эталонных конструкций исходных структур предложений и (2) поиска в словарях эталонных конструкций исходных структур предложений и их переводов, идентичных по структуре и семантическому наполнению конструкциям эталонных предложений, и, при необходимости, коррекции синтаксической структуры и их семантического наполнения.

В рамках рассматриваемой технологии перевода также необходимо разработать процедуры автоматического установления структурного сходства синтаксических конструкций и их текстовых представлений анализируемого исходного предложения с синтаксическими конструкциями предложений-эталонов и их переводных инвариантов. Для решения этой задачи создаются методы и критерии установления структурного и смыслового сходства конструкций предложений.

В процессе сопоставления конструкций в случаях, если критерии сходства конструкций удовлетворяют критериям эквивалентности, этому исходному предложению будут поставлены в соответствие найденные в словаре конструкции переводных эталонных предложений. В противном случае необходимо произвести поэлементное сопоставление локальных конструкций анализируемого исходного предложения с конструкциями предложений-эталонов, выполнить анализ результатов этого сопоставления и произвести последующую коррекцию состава и порядка следования локальных конструкций, построив новое (по отношению к имеющемуся эталонному) текстовое представление переводного предложения.

В табл. 6 приведен пример формализации синтаксической структуры английского предложения в процессе перевода и формализации его глобальных и локальных поисковых структур. В разделе 1 таблицы дано текстовое представление исходного предложения с нумерацией каждого слова. В разделе 2 показаны результаты сборки в одну конструкцию всех преобразованных текстовых форм глобальных и локальных конструкций предложения и их формализованных поисковых представлений. В этой конструкции информация позиционно разделена на блоки грамматической, синтаксической и семантической информации, отделяемые друг от друга знаком «=» (равно). Информация внутри блоков разделяется знаком «-» (тире) или знаком «#» (решетка), однородная информация следует через запятую. Блоки информации расположены в следующем порядке:

- 1) сокращенная поисковая синтагма предикатно-актантной структуры (ПАС) предложения;
- 2) полная поисковая синтагма ПАС предложения;
- 3) текстовое представление поисковой синтагмы ПАС предложения;
- 4) «скелет» исходного предложения в символах обобщенных синтагм;
- 5) «скелет» переводного предложения в символах обобщенных синтагм;
- 6) информация о выполненных (или не выполненных) перестановках локальных конструкций в простом переводном предложении;
- 7) текстовое представление «скелета» исходного предложения;
- 8) текстовое представление «скелета» переводного предложения;
- 9) структура исходного предложения в символах обобщенных синтагм;
- 10) структура переводного предложения в символах обобщенных синтагм;
- 11) текстовое представление исходного предложения;

Более подробно этот процесс описан в работе [7].

Таблица 6

Пример формализации синтаксической структуры английского предложения в процессе перевода и формализации его глобальных и локальных поисковых структур

1. Перенумерованное исходное предложение
#0 The #1 purpose #2 of #3 these #4 tests #5 is #6 to #7 provide #8 a #9 laboratory #10 means #11 of #12 determining #13 the #14 performance #15 characteristics #16 of #17 airborne #18 equipment #19 in #20 environmental #21 conditions #22 representative #23 of #24 those #25 which #26 may #27 be #28 encountered #29 in #30 airborne #31 operation #32 of #33 the #34 equipment #35 .
2. Поисковые представления глобальных и локальных конструкций предложения
z5ŸТЪТТЪŸ=z5ŸТЪŸТvТDЪŸ=«is-to-provide-the-purpose-of-these-tests alaboratorymeansofdeterminingtheperformancecharacteristicsofairborneequipment»=ТЪТŸТЪDЪАЕТЪŸ= AeЮe7АЙSЖАЙЦwАЙЦ7АЙSuqяАыА=01020304050607080910111213=«purpose-of-tests-is-to- provide-means-of-determining-characteristics-of-equipment-of-conditions-representative-of-those=цель- тестов-заключается-в-обеспечении-средств-определения-характеристик-оборудования-в-условии- среды-для-тех»= 3ТЪŸТz5Ÿ4DТЪQ3ADЪDA€vТDЪŸŸbzn€DAЪ3A=twЪaAeжBцAФrЖВЙЦФvwAЧŸЙЦ7 АЙS8KuqФvяАыА.,ФрбOэАцA8ueся7AЧhxСЙЦ=«The-purpose-of-these-tests-is-to-provide-a-laboratory- means-of-determining-the-performance-characteristics-of-airborne-equipment-in-environmental- conditions-representative-of-those-which-may-been-counterred-in-airborne-operation-of-the- equipment.»=«Цель-этих-испытаний-состоит-в-том,-чтобы-предоставить-лабораторные- средства-для-определения-эксплуатационных-характеристик-бортового-оборудования-в-условиях- окружающей-среды,-типичных-для-тех,-которые-могут-встречаться-при-воздушной- эксплуатации-оборудования.»

Преобразование текстового представления анализируемого предложения в совокупность формализованных представлений его глобальных и локальных конструкций выполняется на этапе упрощенного семантико-синтаксического анализа исходного предложения. После получения требуемых поисковых представлений анализируемого предложения последовательно запускаются механизмы поиска на наибольшее совпадение левых частей поисковых синтагм конструкций анализируемого предложения в иерархии структур эталонных переводных словарей: вначале по сформированной на этапе синтаксического анализа поисковой синтагме предикатно-актантной структуры в словаре *PSOTEngRus* устанавливается соответствие с эталоном из словаря предикатно-актантной структуры и соответствующей ей поисковой синтагме «скелета» предложения, далее по этой синтагме в словаре *SntTEngRus* ищется формализованное представление «скелета» и его семантическое наполнение, и, наконец, по поисковым синтагмам слов и словосочетаний и их окружений в словаре *WCTEngRus* находятся эталонные слова и словосочетания, соответствующие структуре и семантическому наполнению «скелета» анализируемого предложения. В случае необходимости всю недостающую информацию для полноценного перевода можно получить из словарной статьи словаря *SentTEngRus*, входом в которую служит поисковая синтагма «скелета» предложения.

В случае совпадения поисковых синтагм конструкций исходного предложения, представлений синтаксических структур и семантического наполнения с аналогичными представлениями эталонных конструкций предложений можно весь набор информации конструкций переводного эталонного предложения приписать анализируемому предложению как результат его семантико-синтаксического синтеза. В

случае частичного совпадения синтаксических структур нужно выполнить дополнительный анализ тех частей формализованных и текстовых представлений конструкций анализируемого предложения, которые не совпали с частями словарных статей эталонных словарей с целью внесения соответствующих корректив в построение переводного предложения.

В табл. 7 представлен фрагмент словаря предикатно-актантных структур (ПАС) исходного предложения и его перевода (словарь *PSOTEngRus*). Входом в эту статью служит усеченная поисковая синтагма ПАС, далее следует идентификационный номер словарной статьи, текстовое представление ПАС и поисковая синтагма «скелета» предложения, соответствующего структуре и семантическому наполнению синтагмы ПАС.

В табл. 8 показан фрагмент словаря структур «скелетов» исходного предложения и его перевода (словарь *SkITEngRus*). Входом является поисковая синтагма «скелета» предложения, далее следует идентификационный номер словарной статьи, текстовое представление исходного «скелета», порядок следования переводов словосочетаний и поисковые синтагмы локальных конструкций (главные слова которых соответствуют структуре «скелета»), а их текстовое представление входит в состав текстового представления соответствующего предложения-эталона).

В табл. 9 приведен фрагмент словаря структур исходных слов и словосочетаний и их переводов (словарь *WCTEngRus*). Вход в эти статьи – усеченные поисковые синтагмы слова или словосочетания с их «стыковочными узлами», показанные в левой колонке таблицы. В средней колонке отражены идентификационные номера словарных статей и в правой колонке – содержание словарной статьи, включающее полную синтагму слова или словосочетания с его контекстом и текстовым представлением.

Таблица 7

Фрагмент словаря предикатно-актантных структур исходного предложения и его перевода – *PSOTEngRus*

z5ЎТЪТТЪй=03528654 =z5ЎТЪйTvTDИЪй=«is to provide- the purpose of these tests - a laboratory means of determining the performance characteristics of airborne equipment»= ТЪТЎТЪДЪАЕТЪй
z5Ў€ТЪТТЪй= 03528654 =z5ЎТЪйTvTDИЪй=«is to provide- the purpose of these tests - a laboratory means of determining the performance characteristics of airborne equipment»= ТЪТЎТЪДЪАЕТЪй
bzn€kDAИЪ3A=0087335 = bzn€kDAИЪ3A =«may be encountered in - which -airborne operation of the equipment»=ÿbzn€АИЪА

Таблица 8

Фрагмент словаря структур «скелетов» исходного предложения и его перевода – *SkITEngRus*

ТЪТЎТЪДЪАЕТЪй=00364876 = «purpose-of-tests-is-to-provide-means-of-determining-characteristics-of-equipment-of-conditions- representative-of-those»= 0102030405060608=T/zT-ÿT /ИЪТ-z5Ў/Tz-DA/ЎА-Q3AD /ИЪQ-DA/ИЪА-vT/ИЪТ-DИЪй/TD = 3ТЪйТ z 5Ў4DTЪQ3ADИЪДА€vTDИЪйÿbzn€DAИЪ3A
ÿн€АИЪА=00784564 =«which-may-be-encountered-in-airborne-operation-of-equipment»=bzn/ÿb-A/ИЪА= 3ТЪйТ z 5Ў4DTЪQ3ADИЪДА€vTDИЪйÿbzn€DAИЪ3A

Фрагмент словаря структур исходных слов и словосочетаний и их переводов – *WCTEngRus*

Поисковая синтагма	№ словарной статьи	Содержание словарной статьи
<i>T/zT</i>	03528654	<i>tw-purpose / цель</i>
<i>ÿT /ЪТ</i>	03845286	<i>ЪаЙс -tests/испытаний – these-tests/этих-испытаний</i>
<i>z5ÿ/Tz</i>	02753528	<i>Юе7АЙС –is-to-provide/заключается-в-обеспечении</i>
<i>DA/ЪA</i>	01743524	<i>ФvЖА – laboratory-means/лабораторное-оборудование</i>
<i>Q3AD /ЪQ</i>	02783352	<i>ЙЦФvwA-determining-the-performance-characteristics/определение-эксплуатационных-характеристик</i>
<i>DA/ЪA</i>	03248651	<i>ЧfЙЦ – airborne-equipment/бортового-оборудования</i>
<i>€ / z5ÿ</i>	01567324	<i>7A-in/в</i>
<i>vT/ЪT</i>	01583635	<i>ЙS8Kuq- environmental-conditions/условиях-окружающей-среды</i>
<i>DЪÿй/TD</i>	02385437	<i>ФvяАыА,,Fr-representative-of-those-which/характерных-для-тех, которые</i>
<i>bzn/ÿb</i>	00837845	<i>боцА –may-be-encountered/могут-встретиться</i>
<i>€ / z5ÿ</i>	02643248	<i>3A-in/при</i>
<i>A/ЪA</i>	03043836	<i>ФhхСЙЦ – airborne-operation-of-the-equipment/воздушной-эксплуатации-оборудования</i>

Таблица 10

Словарная статья словаря структур исходных предложений-эталонов и их переводов – *SentTEngRus*

<p>ЗТЬЪТz5ÿ4DTЪQ3ADЪDA€vTDЪÿйbzn€DAЪ3A= 0564783=</p> <p><i>twЪаАежВцАФрЖВЙЦФvwAЧfЙЦ7АЙС8KuqФvяАыА,,FrбОэАцА8ueся7АЧhхСЙЦ=«The-purpose-of-these-tests-is-to-provide-a-laboratory-means-of-determining-the-performance-characteristics-of-airborne-equipment-in-environmental-conditions-representative-of-those-which-may-been-countered-in-airborne-operation-of-the-equipment./Цель-этих-испытаний-состоит-в-том-, -чтобы-предоставить-лабораторные-средства-для-определения-эксплуатационных-характеристик-бортового-оборудования-в-условиях-окружающей-среды-, -типичных-для-тех, которые-могут-встречаться-при-воздушной-эксплуатации-оборудования.»</i></p>

В табл. 10 продемонстрирован пример словарной статьи словаря структур исходных предложений-эталонов и их переводов (словарь *SentTEngRus*). Входом служит полная поисковая синтагма исходного предложения, далее следуют идентификационный номер словарной статьи в словаре и полная синтагма перевода предложения. Завершает содержание статьи текстовое представление исходного предложения и его перевода.

Формирование эталонных словарей переводов глобальных и локальных конструкций предложений можно обеспечить, используя массивы двуязычных текстов (билингв) и имеющиеся у авторов настоящей статьи сверхбольшие двуязычные фразеологические словари (объемом свыше 3,5 млн словарных статей) [7].

Полученные нами предварительные результаты при использовании предлагаемой технологической схемы фразеологического машинного перевода на основе синтаксических шаблонов показали принципиальную возможность реализации этой схемы при существенных преимуществах по скорости и качеству перевода текстов по сравнению с переводами, выполненными традиционной системой.

Все разработанные модели двуязычных языковых конструкций были преобразованы в их машинное

представление в виде набора двуязычных пользовательских, тематических и политематических словарей.

ПРОГРАММНАЯ РЕАЛИЗАЦИЯ СИСТЕМЫ ФРАЗЕОЛОГИЧЕСКОГО МАШИННОГО ПЕРЕВОДА

Обобщенная архитектура современной системы фразеологического машинного перевода представляет комплекс взаимосвязанных декларативных и процедурных средств, реализующих процесс перевода текстов с одних естественных языков на другие. К декларативным средствам относятся двуязычные словари и различного рода грамматические таблицы, к процедурным – программная реализация, обеспечивающая функционирование системы в целом и ее отдельных модулей.

Процедурные средства ядерной части системы фразеологического машинного перевода – это сложный программный комплекс, включающий в свой состав более двух тысяч функциональных и служебных программных модулей. Из них основными являются:

- Морфологический анализ слов для русского и английского языков;
- Морфологический синтез слов для русского и английского языков;

- Лемматизация английских слов для русского и английского языков;
- Концептуальный анализ текстов для русского и английского языков;
- Синтаксический анализ текстов для русского и английского языков;
- Синтаксический синтез текстов для русского и английского языков;
- Трансфер – процедура замены наименований понятий исходного языка на наименования понятий выходного языка;
- Служебные программы управления работой ядерной частью системы фразеологического машинного перевода, в частности, комплекс программ, обеспечивающий общение пользователя с этой системой (интерфейс пользователя).

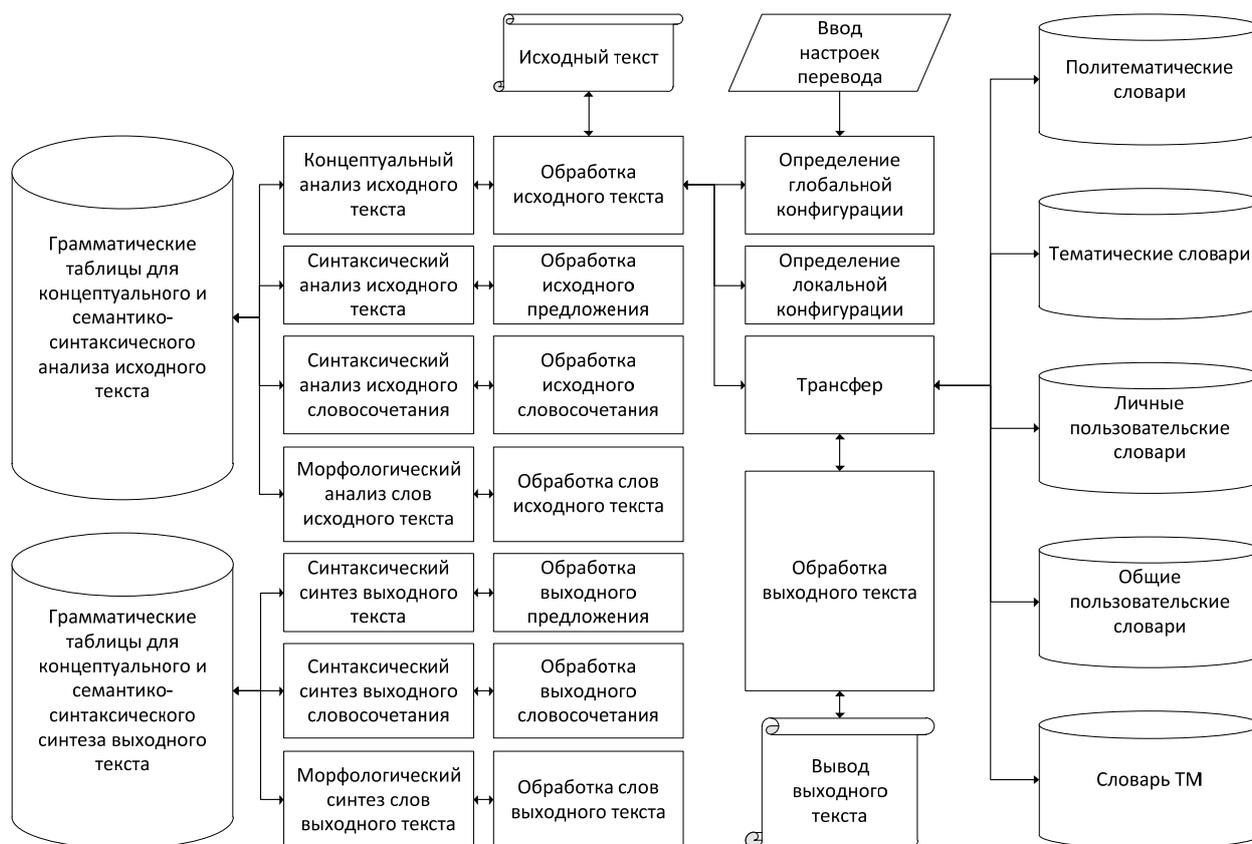
Общая информационно-технологическая архитектура системы фразеологического машинного перевода приведена на рисунке.

Система фразеологического машинного перевода включает основные декларативные средства:

- политематические двуязычные фразеологические машинные словари;
- набор тематических двуязычных фразеологических машинных словарей;

- пользовательские двуязычные фразеологические словари;
- пользовательские двуязычные словари *ТМ* (*Translation Memory*);
- массивы билингв (массивы исходных предложений по узким тематикам и их переводов на выходной язык);
- грамматические таблицы, необходимые для морфологического и синтаксического анализа и синтеза словоформ для каждого языка;
- словарь обобщенных синтагм, необходимый для разрешения грамматической многозначности слов для каждого языка;
- словарь “аномальных” случаев моделей управления слов для каждого языка;
- словарь слов, необходимый для разрешения многозначности моделей управления предлогов для каждого языка.

На основе предлагаемых принципов была разработана программная реализация модернизированной системы фразеологического машинного перевода. Её отличием от традиционной версии является исключение процедуры трансфера и существенное упрощение процедур семантико-синтаксического анализа и синтеза текстов.



Общая информационно-технологическая архитектура системы фразеологического машинного перевода.

ОЦЕНКА ЭФФЕКТИВНОСТИ МЕТОДА

Для оценки эффективности предложенной модели был выбран стандартный метод *BLEU* [10], базирующийся на процедуре сопоставления эталонного и машинного перевода исходного текста. Качество перевода оценивалось по степени их сходства. Но, поскольку, одно и то же смысловое содержание текста может быть передано различным лексическим составом, то возможна ситуация, когда переведенный текст правильно передает смысловое содержание исходного текста лексическим составом, отличным от лексики эталонного перевода. Поэтому оценка качества перевода путем сопоставления текстового представления двух переводов стандартными метриками не всегда является объективной. Для повышения качества оценки перевода необходимо сопоставлять не лексический состав, а смысловое содержание, что требует приведения сравниваемых текстов к их унифицированному семантическому представлению [10]. Для этой задачи использован словарь унифицированных форм представлений наименований понятий, сформированный автоматически путем приведения текстовых форм наименований понятий, выражающих один и тот же смысл, к их унифицированным формам представления.

На основе такого подхода была произведена оценка качества машинного перевода, выполненного традиционной и модернизированной системами фразеологического машинного перевода на основе обобщенных синтагм. Результаты показали повышение качества перевода на 14% на выборке, состоящей из 150 параллельных текстов новостей общественно-политической тематики.

ЗАКЛЮЧЕНИЕ

Если кратко сформулировать основную идею предлагаемого подхода к реализации технологии фразеологического машинного перевода, то она заключается в попытке создать технологию «семантического переводческого конструктора» на основе иерархии шаблонов двуязычных синтаксических конструкций. С его помощью на первом этапе – упрощенный синтаксический анализ – реализуется «разборка» предложения на глобальные и локальные конструкции, сопровождаемые их «стыковочными узлами» – поисковыми синтагмами. На втором этапе в словарной базе – иерархии структур эталонных словарей локальных и глобальных синтаксических конструкций и их переводных соответствий – осуществляется поиск по этим поисковым синтагмам всех необходимых конструкций для окончательной «сборки» текстового представления переводного предложения. И, наконец, на завершающем этапе происходит окончательная «сборка» в соответствии с «технологическими инструкциями», заложенными в правилах последовательности сборки и правилах использования элементов конструкций.

Предварительные сравнительные исследования скорости и качества перевода текстов, выполненные как традиционной, так и модернизированной систе-

мами фразеологического машинного перевода на основе обобщенных синтагм показали, что за счет исключения наиболее ресурсозатратных технологических операций, таких, например, как операции трансфер, и значительного упрощения процедур синтаксического анализа и синтеза текстов, скорость перевода может быть повышена на 30-50%. Качество перевода также существенно улучшается (на 14% и выше). Это обеспечивается за счет синтаксических шаблонов, использование которых приводит к более правильному переводу тех семантически сложных фрагментов текстов, которые ранее невозможно было адекватно перевести традиционной системой фразеологического машинного перевода.

СПИСОК ЛИТЕРАТУРЫ

1. Кулагина О.С. Исследования по машинному переводу. – М.: Наука, 1979.
2. Nagao M. A framework of a mechanical translation between Japanese and English by analogy principle, in *Artificial and Human Intelligence*, ed. A. Elithorn and R. Banerji, p.p. 173-180, North Holland, 1984.
3. Гольдберг Й. Нейросетевые методы в обработке естественного языка. – М.: Изд-во ДМК, 2019.
4. Осинга Д. Глубокое обучение. Готовые решения. – СПб: Изд-во «Диалектика», 2019.
5. Белоногов Г.Г., Калинин Ю.П., Хорошилов А.А. Компьютерная лингвистика и перспективные информационные технологии. Теория и практика построения систем автоматической обработки текстовой информации. – М.: Изд-во Русский мир, 2004.
6. Белоногов Г.Г., Хорошилов Ал-др А., Хорошилов Ал-сей А., Козачук М.В., Рыжова Е.Ю., Гуськова Л.Ю. Каким быть машинному переводу в XXI веке // *Перевод: традиции и современные технологии*. – М.: ВЦП, 2002.
7. Хорошилов Ал-др А., Кан А.В., Хорошилов А.А. Фразеологический машинный перевод. – М.: Изд-во «Директ-Медиа», 2019.
8. Аблов И.В., Козичев В.Н., Ширманов А.В., Хорошилов Ал.-др А., Хорошилов А.А. Средства машинной грамматики русского языка (по Г.Г. Белоногову) // *Научно-техническая информация. Сер. 2*. – 2018. – № 6. – С. 32-46; Ablov I.V., Kozichev V.N., Shirmanov A.V., Khoroshilov Al-dr A., Khoroshilov Al-ey A. The Tools of a Machine Grammar of the Russian Language (based on G.G. Belonogov) // *Automatic Documentation and Mathematical Linguistics*. – 2018. – Vol. 52, № 3. – P. 142-156.
9. Кан А.В., Ревина В. Д., Руснак В.И., Хорошилов Ал-др А., Хорошилов А.А. Автоматическое формирование синтаксической модели языка для задач машинного перевода и информационного поиска // *Научно-техническая информация. Сер. 2*. – 2018. – № 12. – С. 25–41.

10. Руснак В.И. Метод повышения качества оценки систем машинного перевода с использованием фразеологических двуязычных словарей // Тезисы докл. 18-й Международной конференции «Авиация и космонавтика. – М.: НИУ МАИ, 2019

Материал поступил в редакцию 25.12.19.

Сведения об авторах

ХОРОШИЛОВ Алексей Александрович – кандидат технических наук, научный сотрудник Федерального государственного учреждения «Федеральный исследовательский центр «Информатика и управление» Российской академии наук» (ФИЦ ИУ РАН), Москва
e-mail: a.a.horoshilov@mail.ru

КАН Анна Владимировна – кандидат технических наук, начальник аналитического отдела ФГБУ "НИЦ "Институт имени Н.Е. Жуковского"
e-mail: avkan@nrczh.ru

НИКИТИН Юрий Викторович – научный сотрудник ФИЦ ИУ РАН,
e-mail: yuri.v.nikitin@gmail.com

ХОРОШИЛОВ Александр Алексеевич – доктор технических наук, профессор НИУ МАИ, ведущий научный сотрудник ФИЦ ИУ РАН, старший научный сотрудник 27 ЦНИИ МО РФ
e-mail: khoroshilov@mail.ru

ВНИМАНИЮ ЧИТАТЕЛЕЙ!

ИЗДАНИЕ УДК

УНИВЕРСАЛЬНАЯ ДЕСЯТИЧНАЯ КЛАССИФИКАЦИЯ
АЛФАВИТНО-ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ
в 2-х томах

Алфавитно-предметный указатель (АПУ) к 4-му полному изданию УДК на русском языке:

Том I содержит АПУ от буквы А до Н;

Том II содержит АПУ от буквы М до Я и указатель латинских наименований к классам УДК 56 Палеонтология, 57 Биологические науки, 58 Ботаника, 49 Зоология, 61 Медицинские науки.

АПУ содержит около 100 000 понятий, представленных в полных таблицах УДК.

При его составлении были учтены изменения, опубликованные в Выпусках № 1 – 6 «Изменения и дополнения к УДК»

Для подписки необходимо направить заявку для оформления счета по адресу:

125190, Россия, Москва, ул. Усиевича, 20, ВИНТИ РАН

Телефоны: 499 155-42-85, 499 151-78-61

E-mail: feo@viniti.ru

<http://www.udcc.ru>

ВНИМАНИЮ ЧИТАТЕЛЕЙ!

ВИНИТИ РАН, как единственный в России владелец лицензии Консорциума УДК, предлагает издания УДК полного четвертого издания на русском языке в печатном и электронном виде:

1. Таблицы УДК

УДК. Том I Общая методика применения УДК. Вспомогательные таблицы. Основные таблицы. Общий отдел. Алфавитно-предметный указатель к Общему отделу

УДК. Том II 1/3 Философия. Психология. Религия. Богословие. Общественные науки (только электронное издание)

УДК. Том III 5/54 Математика. Естественные науки (только электронное издание)

УДК. Том IV 55/59 Геологические и биологические науки (только электронное издание)

УДК. Том V 6/61 Медицинские науки (только электронное издание)

УДК. Том VI (часть 1) 6/621 Прикладные науки. Технология. Инженерное дело (только электронное издание)

УДК. Том VI (часть 2) 622/629 Техника. Инженерное дело (только электронное издание)

УДК. Алфавитно-предметный указатель к т. VI (1 и 2 части) (только электронное издание)

УДК. Том VII 63/65 Сельское хозяйство. Домоводство. Управление предприятием (только электронное издание)

УДК. Том VIII 66 Химическая технология. Химическая промышленность. Пищевая промышленность. Металлургия. Родственные отрасли (только электронное издание)

УДК. Том IX 67/69 Различные отрасли промышленности и ремесел. Строительство (только электронное издание)

УДК. Том X 7/9 Искусство. Спорт. Филология. География. История.

УДК. АПУ (с в о д н ы й) к полному 4-му изданию

УДК. Изменения и дополнения. Выпуск 2 (к т.т. 1–3) (только электронное издание)

УДК. Изменения и дополнения. Выпуск 3 (к т.т. 1–6) (только электронное издание)

УДК. Изменения и дополнения. Выпуск 4 (к т.т. 1–7) (только электронное издание)

УДК. Изменения и дополнения. Выпуск 5 (к т.т. 1–10)

УДК. Изменения и дополнения. Выпуск 6 (к т.т. 1–10)

УДК. Изменения и дополнения. Выпуск 7 (к т.т. 1–10), 2017 г. (только электронное издание)

Для подписки необходимо направить заявку по адресу:

125190, Россия, Москва, ул. Усиевича, 20, ВИНТИ РАН

Телефоны: 499-155-42-85, 499-151-78-61

E-mail: feo@viniti.ru