

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕ-
ДЕРАЦИИ

федеральное государственное автономное образовательное учреждение высшего об-
разования
«Санкт-Петербургский политехнический университет Петра Великого»

На правах рукописи

Подпись аспиранта

Баскаков Дмитрий Евгеньевич

ФИО АСПИРАНТА

*МЕТОДЫ И АЛГОРИТМЫ ИНТЕЛЛЕКТУАЛЬНОЙ ПОДДЕРЖКИ ПРИ
ПРИНЯТИИ УПРАВЛЕНЧЕСКИХ РЕШЕНИЙ В
ТЕХНИЧЕСКИХ СИСТЕМАХ*

наименование темы научно-квалификационной работы (заглавными буквами)

2.3.1. «Системный анализ, управление и обработка информации»

отрасль науки (шифр и наименование научной специальности)

09.06.01 «Информатика и вычислительная техника»

наименование направленности (шифр и наименование направления)

Академическая степень **Исследователь. Преподаватель-исследователь**

НАУЧНЫЙ ДОКЛАД

Научный руководитель: Доктор технических наук, профессор,
проректор СПбПУ по международной
деятельности,
Арсеньев Дмитрий Германович

Санкт-Петербург, 2022

Научный доклад выполнен в Институте компьютерных наук и технологий федерального государственного автономного образовательного учреждения высшего образования «Санкт-Петербургский политехнический университет Петра Великого»

Научный руководитель:

Доктор технических наук, профессор, проректор СПбПУ по международной деятельности,
Арсеньев Дмитрий Германович

Рецензент:

Кандидат технических наук, старший научный сотрудник, Институт проблем транспорта им. Н. С. Соломенко РАН,
Кузнецова Елена Юрьевна

С научным докладом можно ознакомиться в библиотеке ФГАОУ ВО «Санкт– Петербургский политехнический университет Петра Великого» и на сайте Электронной библиотеки СПбПУ по адресу: <http://elib.spbstu.ru>.

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность работы

На текущий момент в России развивается несколько крупных транспортных проектов федерального масштаба. Одной из ключевых задач является проект освоения Северного Морского Пути (СМП)—кратчайший морской путь между Европейской частью России и Дальним Востоком. В планах экономического развития—навигационное и логистическое оснащение СМП, грузопоток по которому согласно указу Президента РФ должен составить к 2024 году около 80 млн тонн.

Важно понимать, что активное развитие СМП создает не только новые возможности и вызовы для России, но и дополнительно генерирует и новые риски как геополитические, так и природные. Наиболее важным фактором успешного долгосрочного развития СМП является стратегия снижения антропогенной нагрузки на окружающую среду Арктики, обеспечение которой возможно лишь при создании системы комплексной экологической защиты и интеллектуального мониторинга. Интеллектуальная система управления экологической безопасностью (ИСЭБ) СМП в дальнейшем должна стать основой создания Комплексной системы экологической безопасности перспективной интеллектуальной мультимодальной транспортной системы Арктической зоны Российской Федерации

Данная интеллектуальная система управления экологической безопасностью должна включать в себя как традиционные проекты и модели комплексного промышленного экологического мониторинга, так и новые разработанные методы и алгоритмы интеллектуальной поддержки при принятии управленческих решений в технических системах.

Технология построения ИСЭБ для СМП предполагает использование тотального встраивания во все важные элементы системы датчиков и интеллектуальных исполнительных устройств. При этом информационно-

коммуникационная сеть ИСЭБ будет являться интерфейсом и платформой сбора и первичной обработки данных в режиме реального времени. Собранные данные, предполагается, после первичной обработки будут храниться в специальных базах данных на крупных транспортно-логистических узлах.

Действующая сеть государственного экологического мониторинга признана Правительством РФ неэффективной, так как она не позволяет Федеральным органам исполнительной власти оперативно реагировать на экологические происшествия, предотвращать экологические катастрофы и предсказывать некие возможные проблемные ситуации в будущем. Поэтому и принято решение о создании комплексной информационной системы мониторинга состояния окружающей среды. Первые элементы данной системы планируются к внедрению уже в 2024 году с дальнейшей модернизацией системы до 2030 года. На этапе ввода системы в эксплуатацию должны быть реализованы в том числе и интеллектуальные алгоритмы принятия решений на основе следующих видов аналитики:

1. Дескриптивная в части интеллектуального анализа данных экологического мониторинга. Что случилось?
2. Диагностическая аналитика для анализа причин произошедшего.
3. Предиктивная аналитика для целей прогнозирования экологической обстановки с применением перспективных моделей оценки качества экологической среды.
4. Предписывающая аналитика, которая с применением интеллектуальных алгоритмов позволит нам впредь исключать возможные экологические проблемы в будущем.

Таким образом, исследование методов и алгоритмов интеллектуальной поддержки при принятии управленческих решений в технических системах экологического мониторинга становится крайне актуальной и своевременной задачей в соответствии со стратегией развития Северного морского пути.

Цель исследования состоит в снижении рисков и стоимости создания интеллектуальных систем управления экологической безопасностью; разработка новых методов, моделей и алгоритмов принятия управленческих решений на основе машинного обучения; выявление перспективных направлений в области проектирования таких систем; разработка методологической основы проектирования таких систем и доказательство их эффективности.

Для достижения данных целей в диссертации решаются следующие **задачи**:

1. Производится анализ современных киберфизических систем мониторинга и управления/
2. Исследуются современные методов и практики создания интеллектуальных систем экологического мониторинга и управления.
3. Исследуются методы и алгоритмы машинного обучения, которые могут быть использованы при создании интеллектуальных систем экологического мониторинга и управления.
4. Исследуются метрики качества и надежности интеллектуальных систем управления экологической безопасностью.
5. Разрабатываются новые методы, подходы и архитектуры интеллектуальных систем управления экологической безопасностью с применением практик машинного обучения.

Значимым практическим результатом работы является использование предложенных методов и алгоритмов при проектировании перспективной интеллектуальной системы управления экологической безопасностью Северного морского пути (СМП), которая в дальнейшем будет интегрирована в мультимодальный транспортный коридор РФ.

Объекты, (предмет) и методы исследования

Объектом исследования в работе являются интеллектуальные распределенные киберфизические системы мониторинга и управления экологической безопасностью, которые эксплуатируются в особых условиях Крайнего Севера РФ.

Предметом исследования являются методы мониторинга и управления экологической безопасностью, включая методы интеллектуального адаптивного управления на основе машинного обучения.

Методы исследования используются теоретические и эмпирические. В качестве общенаучных теоретических методов применялся метод формализации путем отображения и систематизации существующих структур и методов в виде рисунков, формул и схем. Также широко применялся метод идеализации для роста абстракции исследуемых систем и исключения избыточных свойств и систем. Методы индукции и дедукции использовались для выводов в процессе исследования и заключений частного характера в процессе теоретических и практических изысканий, а анализ и синтез являлся основой всех теоретических изысканий и заключений.

В качестве эмпирических методов применяется метод моделирования и аналогии для описания происходящих процессов, метод сравнения и измерения использовался для сравнения объектов по наиболее существенным признакам.

Научная новизна

Целью исследования является разработка теоретических основ построения интеллектуальных сетей мониторинга и управления экологической безопасностью мультимодальных транспортных систем на пример интеллектуальной системы управления экологической безопасностью Северного морского пути РФ.

В рамках проведенного исследования достигнуты следующие результаты:

1. Разработаны основные методы и принципы построения распределенных интеллектуальных систем управления экологической безопасностью в особых климатических условиях Севера.
2. Разработана математическая модель учета антропогенных и климатических факторов, которые оказывают всестороннее воздействие на эксплуатацию интеллектуальных систем управления экологической безопасностью в особых климатических условиях Севера.
3. Сформулированы условия и метрики эксплуатации интеллектуальных систем управления экологической безопасностью, основанные на решении задачи выбора оптимального управления в зависимости от целевой функции и ограничений в виде равенств и неравенств.
4. Разработаны методы и алгоритмы машинного обучения, которые применяются при принятии управленческих решений в ИСЭБ.
5. Разработана система критериев анализа эффективности эксплуатации ИСЭБ.
6. Разработана метамодель перспективной распределенной интеллектуальной системы управления экологической безопасностью Северного морского пути.
7. Разработан метод интеллектуального управления ИСЭБ в виде программы для ЭВМ.

Положения, выносимые на защиту

1. Системный подход к адаптивному управлению распределенными киберфизическими системами экологической безопасности.
2. Методология и алгоритмы поддержки принятия решений в интеллектуальных системах управления экологической безопасностью.
3. Метамодель распределенной интеллектуальной системы

управления экологической безопасностью Северного морского пути на основе заданных критериев и метрик.

4. Метод управления распределенной киберфизической интеллектуальной системой экологической безопасности Северного морского пути.

Соответствие специальности научных работников. Научные результаты соответствуют паспорту специальности 2.13.1 «Системный анализ, управление и обработка информации»: Разработка проблемно-ориентированных систем управления, принятия решений и оптимизации технических, организационно-технических и информационных систем (п.13); методы и алгоритмы интеллектуальной поддержки при принятии управленческих решений в технических, организационно-технических и информационных системах (п.15); разработка специального математического и алгоритмического обеспечения систем анализа, оптимизации, управления, принятия решений и обработки информации (п.16); методы и алгоритмы прогнозирования и оценки эффективности, качества и надежности сложных технических, организационно-технических и информационных систем (п.17); методы визуализации, трансформации и анализа информации (п. 19);

Теоретическая и практическая значимость

Теоретическая значимость работы заключается в создании нового системного подхода к созданию и управлению интеллектуальными системами управления экологической безопасностью, основанного на принципах распределенности и иерархической связанности.

В рамках созданного подхода разработаны:

1. Методы и алгоритмы поддержки принятия управленческих решений интеллектуальной системой управления экологической

безопасностью.

2. Методы обработки, агрегации и управления данными экологического мониторинга.
3. Методы интеллектуального и визуального анализа экологических данных на основе машинного обучения.
4. Перспективные архитектуры построения интеллектуальных систем управления экологической безопасностью.

Практическая значимость работы определяется возможностью использования разработанных методов и моделей для построения интеллектуальными системами управления экологической безопасностью для Северного морского пути РФ. Результаты работы позволяют:

1. Обеспечить полноту решения при проектировании интеллектуальных систем управления экологической безопасностью в особых климатических условиях Крайнего Севера РФ.
2. Формулировать строгие критерии и метрики оценки функционирования систем экологической безопасности.
3. Снизить время на проектирование и развертывание интеллектуальной системы управления экологической безопасностью.
4. Обеспечить оперативное решение задач интеллектуальной поддержки при принятии управленческих решений в интеллектуальной системе управления экологической безопасностью.

Апробация работы

Практические результаты исследований стали важной частью проекта по созданию интеллектуальной системы управления экологической безопасностью Северного морского пути РФ. Архитектура верхнего уровня интеллектуальной системы управления экологической безопасностью для Северного морского пути представлена на рисунке 1:



Рисунок 1. Архитектура интеллектуальной системы управления экологической безопасностью для СМП

На основании архитектуры верхнего уровня интеллектуальной системы экологической безопасности для СМП в работе предложена перспективная архитектура программной реализации данной системы (рисунок 2).

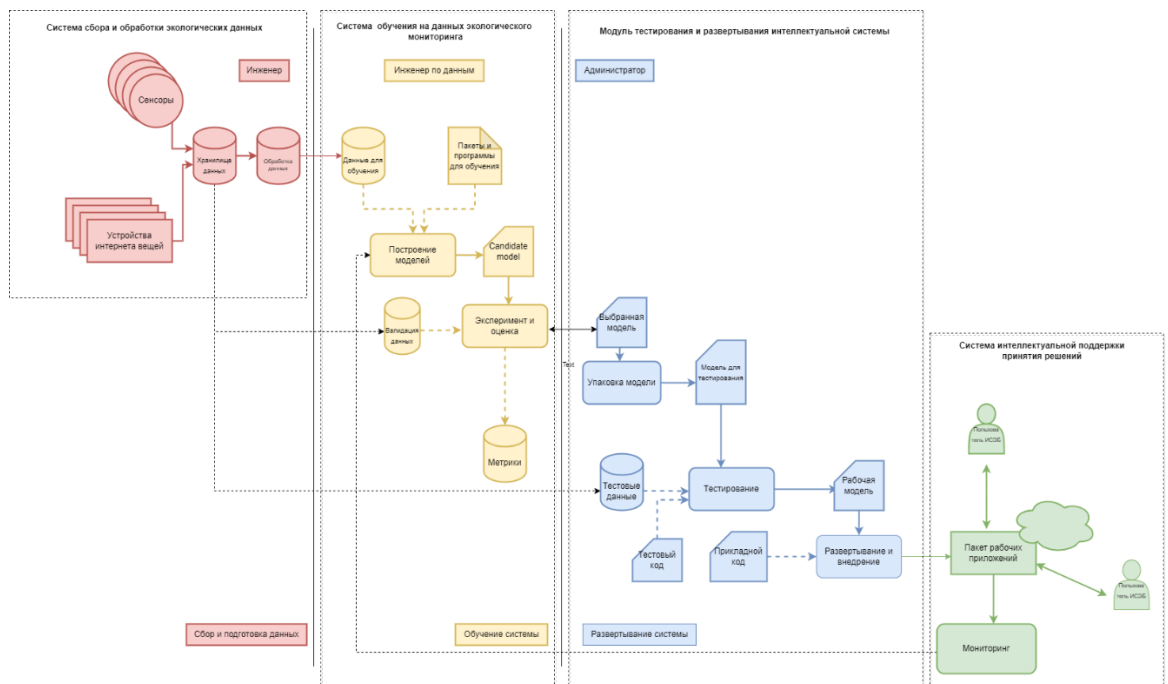


Рисунок 2. Программная архитектура ИСЭБ

В качестве системы управления процессами сбора и обработки данных предполагается использования двигателя Camunda с BPMN (Business Process Model Notation) и DMN (Decision Model and Notation).

BPMN- процесс сбора, обработки и визуализации данных будет выглядеть следующим образом:

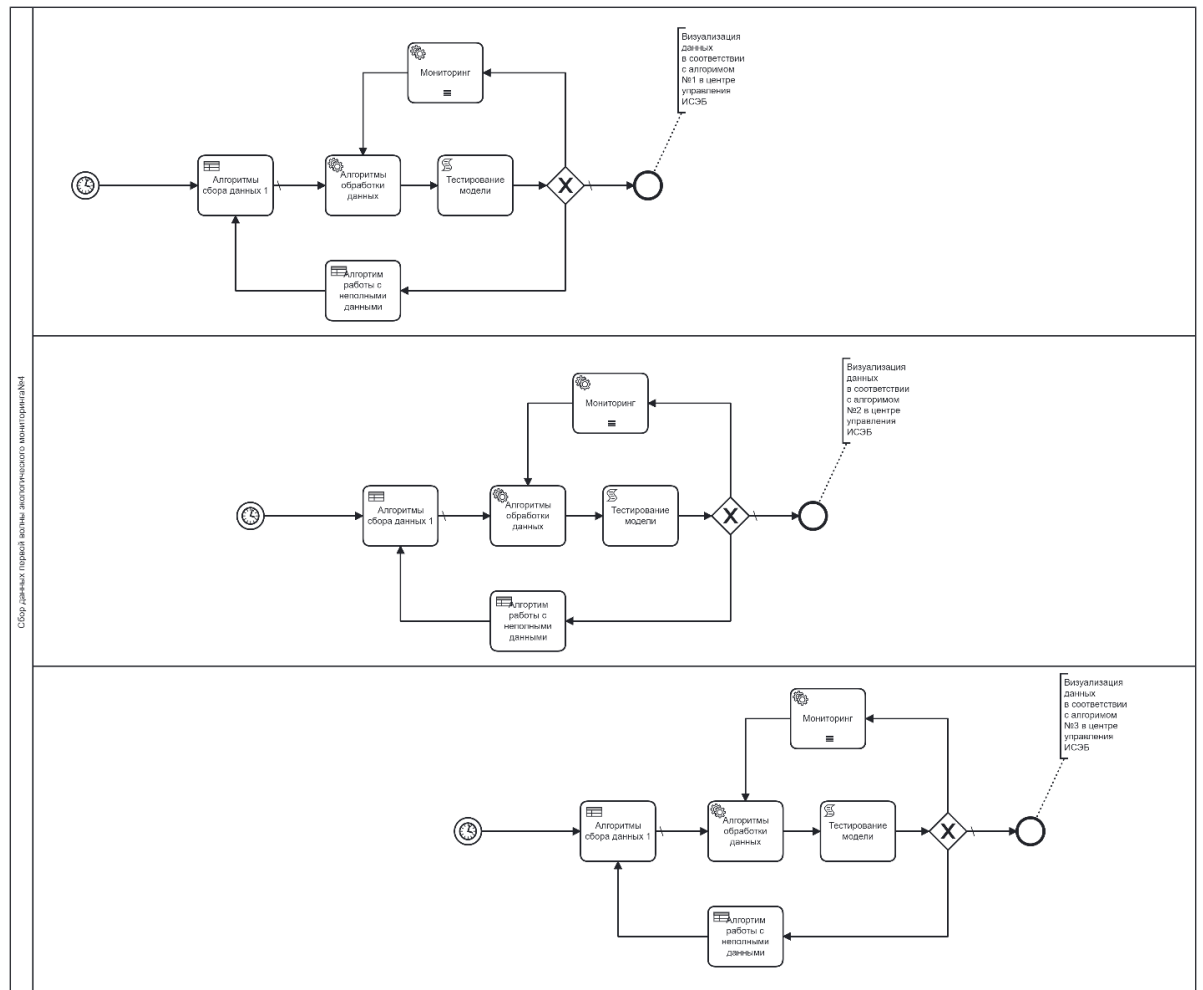


Рисунок 3. BPMN-процесс сбора и визуализации данных ИССЭ

Данные экологического мониторинга предполагается собирать волнами с применением DMN и Camunda. Принцип волновой сборки (рисунок 4) предполагает этапность, временные ограничения и иных требования к получаемым данным, которые и описываются с помощью DMN-нотаций в виде заданных алгоритмов (рисунок 5).

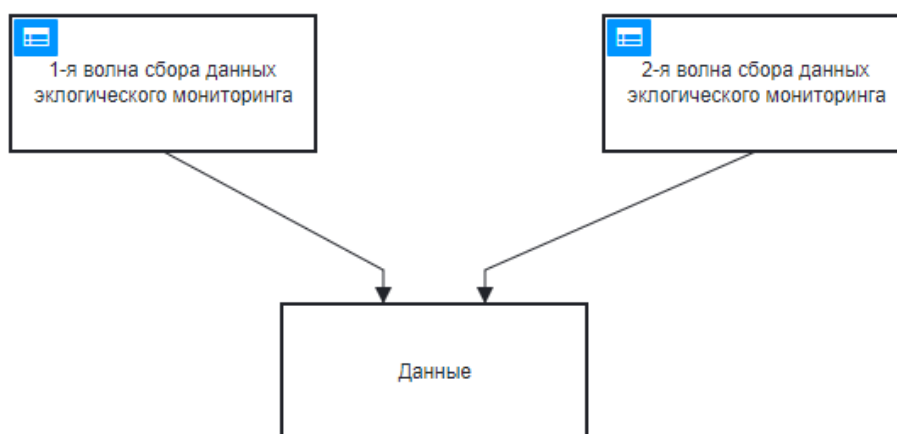


Рисунок 4. Волны сбора экологических данных

Пример алгоритма 1-й волны выглядит следующим образом:

		When		And		And		And		Then	
		IoTs	Time	Number	Boolean	Algorithms	Annotations				
		number	dayTimeDuration	number	boolean	number					
1	Sensor 1		10:00-10:05	> 100	1&2 equals True	a	Алгоритм сбора данных с сенсора №1				
2	Sensor 2		10:00-10:05	< 0.01	True	b	Алгоритм сбора данных с сенсора №2				
3	Sensor 3		10:00-10:05	=11	False	c	Алгоритм сбора данных с сенсора №3				
4	-	-	-	-	-	-					
5	-	-	-	-	-	-					
6	Sensor n		10:00-10:05	>0.001	1&2&3 equals True	n	Алгоритм сбора данных с сенсора №4				
+	-	-	-	-	-	-					

Рисунок 5. 1-я волна сбора данных экологического мониторинга

Использование Camunda для автоматизации процесса сбора и обработки данных оправдано по целому ряду причин и является новой предложенной практикой в процесс создания ИСЭБ:

1. Автоматизация рутинных задач, которые ранее решал человек.
2. Алгоритмизация процесса сбора и обработки данных в соответствии с созданной логикой и правилами применения.
3. Выполнение задач в циклах, по расписанию, с иными условиями.

Важным практическим результатом работы является и то, что процесс тестирования и внедрения интеллектуального обеспечения занимает большую часть релизного цикла разработки, что заметно снижает возможную стоимость

сопровождения программного обеспечения, которая составляет обычно от 40 до 70 процентов стоимости разработки, а часто в некоторых случаях и просто ее превосходит.

Результаты работы апробированы на следующих конференциях и научных мероприятиях:

1. 11-я Российская мультikonференция по проблемам управления. Конференции «Информационные технологии в управлении» (ИТУ-2018).
2. International Conference on Cyberphysical systems & control (CPS&C–2019), 10-12 June 2019, St. Petersburg, Russia.
3. International Scientific Conference "Telecommunications, Computing and Control ("TELECCON-2019).
4. Computational Science and Its Applications - ICCSA 2019. 19th International Conference, Saint Petersburg, Russia, July 1–4, 2019.
5. Конференция «информационные технологии в управлении» (ИТУ-2020).
6. Конференция «Математическая теория управления и ее приложения» (МТУиП-2020).
7. Computational Science and Its Applications – ICCSA 2021. 21th International Conference, Cagliari, Italy.
8. 2nd International Conference Cyber-Physical Systems and Control, St.Petersburg,Russia, 2021.

Публикации

По теме диссертации опубликовано 11 работ, в том числе 3 из них в журналах списка рекомендованных ВАК и 8, входящие в международные реферативные системы цитирования и базы данных Scopus и Web of Science.

1. Контрактные сетевые протоколы взаимодействия когнитивных агентов в распределенных интеллектуальных системах управления. Арсеньев Д. Г., Баскаков Д. Е., Шкодырев В. П. Санкт-Петербург : СПб.: АО «Концерн «ЦНИИ «Электроприбор», 2018. 11-я РОССИЙСКАЯ МУЛЬТИКОНФЕРЕНЦИЯ ПО ПРОБЛЕМАМ УПРАВЛЕНИЯ. Т. МАТЕРИАЛЫ КОНФЕРЕНЦИИ «ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ В УПРАВЛЕНИИ» (ИТУ-2018), стр. 628.
2. Deep Predictive Control. Baskakov, Dmitry. Saint-Petersburg : б.н., 2019. International Scientific Conference "Telecommunications, Computing and Control ("TELECCON-2019).
3. Distributed Ledger Technology and Cyber-Physical Systems. Multi-agent Systems. Concepts and Trends. Dmitry Arsenjev, Dmitry Baskakov, and Vyacheslav Shkodyrev. б.м. : Springer, 2019. Computational Science and Its Applications - ICCSA 2019. T. 19th International Conference, Saint Petersburg, Russia, July 1–4, 2019 Proceedings, Part II, стр. 805.
4. Secure Machine Intelligence and Distributed Ledger. Dmitry Arseniev, Dmitry Baskakov, and Vyacheslav Shkodyrev. Cagliari, Italy : Springer, 2021. Computational Science and Its Applications – ICCSA 2021 . T. Lecture Notes in Computer Science, 12951, стр. 749.
5. Graph Model Approach to Hierarchy Control Network. Dmitry G. Arseniev, Dmitry Baskakov, and Vyacheslav P. Shkodyrev. St. Petersburg, Russia : Springer, 2019. The International Conference on Cyber-Physical Systems and Control (CPS&C'2019). T. Lecture Notes in Computer Science, 11620, стр. 778.
6. Интеллектуальные графовые модели обработки связанных данных. Д.Е. Баскаков, В.П. Шкодырев, Д.Г. Арсеньев. Санкт-Петербург : б.н., 2020. КОНФЕРЕНЦИЯ «ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ В УПРАВЛЕНИИ» (ИТУ-2020).
7. Иерархическая кластеризация: алгоритмы и результаты. Д.Г. Арсеньев, Д.Е. Баскаков, В.П. Шкодырев. Санкт-Петербург : б.н. КОНФЕРЕНЦИЯ

«МАТЕМАТИЧЕСКАЯ ТЕОРИЯ УПРАВЛЕНИЯ И ЕЕ ПРИЛОЖЕНИЯ»
(МТУиП-2020).

8. Mutation Testing of Software Quality on the System-Level. Dmitry Arseniev, Dmitry Baskakov, Jussi Kasurinen, Vyacheslav Shkodyrev, Alexey Mergasov. St.Petersburg : б.н., 2021. 2nd International Conference Cyber-Physical Systems and Control-CPS&C'2021.

9. Software Engineering Principles Apply to Artificial Intelligence Systems. Dmitry Arseniev, Dmitry Baskakov, Jussi Kasurinen, Vyacheslav Shkodyrev. St.Petersburg, Russia : б.н., 2021. 2nd International Conference Cyber-Physical Systems and Control - CPS&C'2021.

11. Robust models of distance metric learning by interval-valued training data . Lev Utkin, Andrei Konstantinov, Dmitry Baskakov, Vladimir Muliukha, Natalia Politaeva. St.Petersburg,Russia : б.н., 2021. 2nd International Conference Cyber-Physical Systems and Control.

Представление научного доклада: основные положения

1. Краткое состояние отрасли интеллектуального экологического мониторинга
2. Архитектура интеллектуальной системы экологической безопасности (ИСЭБ)
3. Программная архитектура ИСЭБ
4. Применение BPMN и DMN нотаций для моделирования процессов и описания алгоритмов волнового сбора информации
5. Основные результаты и заключения

СОДЕРЖАНИЕ РАБОТЫ

(по главам)

Во введении рассмотрены современные подходы, методы и практики в области построения интеллектуальных систем управления. Исследованы основные парадигмы, проведен критический анализ существующих подходов. Выявлены слабые места и недостатки в действующих методологиях.

В первой главе проведен обзор литературы по методам и алгоритмам интеллектуальной поддержки при принятии управленческих решений в технических системах. Данная глава состоит из четырех основных разделов, а именно:

1. Обзор литературы в области методик и практик построения систем на основе искусственного интеллекта.
2. Обзор литературы по алгоритмам машинного и глубокого обучения, которые получили наибольшее распространение при построении индустриальных, коммерческих решений с открытым исходным кодом.
3. Обзор литературы по практикам и методам разработки и создания предметной области (Domain Model) на основе принципов программной инженерии при проектировании корпоративных программных приложений в том числе и с применением искусственного интеллекта.
4. Обзор литературы по метрикам качества программных приложений, которые могут быть применены при проектировании интеллектуальных решений, а именно: доступность, возможность развертывания, энергоэффективность, интегрируемость, модифицируемость, производительность, безопасность, тестируемость и удобство использования.

Во второй главе описываются методики эмпирического и теоретического исследований. В качестве основного метода эмпирического

исследования автор применяет сравнение методов, подходов и алгоритмов при проектировании интеллектуальных систем. Дополнительно сравниваются и результаты экспериментов и тестов из открытых источников и литературы.

В качестве теоретических методов исследования автор использует анализ и синтез методов и алгоритмов, применяет классификацию и моделирование протекающих процессов. Выводы и заключения носят индукционный характер. Большинство значимых теоретических результатов формализуется в виде рисунков, формул и схем.

В третьей главе приводятся теоретические и практические результаты проведенного исследования. В процессе исследования выявлено, что от качества работы методов и алгоритмов зависят целые перспективные направления и даже отрасли, а именно:

1. Пилотирование автомобилей без водителя
2. Финансы и банковские системы
3. здравоохранение
4. Сбор и обработка данных

При критическом анализе литературы выявлена необходимость создания новых подходов и практик в области проектирования и эксплуатации систем на основе искусственного интеллекта с применением машинного обучения. Обозначены направления будущего, например, Machine Learning Engineering или совершенно новая профессия- Machine Learning Engineer.

С практической точки зрения инженерия машинного обучения (Machine Learning Engineering) как новая дисциплина при разработке интеллектуальных систем требует несколько иных, а, зачастую, и вовсе новых подходов при разработке и построении систем, что не позволяет в теории и практике опираться на классические подходы и концепции программной инженерии [2]. В работе систематизированы основные отличия, которые присущи системам и моделям на основе искусственного интеллекта:

1. *Данные.* Данные являются основой всех систем искусственного интеллекта. При их недостаточном количестве или должном качестве просто не имеет смысла даже планировать и создавать подобные системы. Управление данными является отличительной чертой всех систем на основе искусственного интеллекта. Для подобных целей используется много специальных средств, например: TFX's ExampleGen, StatisticGen, SchemaGen, ExampleValidator.
2. *Модель.* Как разработчик программного обеспечения создает код, так и инженер по машинному обучению в итоге на основе данных, своих знаний и подходов создает в конечном итоге модель. При этом, как и данные, модели постоянно эволюционируют и развиваются со временем. Средства, которые здесь могут использоваться, например, TFX's Evaluator и Infra Validator.
3. *Объединение фрагментов.* Как программный инженер создает свой продукт, используя разные библиотеки, так и инженер по машинному обучению используя разные данные, фрагменты разного кода и библиотек, анализируя разные библиотеки и алгоритмы в итоге создает совершенно новую систему принятия решений на основе искусственного интеллекта. Здесь могут использоваться следующие программные средства: TFX's ExampleGen, Transform, Trainer и Tuner, а также TensorFlow Hub.
4. *Трудности тестирования и отладки.* Особенностью систем на основе машинного обучения является их весьма трудоемкая отладка, где ошибка в конкретном наборе данных или небольшая неточность в настройке, например, гиперпараметров, может приводить в итоге к полностью неработоспособной системе, что заметно отличает этот процесс от классической разработки программных систем. Для этих целей можно использовать TFX's ML Metadata, например.

5. *Непрерывное обучение.* Системы на основе искусственного интеллекта должны непрерывно обучаться в этом изменчивом мире. Данные и модели постоянно обновляются, появляются новые метрики и подходы, происходит динамическое обновление окружающей среды, что требует создания механизмов как непрерывного развития таких систем, так и внедрения моделей версионирования и обработки изменений, например.

В работе предлагается описывать любую техническую интеллектуальную систему на четырех уровнях, а именно:

1. Математический уровень (используется для построения математических моделей работы системы)
2. Архитектурный уровень (используется для описания функционирования составных частей системы)
3. Экономический уровень (стоимостное описание системы)
4. Физический уровень (описание составных частей системы во времени и пространстве)

Но с практической точки зрения исследования интеллектуальных систем проводятся чаще всего на втором, архитектурном уровне. В процессе исследования постулируется, что на текущий момент нет систем закрытого типа на основе искусственного интеллекта. Таким образом трудно провести грань между испытаниями системы и ввода ее в эксплуатацию и ее дальнейшей непрерывной доработкой и модернизацией, поскольку данные этапы работы становятся неразличимы на уровне процессов.

Основные практические выводы, которые можно сделать на основании проведенного исследования, следующие:

1. *Область применений.* Основной фокус текущих исследований и применений систем на основе машинного обучения это: автомобильный транспорт, финансы, здравоохранение и обработка данных. При этом большинство подобных систем разработано либо

в лабораториях (41%) или непосредственно в больших компаниях (33%). Большинство значимых исследований и результатов достигается как раз либо в рамках исследовательских лабораторий больших компаний, либо как результат проектной работы в этих же компаниях. Случаи успешных решений в рамках небольших компаний или стартапов очень малы (несколько процентов).

2. *Качество и тестирование.* Тестирование реальных систем на основе машинного обучения представляет собой весьма нетривиальную задачу как с точки зрения уникальности доменных областей, так и отсутствием единого подхода к практикам оценки качества и тестирования с точки зрения методологий и знаний. Важно отметить, что данный раздел представляет собой важнейший большой пласт будущей теоретической научной работы.
3. *Управление данными.* В отличие от традиционных технических систем системы с применением машинного обучения существенно зависят как от качества данных, так и их предварительной обработки и дальнейшей интерпретации. Следует добавить и важность постоянного обновления данных, непрерывного процесса обучения, а также этические и правовые нюансы использования таких систем, например, в медицине или в автомобильном транспорте.
4. *Инфраструктура.* Машинное обучение требует создания отдельной инфраструктуры, стоимость которой существенно выше в сравнении с традиционными программными средами.
5. *Программная и аппаратная сложность.* Важно понимать, что системы с применением машинного обучения обладают всеми свойствами традиционных программных систем, но при этом дополнительно существенно сложнее в силу свойств и методов, которые присущи системам на основе искусственного интеллекта.

6. *Новая отрасль науки.* Современный контекст развития систем с применением машинного обучения требует выделения в отдельную отрасль знаний некоторых подходов в области разработки и эксплуатации таких систем, а именно: тестирование, качество программного обеспечения, управления данными, разработка моделей, инфраструктура и работа с требованиями применительно исключительно к системам искусственного интеллекта.

В процессе исследования систем с искусственным интеллектом описаны и некоторые ограничения, которые при практическом применении присущи как данным методам, так и в целом интеллектуальным системам:

1. *Задача.* Убедитесь, что задача, которую вы решаете, должна иметь интеллектуальную поддержку принятия решений. У вас достаточно данных для обучения такой системы и нет более простых вариантов решения. Важно понимать, что подобные решения с использованием интеллектуальных алгоритмов на основе машинного обучения, зачастую, существенно дороже обычных систем управления.
2. *Команда.* Прежде чем приступить к разработке систем и методов с применением искусственного интеллекта и машинного обучения убедитесь, что у вас есть соответствующая команда специалистов, которая располагает соответствующими знаниями и опытом в исследуемой предметной области.
3. *Данные.* Убедитесь, что вы располагаете необходимым набором данных для обучения системы. Вы можете получать эти данные и далее, у вас нет проблем с безопасностью и доступом к этим данным и методам их хранения. В случае отрицательного ответа на данные вопросы лучше не приступать к такой работе, ибо доступность и безопасность данных является ключевой проблемой большинства систем на основе машинного обучения.

4. *Алгоритм.* Вы располагаете доступным алгоритмом или моделью, либо у вас есть команда и сотрудники, которые способны решать подобные задачи.
5. *Архитектура.* Вы располагаете архитектурой будущего решения или понимаете необходимость такого подхода. Важно знать и понимать высокую зависимость систем искусственного интеллекта от интеграционных потоков со смежными системами, качества данных и их соответствия моделям.
6. *Верификация и валидация.* Вы располагаете инструментами и практиками будущей верификации и валидации вашего решения. При этом именно валидация будет иметь решающее значение для оценки качества вашего проекта.
7. *Неопределенность.* Большинству решений на основе машинного обучения и искусственного интеллекта присуща достаточно высокая степень неопределенности, основанная прежде всего на интерпретации получаемых результатов. Вас данная степень неопределенности изначально устраивает и/или вы с ней знакомы. В противном случае лучше отложить применение таких систем в пользу традиционных систем без машинного обучения.
8. *Обслуживание системы.* Системы искусственного интеллекта весьма дороги в обслуживании в сравнении с традиционными системами. Стоимость данного обслуживания вы учитываете, умеете считать и данные оценки вас устраивают.

Далее в работе указывается несколько важных особенностей создания и построения практических методов и алгоритмов интеллектуальной поддержки при принятии решений в технических системах.

Архитектурные особенности

Традиционный цикл создания программного продукта с применением систем машинного обучения выглядит следующим образом:

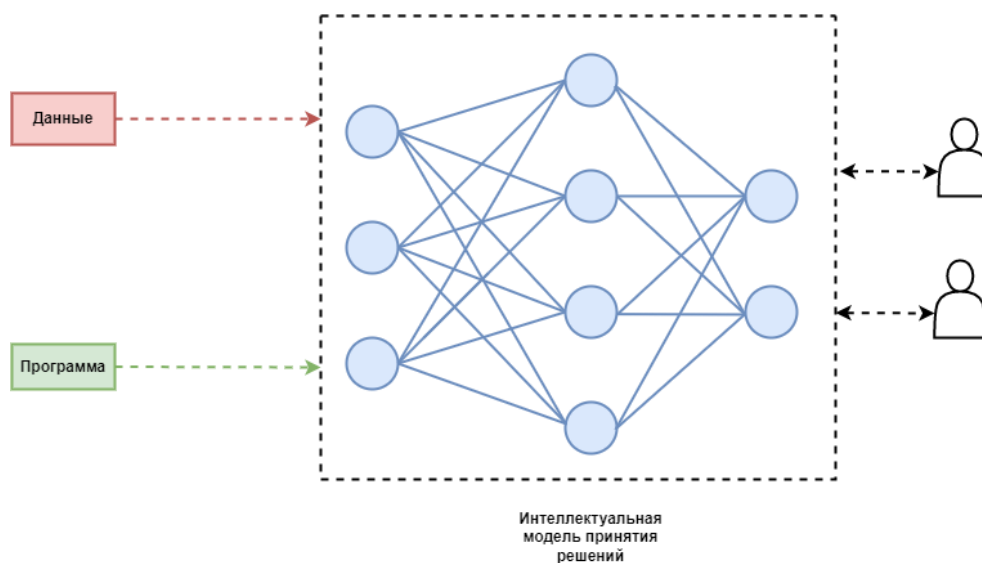


Рисунок 6. Традиционный подход машинного обучения

В случае построения интеллектуальных систем управления взаимодействие разработки и данных происходит в режиме реального времени и выглядит следующим образом:

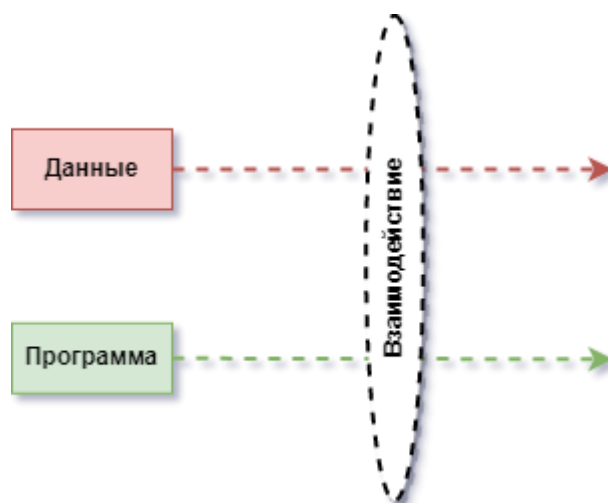


Рисунок 7. Индустриальное машинное обучение

Источники данных

Интеллектуальные системы управления работают со многими источниками данных, что заметно отличает от обычных систем управления:

1. Пользовательские данные, включая данные для машинного обучения и тестирования.
2. Системные данные, которые генерируются системой (логи,

результаты обучения, предсказания моделей).

3. Внешние источники данных, которые генерируются различными сторонними сервисами для успешной работы системы

На самом деле в академическом сообществе принято обращать внимание только на непосредственно алгоритмы машинного обучения и принятия решений, что в корне не соответствует инфраструктуре таких решений в целом.

При проектировании ИСЭБ СМП непосредственно код и алгоритмы, например, занимают не более 5% от всей инфраструктуры интеллектуальной системы (рисунок 3).



Рисунок 8. Инфраструктура ИСЭБ

Традиционный процесс разработки систем и приложений на основе машинного обучения начинается с преобразования данных из различных источников в набор размеченных данных. Обычно данную процедуру выполняют инженеры. Далее в работу включаются специалисты по данным и только потом разработчики. Процесс разработки выглядит согласно рисунку 9:

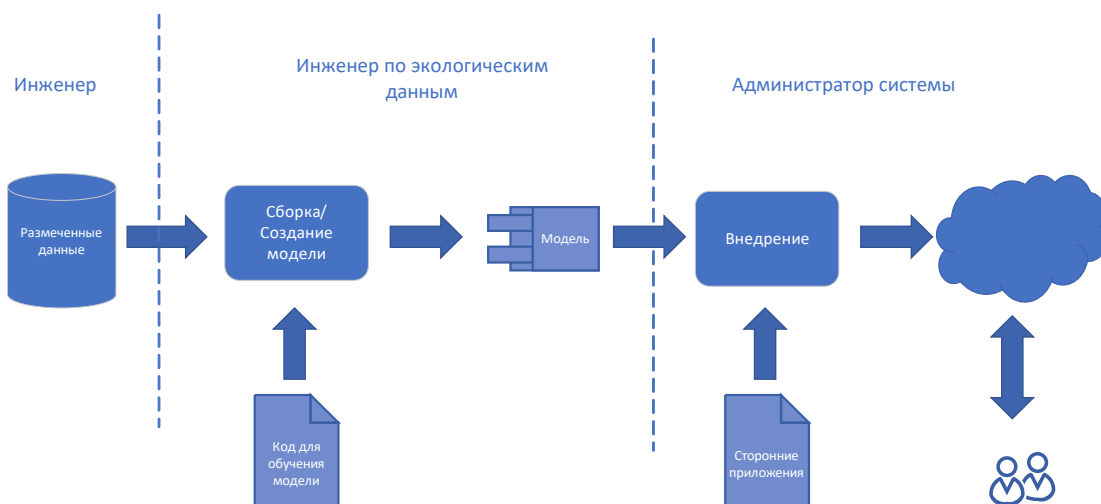


Рисунок 9. Цикл разработки ИСЭБ на основе машинного обучения

В работе предложен новый подход к созданию методов и алгоритмов интеллектуальной поддержки при принятии управленческих решений в технических системах (рисунок 10). Ключевым моментом является внедрение современных практик программной инженерии при проектировании и внедрении интеллектуальных систем на примере проектирования ИСЭБ.

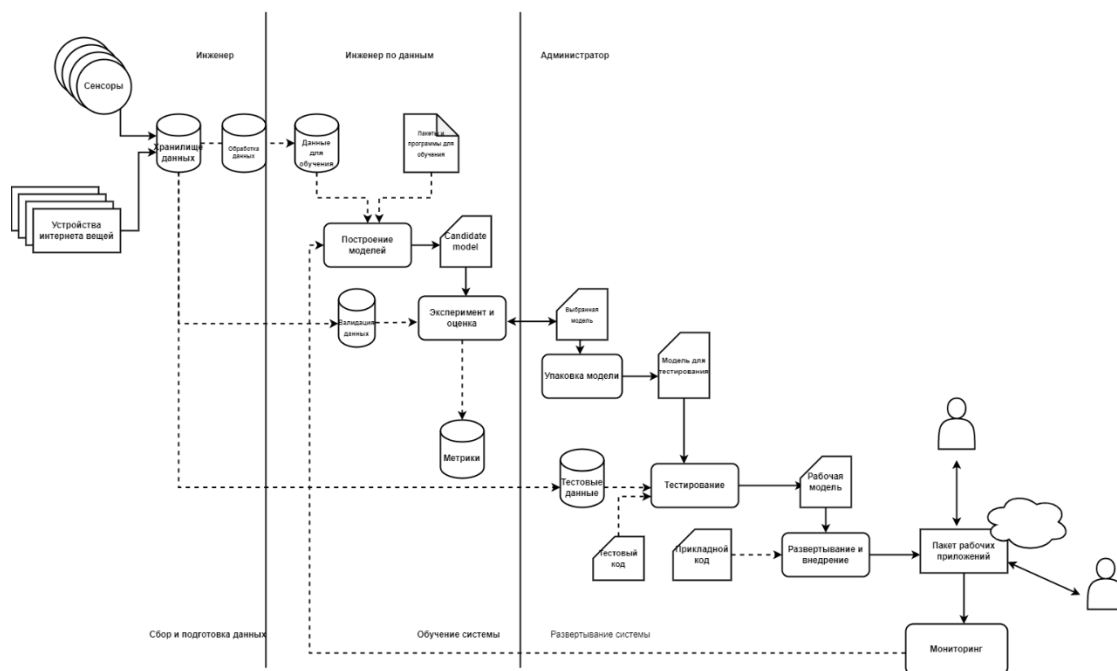


Рисунок 10. Процесс проектирования и внедрения ИСЭБ

Человеко-центристская модель машинного обучения

Комплекс ИСЭБ представляет собой реализацию парадигмы активного вовлечения реального пользователя в совершенствование алгоритмов и методов интеллектуальной поддержки при принятии управленческих решений в технических системах с созданием человеко-центристской модели машинного обучения (Human Centric Model Learning, HCML) (рисунок 11):



Рисунок 11. Человеко-центристская модель машинного обучения

Программно-аппаратная часть ИСЭБ должна обладать и поддерживать следующие архитектурные характеристики, критерии и метрики, которые обеспечат заданное качество функционирования системы, а именно: доступность (*Availability*), масштабируемость (*Scalability*), устойчивость к ошибкам (*Fault tolerance*), производительность (*Performance*), надежность (*Reliability*), безопасность (*Security*), гибкость (*Elasticity*), возможность развертывания (*Deploy ability*), тестируемость (*Testability*), восстанавливаемость (*Recoverability*) и обучаемость (*Learnability*).

В работе вводится интегральная целевая функция качества системы:

$$\begin{aligned}
 Q_1(x) = & A(x_a) + S(x_s) + F(x_f) + P(x_p) + R1(x_r) + S(x_s) + E(x_e) \\
 & + D(x_d) + T(x_t) + R2(x_{r2}) + L(x_l)
 \end{aligned}
 \quad (1)$$

где: $A(x_a)$ — функция доступности системы, $S(x_s)$ — функция масштабируемости системы, $F(x_f)$ — функция устойчивости к ошибкам, $P(x_p)$ — функция производительности системы, $R1(x_r)$ — функция надежности системы, $S(x_s)$ — функция безопасности системы, $E(x_e)$ — функция гибкости системы, $D(x_d)$ — функция развертывания системы, $T(x_t)$ — функция тестируемости системы, $R2(x_{r2})$ — функция восстанавливаемости системы и $L(x_l)$ — функция обучаемости системы.

Классическая CAP (*Consistency, Availability, Partition Tolerance*)-теорема выглядит графически следующим образом:

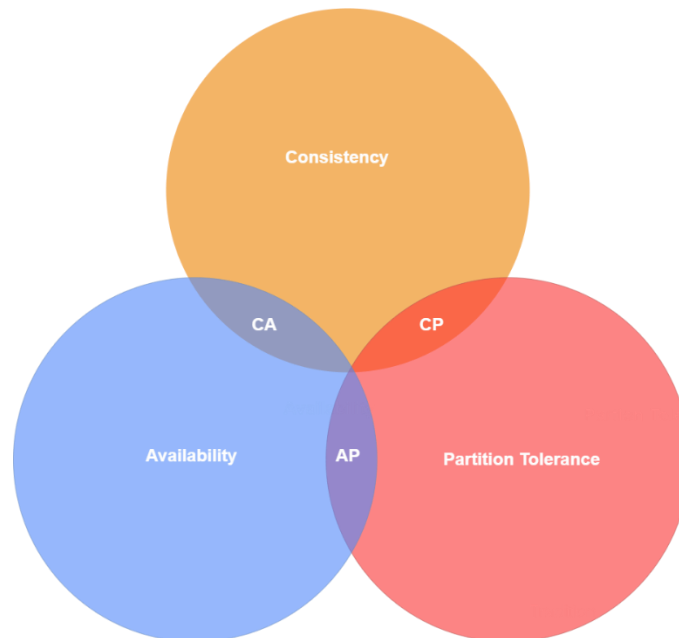


Рисунок 12. CAP-Теорема

В исследовании предлагается переформулировать CAP-теорему для распределенных систем (рисунок 13), которая постулирует, что система не может предоставлять более чем 2 из трех характеристик, путем введения функций, а именно:

$$Q_2(x) = C(x_c) + P(x_p) \parallel P(x_p) + A(x_a) \parallel C(x_c) + A(x_a) \quad (2)$$

где: $C(x_c)$ — функция согласованности данных (*Consistency*), $A(x_a)$ — функция

доступности данных и системы (*Availability*) и $P(x_p)$ — функция устойчивости системы к сбоям/разделению (*Partition tolerance*).

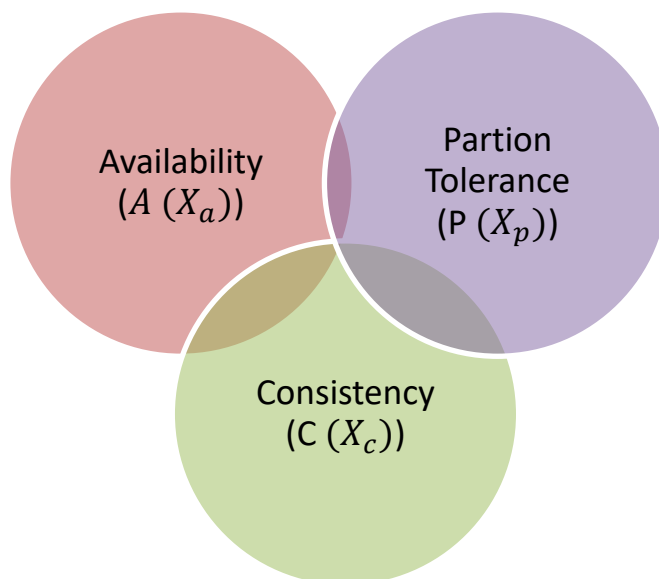


Рисунок 13. Переформулируемая CAP-теорема

Развитием CAP-теоремы является ее расширение в виде PACELC (Partitioning, Availability, Consistency, End, Latency, Consistency)—теоремы, которую математически сформулируем следующим образом:

$$Q_3(x) = \begin{cases} C(x_c) + P(x_p) \parallel P(x_p) + A(x_a) \parallel C(x_c) + A(x_a), & E \neq 0 \\ L(x_l) + C(x_c), & E = 0 \end{cases} \quad (3)$$

где: $C(x_c)$ — функция согласованности данных (*Consistency*), $A(x_a)$ — функция доступности данных и системы (*Availability*), $P(x_p)$ — функция устойчивости системы к сбоям/разделению (*Partition tolerance*) и $L(x_l)$ — функция задержки передачи данных (*Latency*).

Также в исследовании предлагается дополнить интегральную целевую функцию качества ИСЭБ следующими характеристиками, которые имеют

вторичное значение, но, тем не менее, оказывают существенное влияние на функционирование системы в целом:

1. Конфигурируемость (*Configurability*)
2. Расширяемость (*Extensibility*)
3. Возможность установки (*Installability*)
4. Эффективность и повторное использование (*Leverageability/reuse*)
5. Локализация (*Localization*)
6. Ремонтопригодность (*Maintainability*)
7. Портативность (*Portability*)
8. Поддержка (*Supportability*),
9. Обновление (*Upgradeability*).

Основные выявленные проблемы при проектировании и создании ИСЭБ:

1. Требования доступности сетевых сервисов 365*24*7.
2. Высокое время отклика.
3. Ограниченная пропускная способность.
4. Сетевая безопасность.
5. Архитектурная и топологическая изменчивость.
6. Сложность администрирования.
7. Высокая стоимость.
8. Сетевая неоднородность.

Практические и теоретические выводы, которые дополнительно можно сделать на основании проведенного исследования, следующие:

- ✓ *Связывающий код (Glue Code)*. Использование пакетов машинного обучения с открытым исходным кодом или общедоступных решений часто несет в себе риск существенного увеличения стоимости проекта по причине использования обобщенных метрик и алгоритмов без применения к конкретной доменной области. Универсальность в машинном обучении не приводит к значимым результатам,

зачастую. В зрелых системах только около 5% процентов всего машинного кода относится в той или иной степени к алгоритмам. Весь остальной код примерно на 95% является связующим с другими системами и модулями. Создавать изначально собственное решение более эффективно и часто менее затратно в перспективе.

- ✓ *Проблемы тестирования и интеграции.* В системах машинного обучения промышленного назначения существенная роль отводится процессам тестирования и интеграции. В случае отсутствия детальной проработки архитектуры такой системы на самом раннем этапе процессы дальнейшего сбора и обработки данных превращаются в трудноразрешимую задачу, особенно в режиме реального времени. Сквозные и нагрузочные тесты в случае использования каких-то универсальных решений с открытым кодом, зачастую, просто невозможно использовать. Единственным разумным путем в таком случае является создание и проектирование системы изначально с нуля.

В работе систематизирована и предложена практика создания интеллектуальных систем управления на основе искусственного интеллекта на примере ИСЭБ с внедрением современных моделей и методик проектного менеджмента. Предложен рабочий процесс по созданию таких систем (рисунок 14).



Рисунок 14. Дорожная карта проекта создания ИСЭБ СМП

В исследовании на примере ИСЭБ делается важнейший теоретический и практический вывод о том, что эксплуатация систем на основе искусственного интеллекта происходит уже на этапе, фактически, ее мониторинга и пред

продуктовой эксплуатации. То есть достижение требуемых метрика качества на данном этапе, фактически, нас приближает к системе с реальными характеристиками эксплуатации.

В проведенном исследовании делается вывод и о том, что процесс разработки, эксплуатации и развития интеллектуальных систем превращается в непрерывный итеративный процесс (рисунок 15).

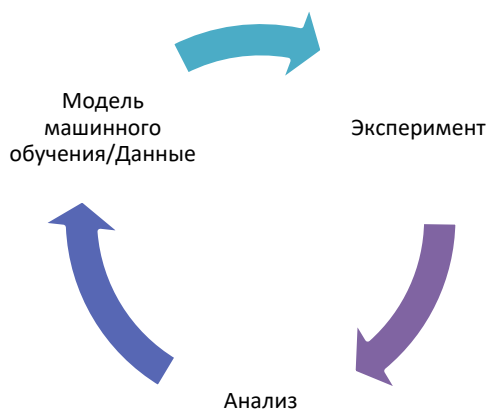


Рисунок 15. Итеративный процесс работы с данными и моделью в машинном обучении

Также в работе предложен инженерный подход по итеративному процессу анализа производительности таких систем, рассмотрены метрики качества функционирования таких систем в том числе с применением практик программной инженерии (рисунок 16).

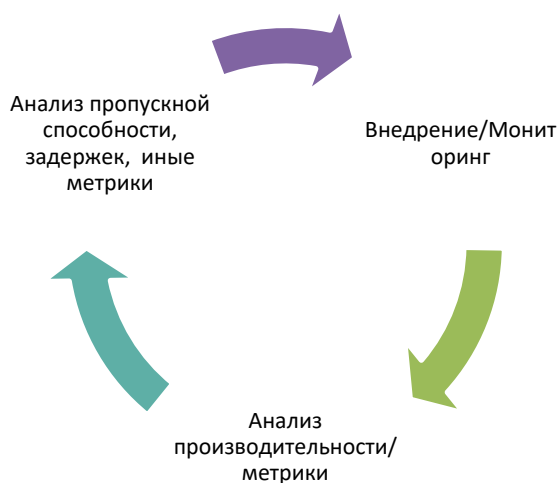


Рисунок 16. Итеративный процесс деплоя и анализа производительности систем на основе машинного обучения

В исследовании делается вывод и на предмет создания методологии работы с требованиями при программной реализации таких систем, которая существует в программной инженерии.

С точки зрения научной новизны в работе проведена базовая систематизация требований при создании интеллектуальных систем управления, в каких случаях применение искусственного интеллекта, например, оправданно, а в каких случаях от создания и эксплуатации таких систем разумнее воздержаться на примере проектирования ИСЭБ СМП.

В работе исследуется использование нескольких архитектурных подходов. В качестве предположения рекомендовано отказаться от практики применения архитектуры on-line вывода при создании ИСЭБ (рисунок 17) в пользу использования архитектуры на основе виртуальных машин (рисунок 18) или на основе контейнеризации с применением приложения KuberFlow от Kubernetes или DockerSwarm для небольших команд и проектов (рисунок 19).

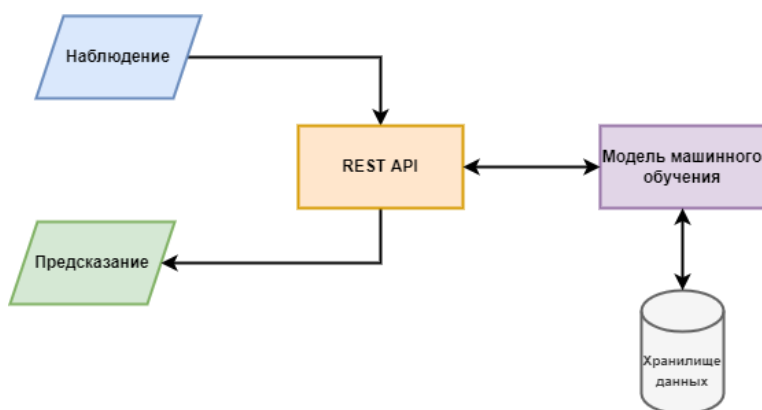


Рисунок 17. Онлайн вывод интеллектуальной системы

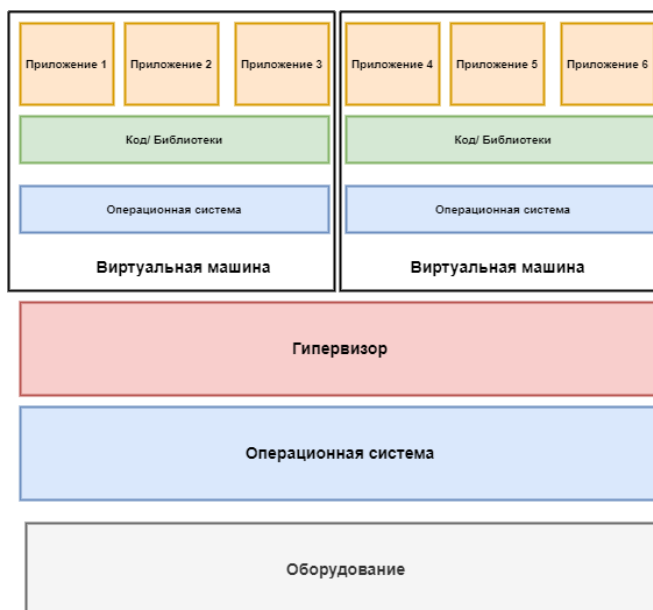


Рисунок 18. Архитектура на основе виртуальных машин

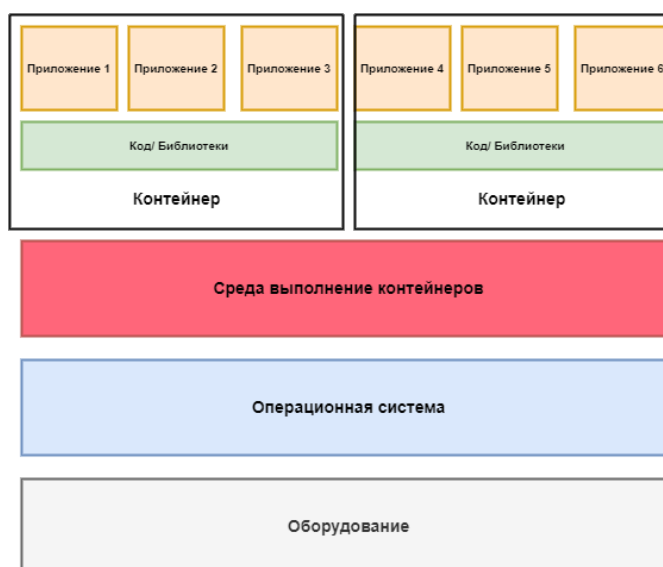


Рисунок 19. Контейнерная архитектура для машинного обучения

Автор анализирует преимущества использования контейнерной архитектуры для интеллектуальной системы экологического мониторинга, делая соответствующие выводы, а именно:

- ✓ Небольшое количество дополнительных программных средств (операционные системы, например)

- ✓ Высокий уровень абстракции при построении интеллектуальных систем
- ✓ Простое и быстрое развертывание и масштабирование

В качестве возможного перспективного инженерного решения предложена расширенная архитектура интеллектуальной системы управления экологической безопасностью реального времени для Северного морского пути РФ, которая способна обрабатывать в онлайн режиме стриминговые данные от различных сервисов и датчиков, например устройств интернета-вещей или каких-то сенсоров (рисунок 20).

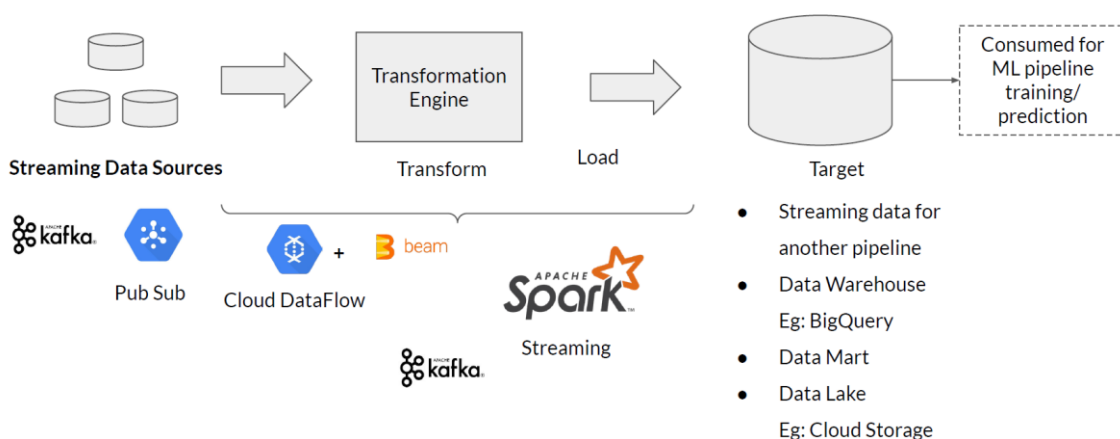


Рисунок 20. Процесс стриминговой обработки данных для экологического мониторинга

На основании проведенного исследования при проектировании ИСЭБ предлагаются к использованию следующие практики построения перспективных систем интеллектуального принятия решений, а именно:

1. *Непрерывная интеграция (Continuous Integration, CI)*: тестирование и проверка кода, компонентов, данных, данных схемы и модели на постоянной основе.
2. *Непрерывная доставка (Continuous Delivery, CD)*: не только развертывание одного программного пакета или сервиса, а система, которая автоматически развертывает другой сервис (например,

модель обработки и визуализации экологических данных с сенсоров).

3. *Непрерывное обучение (Continuous Training, CT)*: новый процесс, уникальный для систем ML, который автоматически переобучает модели-кандидаты для тестирования и обслуживания.
4. *Непрерывный мониторинг (Continuous Monitoring, CM)*: выявление ошибок в системе и мониторинг данных экологического вывода и показателей производительности модели, привязанных к результатам ИСЭБ.

Также в работе предлагается применение обновленного подхода к построению жизненного цикла систем на основе машинного обучения на примере ИСЭБ (рисунок 21).

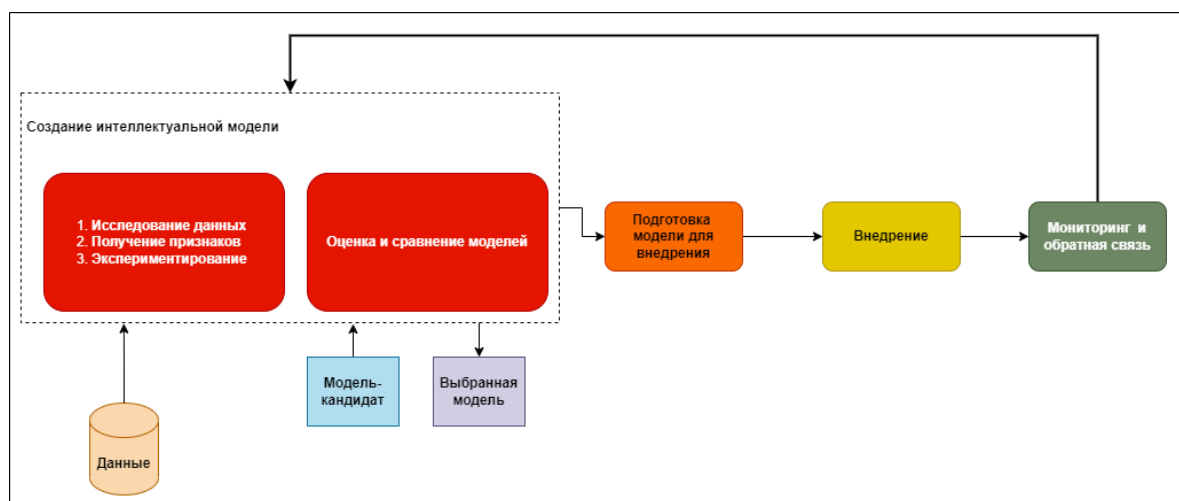


Рисунок 21. Жизненный цикл проекта на основе машинного обучения

Результаты и их обсуждение

В рамках проведенного исследования выполнена работа по созданию новых методов и алгоритмов интеллектуальной поддержки принятия управленческих решений в технических системах. Предложены новые научные методы и инженерные практики проектирования и внедрения в эксплуатацию

киберфизической распределенной интеллектуальной системы управления экологической безопасностью реального времени для Северного морского пути РФ

Наиболее значимым практическим результатом исследования является вывод о том, что непосредственно алгоритмы и программный код интеллектуального принятия решений составляют лишь небольшую часть (около 5%) от всей области методологий, работ и знаний интеллектуальной поддержки при принятии решений в технических системах на примере ИСЭБ для СМП.

Важным в исследовании является и то, что намечены будущие направления теоретических и практических работ в области построения методов и алгоритмов интеллектуальной поддержки принятия решений в технических системах, например, создание единой методологии оценки качества функционирования таких систем или методов и практик тестирования их работоспособности.

Заключение

Основные результаты работы:

1. Разработан системологический подход к адаптивному управлению распределенными киберфизическими системами экологической безопасности.
2. Разработана методология и алгоритмы поддержки принятия решений в интеллектуальных системах управления экологической безопасностью.
3. Разработана метамодель распределенной киберфизической интеллектуальной системы управления экологической безопасностью Северного морского пути на основе заданных критериев и метрик.
4. Разработан метод управления распределенной киберфизической интеллектуальной системой экологической безопасности Северного морского пути РФ.

Теоретические и практические результаты исследования стали важной частью проекта по созданию интеллектуальной системы управления экологической безопасностью Северного морского пути РФ, которая может быть интегрирована в мультимодальный коридор транспортной безопасности РФ. Дополнительно важно отметить, что теоретические и практические результаты работы могут быть использованы:

1. При проектировании информационных и коммуникационных процессов мониторинга экологической обстановки для различных транспортных мод (автомобильной, водной, авиационной и железнодорожной).
2. Технологии обработки Больших данных об экологической обстановке на инфраструктурных транспортных объектах различной модальности и преобразования их в точные знания об экологической обстановке на транспорте;

Список работ, опубликованных по теме научно-квалификационной работы (диссертации)

Публикации в изданиях, рецензируемых ВАК

1. *Контрактные сетевые протоколы взаимодействия когнитивных агентов в распределенных интеллектуальных системах управления*. Арсеньев Д. Г., Баскаков Д. Е., Шкодырев В. П. Санкт-Петербург : СПб.: АО «Концерн «ЦНИИ «Электроприбор», 2018. 11-я РОССИЙСКАЯ МУЛЬТИКОНФЕРЕНЦИЯ ПО ПРОБЛЕМАМ УПРАВЛЕНИЯ. Т. МАТЕРИАЛЫ КОНФЕРЕНЦИИ «ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ В УПРАВЛЕНИИ» (ИТУ-2018), стр. 628.
2. *Deep Predictive Control*. Baskakov, Dmitry. Saint-Petersburg : б.н., 2019. International Scientific Conference "Telecommunications, Computing and Control ("TELECCON-2019).
3. *Distributed Ledger Technology and Cyber-Physical Systems. Multi-agent Systems. Concepts and Trends*. Dmitry Arsenjev, Dmitry Baskakov, and Vyacheslav Shkodryev. б.м. : Springer, 2019. Computational Science and Its Applications - ICCSA 2019. Т. 19th International Conference, Saint Petersburg, Russia, July 1–4, 2019 Proceedings, Part II, стр. 805.

4. *Secure Machine Intelligence and Distributed Ledger*. Dmitry Arseniev, Dmitry Baskakov, and Vyacheslav Shkodyrev. Cagliari, Italy : Springer, 2021. Computational Science and Its Applications – ICCSA 2021 . T. Lecture Notes in Computer Science, 12951, стр. 749.
5. *Graph Model Approach to Hierarchy Control Network*. Dmitry G. Arseniev, Dmitry Baskakov, and Vyacheslav P. Shkodyrev. St. Petersburg, Russia : Springer, 2019. The International Conference on Cyber-Physical Systems and Control (CPS&C'2019). T. Lecture Notes in Computer Science, 11620, стр. 778.
6. *Интеллектуальные графовые модели обработки связанных данных*. Д.Е. Баскаков, В.П. Шкодырев, Д.Г. Арсеньев. Санкт-Петербург : б.н., 2020. КОНФЕРЕНЦИЯ «ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ В УПРАВЛЕНИИ» (ИТУ-2020).
7. *Иерархическая кластеризация: алгоритмы и результаты*. Д.Г. Арсеньев, Д.Е. Баскаков, В.П. Шкодырев. Санкт-Петербург : б.н. КОНФЕРЕНЦИЯ «МАТЕМАТИЧЕСКАЯ ТЕОРИЯ УПРАВЛЕНИЯ И ЕЕ ПРИЛОЖЕНИЯ» (МТУиП-2020).
8. *Mutation Testing of Software Quality on the System-Level*. Dmitry Arseniev, Dmitry Baskakov, Jussi Kasurinen, Vyacheslav Shkodyrev, Alexey Mergasov. St.Petersburg : б.н., 2021. 2nd International Conference Cyber-Physical Systems and Control-CPS&C'2021.
9. *Software Engineering Principles Apply to Artificial Intelligence Systems*. Dmitry Arseniev, Dmitry Baskakov, Jussi Kasurinen, Vyacheslav Shkodyrev. St.Petersburg, Russia : б.н., 2021. 2nd International Conference Cyber-Physical Systems and Control - CPS&C'2021.
11. *Robust models of distance metric learning by interval-valued training data* . Lev Utkin, Andrei Konstantinov, Dmitry Baskakov, Vladimir Muliukha, Natalia Politaeva. St.Petersburg,Russia : б.н., 2021. 2nd International Conference Cyber-Physical Systems and Control.

Аспирант

(подпись)

Баскаков

Д.Е.

