

Организация хранения данных функционирования объектов киберфизических систем

Д.А. Коростелев^{1,2}, А.Г. Подвесовский^{1,2}✉, А.А. Захарова¹

¹ Институт проблем управления им. В.А. Трапезникова РАН,
г. Москва, 117997, Россия

² Брянский государственный технический университет,
г. Брянск, 241035, Россия

Ссылка для цитирования

Коростелев Д.А., Подвесовский А.Г., Захарова А.А. Организация хранения данных функционирования объектов киберфизических систем // Программные продукты и системы. 2024. Т. 37. № 3. С. 334–343. doi: 10.15827/0236-235X.142.334-343

Информация о статье

Группа специальностей ВАК: 2.3.8

Поступила в редакцию: 16.04.2024

После доработки: 05.05.2024

Принята к публикации: 14.05.2024

Аннотация. В статье представлены результаты исследований в области организации и программной поддержки хранения данных функционирования объектов киберфизических систем, получаемых в ходе экспериментов. Определены основные особенности их хранения: использование специализированных форматов, хранение больших объемов данных разной степени структурированности, поддержка универсальных структурированных форматов данных, хранение сведений о структуре самих экспериментов и сценариях их проведения. Проанализированы технологии построения систем хранения и специализированные форматы данных, учитывающие эти особенности. Предложен подход к построению программной системы хранения и предобработки экспериментальной информации о функционировании объектов киберфизической системы. Архитектура и особенности реализации системы рассмотрены на примере хранения данных экспериментов с группами совместно действующих беспилотных мобильных средств при решении задач перемещения в пространстве и транспортировки грузов. Архитектура предполагает наличие серверного приложения и веб-клиента, взаимодействующего с ним с помощью API-интерфейса, построенного на основе принципов REST. В состав серверного приложения также входят подсистема предобработки результатов экспериментов и конвертер форматов источников данных. Для хранения экспериментальной информации предложено использовать двухуровневую структуру модели данных. Верхний уровень предназначен для хранения информации о структуре и параметрах экспериментов и реализован в виде БД в системе PostgreSQL. На нижнем уровне осуществляется хранение потоковой информации, получаемой от первичных источников данных. Данный уровень реализован в виде облачного хранилища, построенного на основе Яндекс Облака с использованием технологии S3. Результаты апробации разработанной системы при проведении реальных экспериментов подтвердили обоснованность выбранных архитектурных решений. Практическая значимость выполненных исследований и разработок заключается в развитии подходов к построению масштабируемых программных систем хранения данных функционирования объектов киберфизических систем, что способствует систематизации задач и сценариев управления этими объектами при разработке новых технологий и инструментов управления.

Ключевые слова: киберфизическая система, объекты киберфизических систем, экспериментальные данные, форматы данных, модель данных, СУБД, REST API, S3-хранилище

Благодарности. Исследование выполнено при поддержке РФФ, проект № 23-19-00342, <https://rscf.ru/project/23-19-00342/>

Введение. Одно из ключевых направлений цифровизации любой сферы деятельности связано с развитием и повсеместным внедрением *киберфизических систем* (КФС). В общем случае под КФС понимаются комплексные системы, объединяющие информационно-вычислительные и физические компоненты, находящиеся в тесном взаимодействии между собой и функционирующие посредством интеграции информационно-коммуникационных, вычислительных и физических технологий. Примерами КФС могут служить различные технические устройства и комплексы: робототехнические системы, беспилотные транспортные системы и средства, системы умного города, сенсорные

системы и др. Характерной чертой современных КФС является взаимодействие большого числа объектов различных типов, как полностью автономных, так и частично автоматизированных и управляемых операторами. За счет автономности объектов КФС, а также интеллектуальности и распределенного характера управления ими становятся возможными решение принципиально новых задач автоматизации производственных и технологических процессов, а также значительное упрощение, ускорение и удешевление подходов к решению.

С другой стороны, развитие индустрии КФС, повышение их технического уровня и уровня интеллектуальности обуславливают актуаль-

ность задач управления такими системами, включая управление как отдельными объектами КФС, так и группой взаимодействующих объектов, в том числе в условиях неполноты и разномасштабности информации, а также распределенности и разнородности источников ее получения.

Исследования в области создания новых и совершенствования существующих подходов, методов и средств управления объектами КФС неизбежно связаны с необходимостью проведения большого количества и вычислительных, и натуральных экспериментов. При этом основу управления объектами КФС всегда составляют процессы сбора, обработки и анализа разнородной информации, получаемой непосредственно от самих объектов, а также от внешних источников данных, осуществляющих мониторинг их работы. В подобной ситуации становятся важными задачи накопления и упорядочения информации, получаемой от объектов КФС и в ходе постановки экспериментов, и при выполнении реальных заданий.

Таким образом, необходим инструмент, позволяющий осуществлять сбор, предварительную обработку и хранение информации, получаемой от объектов КФС в ходе выполнения экспериментов с ними. Важным требованием к указанному инструменту является поддержка набора форматов данных, учитывающих особенности представления и хранения такой информации и обеспечивающих возможность ее дальнейшего использования для обработки и анализа в рамках создания методов и построения систем управления КФС.

Анализ особенностей хранения экспериментальной информации, получаемой от объектов КФС

С точки зрения процессов сбора и хранения информации КФС можно рассматривать как совокупность мобильных и стационарных источников разнородных данных, относящихся к различным областям. Для хранения таких данных, как правило, существуют специализированные форматы, традиционно применяемые в соответствующей области. Так, одним из наиболее распространенных форматов хранения данных в КФС, основанных на беспилотных объектах, является ROS Bag. Данный формат используется в системе ROS и обеспечивает хранение данных изображений, видеопотоков, лидарных данных и других сообщений от различных сенсоров и устройств [1, 2]. Помимо

этого, для хранения изображений и видеопотоков широко используются специально разработанные для этого графические и видеоформаты. Применительно к КФС можно отдельно отметить формат GeoTIFF [3, 4], предназначенный для хранения изображений с геопространственной привязкой – этот формат может применяться для хранения данных беспилотных транспортных систем, работающих в таких приложениях, как картография и навигация.

Таким образом, можно выделить *первую особенность* хранения экспериментальной информации, получаемой от объектов КФС, – возможность использования специализированных форматов, применяемых в соответствующих областях.

Информация, получаемая от КФС, характеризуется большими объемами данных различной степени структурированности с преобладанием полуструктурированных и неструктурированных данных. Для хранения подобных данных существуют специально разработанные подходы и форматы хранения [5]. Следует отметить формат HDF5. Он предназначен для хранения больших объемов данных с поддержкой иерархической структуры и системы метаданных. Известен опыт успешного применения данного формата для хранения информации, которая получается с датчиков беспилотных аппаратов [6, 7]. Кроме того, в качестве средств хранения больших объемов информации можно указать облачные хранилища, построенные на базе технологий AWS S3 (<https://aws.amazon.com/s3/>) и характеризующиеся высокой доступностью и масштабируемостью. Также популярным и востребованным подходом хранения данных сложной структуры и больших объемов является использование NoSQL СУБД [8, 9].

С учетом этого *вторая особенность* хранения экспериментальной информации, получаемой от объектов КФС, состоит в необходимости хранения больших объемов данных разной степени структурированности.

Наряду со специализированными форматами представления данных, получаемых от сенсоров и устройств, для хранения данных КФС также требуются универсальные структурированные форматы данных, такие как CSV, XML, JSON [10]. Они активно используются при анализе данных, а также удобны для передачи и обмена данными. Это обуславливает *третью особенность* хранения экспериментальной информации, получаемой от объектов КФС, – необходимость поддержки универ-

сальных структурированных форматов данных.

В контексте хранения информации, получаемой от объектов КФС при проведении экспериментов с ними, приходится сталкиваться с *четвертой особенностью* – необходимость хранения сведений о структуре самого эксперимента (сценариях его проведения, участвующих в нем объектах, используемых источниках данных и проч.).

Учитывая высокую степень разнородности информации, используемой в управлении объектами КФС, многообразие механизмов ее получения, а также форматов представления и хранения, а также принимая во внимание недостаток универсальных программных инструментов, поддерживающих хранение и преобразование такой информации при проведении экспериментов с объектами КФС, можно сделать вывод об актуальности исследований, направленных на создание методов и средств хранения и преобразования экспериментальной информации, получаемой от объектов КФС и используемой для моделирования сценариев управления этими объектами.

Фактически речь идет о проектировании и разработке системы поддержки хранения экспериментальной информации, получаемой от объектов КФС, с учетом выделенных особенностей хранения такой информации, способов ее получения и механизмов взаимодействия с данными.

Отдельно следует отметить требование, связанное с обеспечением удобства использования данных при их автоматизированной обработке, суть которого заключается в обеспечении возможности оперировать универсальными форматами данных при их анализе. Иными словами, требование состоит в предоставлении дополнительных инструментов преобразования информации путем ее конвертации в соответствующие форматы.

Предметная область и терминология

Описание предлагаемого подхода к хранению экспериментальной информации, получаемой от объектов КФС, и программных решений, реализующих данный подход, проведем на примере КФС, представляющих собой группы совместно действующих беспилотных мобильных средств различного уровня автономности, создаваемые для выполнения некой задачи или комплекса задач. В качестве задач рассмотрим перемещение мобильных средств

в пространстве и транспортировку грузов с их помощью.

Введем ряд понятий, которые далее будут использованы для выделения сущностей, необходимых при построении модели данных.

Объекты КФС, задействованные в решении целевой задачи, будем называть *актерами*. Примерами акторов в контексте рассматриваемой предметной области могут являться беспилотные мобильные средства различных типов (летательные аппараты, платформы и др.), объекты инфраструктуры (погрузчики, зарядные станции) и т.д. Предполагается, что каждый актер имеет некоторый уровень автономности, который может зависеть от конкретной выполняемой задачи и от характера взаимодействия с другими актерами.

Объединение задач, решаемых группой акторов, по некоторому целевому признаку будем называть *миссией*. Примером миссии может быть транспортировка разнородных грузов в маршрутной сети, содержащей источники и пункты назначения, с использованием комплекса беспилотных летательных аппаратов и наземных транспортных средств [11, 12].

Каждая миссия характеризуется своей *структурой*, определяемой типом и характером взаимосвязи решаемых задач, и *набором параметров*, которые можно задавать при планировании миссии, а также при необходимости варьировать в процессе ее исполнения.

Планом миссии будем называть ее структурное представление, в котором определены все параметры. В общем случае план миссии строится с учетом ее цели, начальных условий, ограничений и множества акторов. При этом построение плана осуществляется на двух уровнях: планирование заданий отдельным актерам (уровень задания, исполнительный) и планирование взаимодействия между актерами (уровень миссии, тактический). Также при построении плана могут быть заданы механизмы оценки эффективности выполнения миссии в целом или ее отдельных составляющих.

Под *сценарием управления* понимаем набор взаимосвязанных управляющих воздействий, направленных на реализацию плана миссии, предназначенный для обеспечения эффективного решения комплекса задач, с которыми связана миссия.

Наконец, *эксперимент* – специально организованная миссия, в рамках которой при реализации определенного сценария управления осуществляется сбор данных, характеризующих состояние и функционирование акторов,

а также объектов внешней среды в процессе исполнения миссии. Таким образом, для эксперимента план миссии и сценарии управления дополняются требованиями к составу, источникам и форматам собираемых в ходе эксперимента данных.

Проектирование и реализация программной системы

Для хранения и предобработки данных функционирования объектов КФС разработана программная система на основе клиент-серверной архитектуры (рис. 1). В качестве серверной подсистемы выступает .NET-приложение, предоставляющее для взаимодействия с ней API-интерфейс, построенный с использованием принципов REST (<https://restapitutorial.ru/>).

Применение REST API позволяет в полной мере задействовать возможности протокола HTTP за счет применения различных поддерживаемых им методов для разных типов действий над сущностями, используемыми в приложении. Основные методы HTTP:

- GET для получения как сведений о конкретной сущности, так и списка соответствующих сущностей с возможностью фильтрации и постраничной навигации; результат запроса

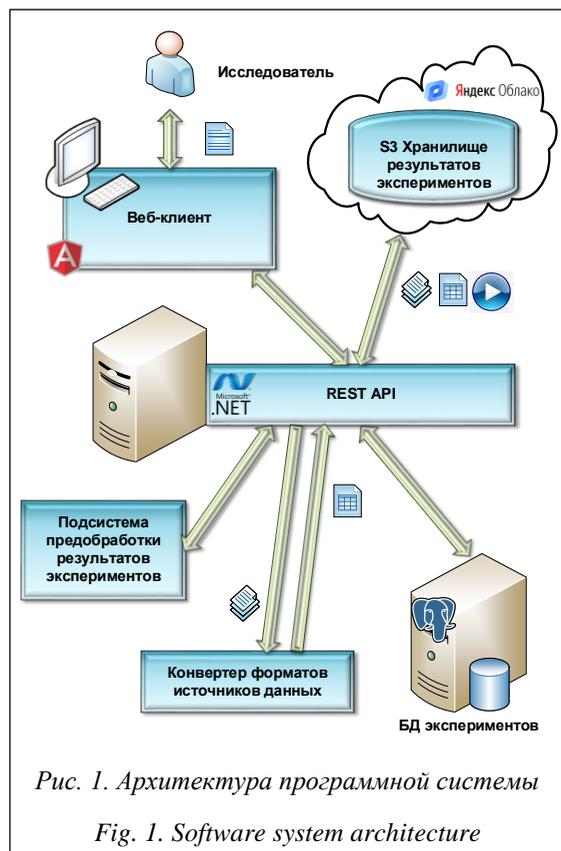


Рис. 1. Архитектура программной системы

Fig. 1. Software system architecture

представляет собой описание одной или нескольких сущностей в JSON-формате;

- POST для создания новых сущностей; описание сущности представляется в JSON-формате в теле запроса;

- PUT для обновления уже имеющихся сущностей; обновленное описание сущности представляется в JSON-формате в теле запроса;

- DELETE для удаления сущности; идентификатор удаляемой сущности передается как часть URL в запросе.

Для каждой сущности используется свой базовый адрес URL, что позволяет структурировать REST API, обеспечивая тем самым более простое и понятное взаимодействие с ним со стороны веб-клиента. Полное описание API доступно по ссылке <https://cps-control.ru/swagger/index.html>.

Веб-клиент представляет собой одностороннее веб-приложение, выполненное на базе платформы Angular в виде набора HTML-страниц, JS-скриптов и таблиц стилей CSS (<http://www.swsys.ru/uploaded/image/2024-3/12.jpg>). Веб-клиент изначально загружается с сервера и выполняется в браузере исследователя. Взаимодействие веб-клиента и REST API осуществляется с помощью указанных выше видов асинхронных HTTP-запросов.

В состав серверной компоненты также входят два дополнительных модуля:

- подсистема предобработки результатов экспериментов, предназначенная для первичной обработки загружаемых файлов с экспериментальными данными, включая сжатие, объединение, разбиение и очистку данных;

- конвертер форматов источников данных, используемый для преобразования информации из исходных форматов, поддерживаемых датчиками объектов КФС, в табличный формат, пригодный для дальнейшей обработки и анализа.

С учетом особенностей хранения информации, получаемой в ходе экспериментов с объектами КФС, для ее обработки и хранения предлагается использовать двухуровневую структуру модели данных. На верхнем уровне данная модель определяет структуру данных базовых сущностей и включает следующую информацию:

- набор взаимосвязанных параметров, определяющих структуру и план миссии;
- начальные условия, задающие состояние КФС на момент начала исполнения миссии;
- состав акторов, участвующих в миссии;
- ограничения, которые необходимо учитывать при построении плана миссии;

– сведения о первичных источниках данных, используемых при реализации сценария управления и проведении эксперимента.

Этот уровень может быть представлен в формализованном виде. Для выполнения подобной формализации необходимо задавать структуру сценария управления с помощью схем данных, характерных для реляционных или объектно-реляционных СУБД, поскольку это позволяет однозначно выделять элементы структуры модели данных и устанавливать взаимосвязь между ними. Подобное требование также обусловлено необходимостью в дальнейшем осуществлять выборки сценариев управления по различным наборам условий с целью дальнейшего анализа и оптимизации данных сценариев.

Таким образом, при описании схемы данных на верхнем уровне должны быть предусмотрены следующие сущности: справочники типов акторов, акторов, грузов, форматов первичных источников данных, сценарии управле-

ния, сведения об использовании грузов в плане миссии, об использовании акторов в плане миссии; о заданиях, об этапах заданий, о первичных источниках данных.

Учитывая эти требования, для хранения данных на верхнем уровне была разработана БД в СУБД PostgreSQL. Физическая модель БД представлена на рисунке 2. В ее структуре предусмотрены таблицы-справочники с информацией о транспортируемых грузах (габариты, тип крепления, масса и т.п.) – Cargoes, об акторах (объектах КФС) – CPSObjects, о типах акторов – CPSObjectTypes, о форматах, используемых в источниках данных, – DataSourceFormats.

Также в структуре БД предусмотрены таблицы для хранения сведений об экспериментах и связанных с ними сценариях управления.

На нижнем уровне модель данных оперирует гетерогенной информацией, получаемой от первичных источников данных. Этими источниками являются датчики акторов и внеш-

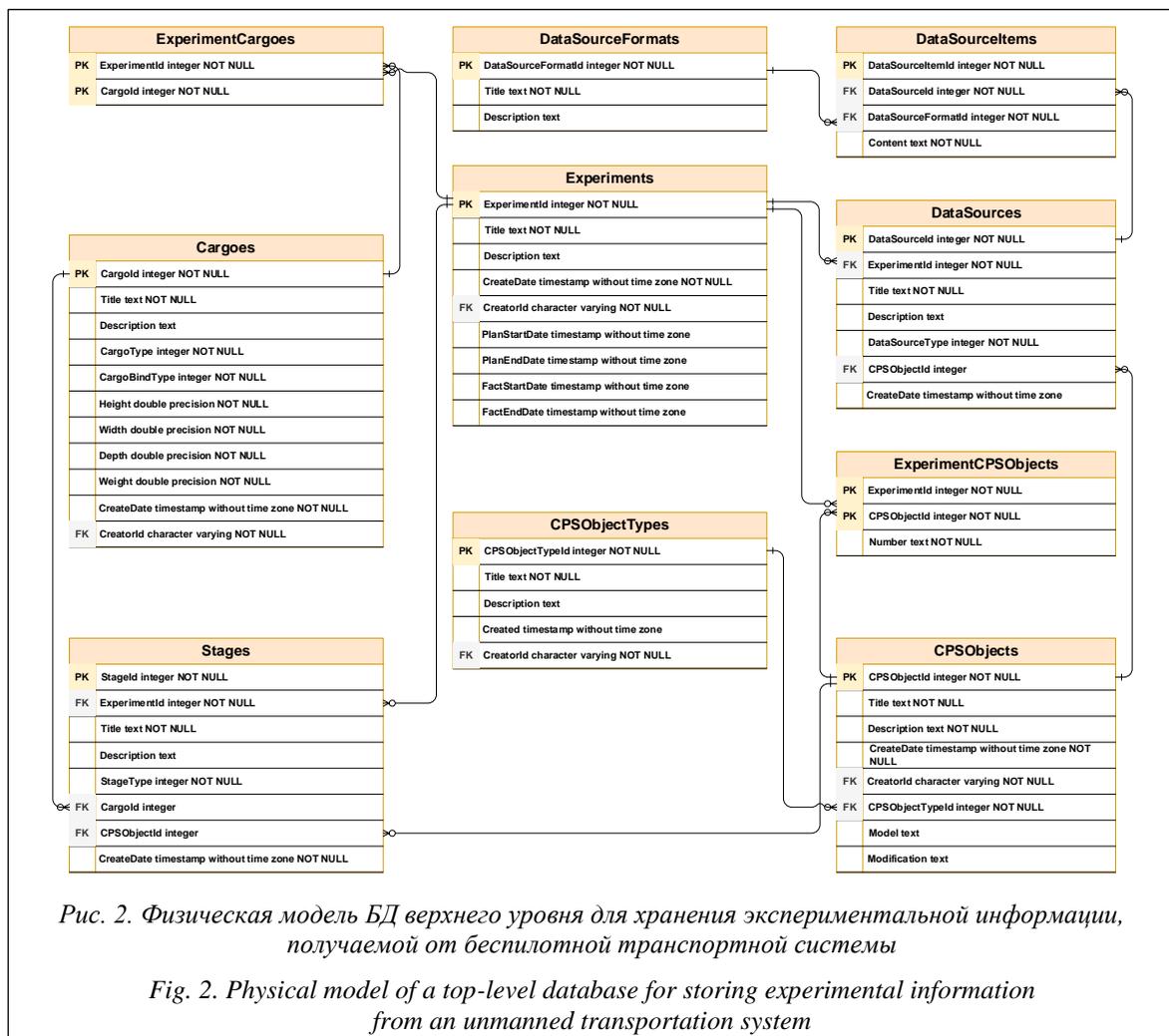


Рис. 2. Физическая модель БД верхнего уровня для хранения экспериментальной информации, получаемой от беспилотной транспортной системы

Fig. 2. Physical model of a top-level database for storing experimental information from an unmanned transportation system

них систем наблюдения. Получаемая от них информация может представляться в форме видеопотока, аудиопотока, потока числовых векторов и т.п. При этом обязательным требованием ко всем потокам данных, используемых в рамках сценария, является наличие привязки каждого фрейма данных ко времени. Объем подобных данных достаточно велик и не всегда структурирован, поэтому для дальнейшей работы с ним должны быть предусмотрены средства хранения его в первоначальном виде, а также методы очистки и предобработки данных.

С учетом того, что сами потоки данных зачастую могут быть представлены в виде файлов соответствующего формата, для их хранения целесообразно использовать облачные хранилища на основе технологии S3. Соответственно, для работы с файлами, содержащими показания датчиков, использовано S3-хранилище, построенное на базе Яндекс Облака. Взаимодействие с S3-хранилищем реализовано с помощью библиотеки AWS SDK для .NET и осуществляется путем преобразования запросов от веб-клиента к REST API в запросы к S3 API Яндекс Облака. Поддерживаются запросы на загрузку файла в бакет, удаление файла из бакета и обновление файла в бакете.

С целью обеспечения взаимосвязи между моделями верхнего и нижнего уровней на верхнем уровне должна быть предусмотрена дополнительная сущность – сведения об объекте S3-хранилища, используемого для хранения первичных данных определенного источника в указанном формате. Для этого в таблице DataSourceItems размещено поле Content строкового типа, в котором хранится в сериализованном виде (на базе формата JSON) описание связанного объекта бакета в S3-хранилище.

Описание и обсуждение экспериментов

В рамках тестирования и апробации разработанной системы был проведен ряд экспери-

ментов с участием следующих беспилотных мобильных средств: планер (с установленным GPS-датчиком), мультикоптер (квадрокоптер с установленными датчиками следующих видов: барометр, RGB-камера, инерциальная система наклона, GPS-датчик) и мобильный ровер (с установленными датчиками следующих видов: GPS-модуль, лазерный дальномер, модуль технического зрения).

Информация об одной из проведенных серий экспериментов, результаты которых были помещены в разработанное хранилище, представлена в таблице. В качестве примера можно привести эксперимент по сбору данных от сенсоров (координат объекта в пространстве и наблюдаемых с него видеопотоков) при перемещении мобильного ровера с установленной стереопарой. При продолжительности эксперимента 5 мин. были получены log-файл состояния размером 1,5 МБ и файл ROS-bag со стереопары в формате RGB и RGBD размером файла 5,57 ГБ.

Полученные файлы с экспериментальной информацией были помещены в хранилище системы: метаданные сохранялись в таблицах БД PostgreSQL, а сами файлы помещались в S3-хранилище Яндекс Облака. Связь между хранилищами обеспечивалась за счет структурированного в формате JSON объекта в таблице базы PostgreSQL.

Отметим, что для некоторых форматов (например, log-файлов) имеется возможность конвертировать их в CSV-формат, более удобный для последующей автоматизированной обработки и анализа данных.

Как можно заметить, даже небольшой по продолжительности эксперимент (длительностью порядка 5 мин.) может порождать значительный объем информации (более 5 ГБ), что подтверждает адекватность решения об использовании специализированного S3-хранилища исходной экспериментальной информации. Вместе с тем хранение файлов без метаданных о структуре и параметрах проведенных

Информация о результатах одной серии экспериментов с объектами КФС

Information on the results of one series of experiments with CPS objects

Тип объекта КФС	Количество экспериментов	Средняя продолжительность эксперимента, мин.	Количество источников информации	Суммарный объем полученных экспериментальных данных
Планер	3	19,3	1	154,7 Мб
Мультикоптер	2	10	1	17,78 Мб
Ровер	1	5	2	5,57 Гб

экспериментов и используемых форматах файлов существенно затруднило бы дальнейшее использование накопленной информации в автоматизированных системах анализа данных. Однако в рамках разработанной программной системы указанная проблема решается за счет структурированного хранения сведений об экспериментах и объектах КФС путем установления однозначного соответствия между записями в БД PostgreSQL и объектами (файлами) S3-хранилища. Таким образом, можно утверждать, что предложенная архитектура системы пригодна для решения задач хранения экспериментальной информации при исследовании моделей управления объектами КФС.

В целом следует отметить, что для хранения экспериментальной информации в разных предметных областях большинство исследователей предпочитают использовать либо локальные файловые хранилища, либо облачные (<https://www.stemcell.com/efficient-research/storing-data>). В этом случае структурирование экспериментов осуществляется за счет иерархического упорядочения файлов и каталогов. Данный подход имеет ряд существенных недостатков: низкая надежность хранения информации и сложность организации удаленного доступа при использовании собственных файловых хранилищ; ограниченный объем хранилищ (как локальных, так и облачных); сложность поиска информации по метаданным при удаленном доступе к хранилищу. С другой стороны, наблюдается рост интереса к использованию S3-хранилищ для хранения экспериментальных данных (<https://aws.amazon.com/ru/blogs/robotics/record-store-robot-data-rosbag/>). Однако хранение только самих данных эксперимента без метаданных затрудняет широкое применение подобных подходов, поскольку это существенно ограничивает возможности поиска и выборки экспериментов по заданным критериям. В то же время существуют и специализированные онлайн-хранилища, учитывающие метаданные. Например, подобные системы используются при проведении биологических [13] или обобщенных вычислительных экспериментов [5]. Однако хранилища подобного типа ориентированы на иную структуру метаданных и не подходят для применения в задачах хранения данных функционирования объектов КФС.

Предложенные в статье подходы и описанная система лишены указанных выше ограничений за счет

– использования структурированного хранения метаданных о проведенных эксперимен-

тах, что обеспечивает возможности поиска и выборки экспериментов по заданным критериям для их дальнейшего анализа;

– применения надежных и масштабируемых S3-хранилищ, гарантирующих сохранность результатов экспериментов, а также обеспечивающих возможность удаленного доступа и отсутствие ограничений по объему хранимой информации.

Заключение

Создание технологий и средств хранения информации, получаемой от объектов КФС в процессе их функционирования, является важным и перспективным направлением в рамках создания подходов и методов управления объектами КФС на основе анализа и интерпретации разнородных данных разного уровня структуризации. Предложенная система позволяет решать задачи хранения данных, получаемых в ходе экспериментов с объектами КФС, как в исходном, сыром, виде, так и в форматах, пригодных для автоматизированной обработки. Одним из перспективных путей использования этих данных является создание наборов данных для тестирования различных сценариев управления. Дополнительно за счет хранения в структурированном виде данных об экспериментах система позволяет накапливать соответствующий опыт, который в дальнейшем может быть использован при постановке и планировании новых экспериментов.

Принципы построения системы хранения данных были описаны на примере задач перемещения беспилотных мобильных средств в пространстве при решении задач транспортировки грузов. Вместе с тем с точки зрения архитектуры, применяемых подходов к хранению данных и форматов для их представления система в достаточной мере не зависит как от конфигурации КФС, так и от области ее применения. Благодаря этому имеется возможность масштабирования системы за счет как распределенного накопления и хранения разнородных данных, получаемых от объектов КФС и иных источников информации, так и расширения состава поддерживаемых типов объектов, форматов источников данных и методов их предварительной обработки, что дает возможность создать универсальную систему, позволяющую решать более широкий класс задач в области хранения и предварительной обработки разнородных данных. В свою очередь, данное обстоятельство будет способствовать

систематизации различных классов задач и сценариев управления объектами КФС, что может быть полезным как при разработке новых технологий и инструментов управления, так и при адаптации существующего математического аппарата и алгоритмического обеспечения при создании таких инструментов.

Еще одним актуальным направлением развития предложенного подхода является его расширение до уровня хранения не только самих данных функционирования объектов КФС,

но и моделей сценариев управления объектами. Таким образом, от системы хранения данных можно перейти к созданию на ее основе библиотек и репозиториев моделей по разным классам ситуационных задач управления объектами КФС. Наличие подобных репозиториев является необходимым условием развития ситуационного подхода к управлению объектами КФС, поскольку именно таким образом возможно накопление необходимой базы прецедентов.

Список литературы

1. Sánchez M., Morales J., Martínez J.L., Fernández-Lozano J.J., García-Cerezo A. Automatically annotated dataset of a ground mobile robot in natural environments via gazebo simulations. *Sensors*, 2022, vol. 22, no. 15, art. 5599. doi: 10.3390/s22155599.
2. Bosello M., Aguiari D., Keuter Y., Pallotta E. et al. Race against the machine: A fully-annotated, open-design dataset of autonomous and piloted high-speed flight. *IEEE Robotics and Automation Letters*, 2024, vol. 9, no. 4, pp. 3799–3806. doi: 10.1109/LRA.2024.3371288.
3. Bennett M.K., Younes N., Joyce K. Automating drone image processing to map coral reef substrates using google earth engine. *Drones*, 2020, vol. 4, no. 3, art. 50. doi: 10.3390/drones4030050.
4. Zhuo X., Koch T., Kurz F., Fraundorfer F., Reinartz P. Automatic UAV image geo-registration by matching UAV images to georeferenced image data. *Remote Sensing*, 2017, vol. 9, no. 4, art. 376. doi: 10.3390/rs9040376.
5. Подвесовский А.Г., Коростелёв Д.А., Лупачёв Е.А., Беляков Н.В. Построение хранилища обобщенных вычислительных экспериментов на основе онтологического подхода // Онтология проектирования. 2022. Т. 12. № 1. С. 41–56. doi: 10.18287/2223-9537-2022-12-1-41-56.
6. Wiemann T., Igelbrink F., Pütz S., Hertzberg J. A file structure and reference data set for high resolution hyperspectral 3D point clouds. *IFAC-PapersOnLine*, 2019, vol. 52, no. 8, pp. 403–408. doi: 10.1016/j.ifacol.2019.08.101.
7. Morast E. A deep learning based approach to object recognition from LiDAR data along Swedish railroads. TRITA, TRITA-ABE-MBT-22448. URL: <https://www.diva-portal.org/smash/get/diva2:1678019/FULLTEXT01.pdf> (дата обращения: 15.04.2024).
8. Wang S., Li G., Yao X., Zeng Y., Pang L., Zhang L. A distributed storage and access approach for massive remote sensing data in MongoDB. *ISPRS Int. J. Geo-Inf.*, 2019, vol. 8, no. 12, art. 533. doi: 10.3390/ijgi8120533.
9. Béjar-Martos J.A., Rueda-Ruiz A.J., Ogayar-Anguita C.J., Segura-Sánchez R.J., López-Ruiz A. Strategies for the storage of large LiDAR datasets – a performance comparison. *Remote Sensing*, 2022, vol. 14, no. 11, art. 2623. doi: 10.3390/rs14112623.
10. Allahham MHD S., Al-Sa'd M.F., Al-Ali A. et al. DroneRF dataset: A dataset of drones for RF-based detection, classification and identification. *Data in Brief*, 2019, vol. 26, art. 104313. doi: 10.1016/j.dib.2019.104313.
11. Захарова А.А., Кутахов В.П., Мещеряков Р.В., Подвесовский А.Г., Смолин А.Л. Моделирование задач транспортировки грузов в беспилотной авиационной транспортной системе // Авиакосмическое приборостроение. 2023. № 3. С. 3–15.
12. Zakharova A., Podvesovskii A. Model for optimization of heterogeneous cargo transportation using UAVs, taking into account the priority of delivery tasks. In: *SIST. Proc. ADOP*, 2023, vol. 362, pp. 257–268. doi: 10.1007/978-981-99-4165-0_24.
13. Morrell W.C., Birkel G.W., Forrer M. et al. The experiment data depot: A web-based software tool for biological experimental data storage, sharing, and visualization. *ACS Synth. Biol.*, 2017, vol. 6, no. 12, pp. 2248–2259. doi: 10.1021/acssynbio.7b00204.

Cyber-physical system objects: Organizing operation data storage

Dmitry A. Korostelyov^{1,2}, Aleksandr G. Podvesovskii^{1,2} ✉, Alena A. Zakharova¹

¹ V.A. Trapeznikov Institute of Control Sciences of Russian Academy of Sciences, Moscow, 117997, Russian Federation

² Bryansk State Technical University, Bryansk, 241035, Russian Federation

For citation

Korostelyov, D.A., Podvesovskii, A.G., Zakharova, A.A. (2024) 'Cyber-physical system objects: Organizing operation data storage', *Software & Systems*, 37(3), pp. 334–343 (in Russ.). doi: 10.15827/0236-235X.142.334-343

Article info

Received: 16.04.2024

After revision: 05.05.2024

Accepted: 14.05.2024

Abstract. The paper presents the results of research in the field of organization and software support when storing operation data from cyber-physical system objects obtained during experiments. The identified features of their storage are the following: special data formats, large amounts of data with different structuring degrees, supporting universal structured data formats, storing information about the structure and scenarios of experiments. The authors of the paper have analyzed technologies for building storage systems and special-purpose data formats that consider these features. They also proposed an approach to building a software system for storing and preprocessing experimental information about the functioning of cyber-physical system objects. The architecture and peculiarities of the system implementation are considered on the example of storing data of experiments with groups of jointly operating unmanned mobile vehicles when solving tasks of moving in space and transporting cargoes. The architecture assumes a server application and a web-client interacting with it by means of an API-interface based on REST principles. The server application also includes a subsystem for preprocessing experimental results and a data source format converter. The paper proposes to use a two-level structure of the data model for storing the experimental information. The upper level is intended for storing information about the structure and parameters of experiments and is implemented as a database in the PostgreSQL system. The lower level stores streaming information received from primary data sources. This level is implemented as a cloud storage built based on Yandex Cloud using S3 technology. The results of testing the developed system in real experiments confirmed the validity of the selected architectural solutions. Practical significance of the performed researches consists in developing approaches to the construction of scalable software systems for storing operating data of cyber-physical system objects. This contributes to organizing tasks and control scenarios of these objects when developing new control technologies and tools.

Keywords: cyberphysical system, cyberphysical system objects, experimental data, data formats, data model, DMS, REST API, S3 storage

Acknowledgements. This study was funded by the Russian Science Foundation, project no. 23-19-00342, <https://rscf.ru/en/project/23-19-00342/>

References

1. Sánchez, M., Morales, J., Martínez, J.L., Fernández-Lozano, J.J., García-Cerezo, A. (2022) 'Automatically annotated dataset of a ground mobile robot in natural environments via gazebo simulations', *Sensors*, 22(15), art. 5599. doi: 10.3390/s22155599.
2. Bosello, M., Aguiari, D., Keuter, Y., Pallotta, E. et al. (2024) 'Race against the machine: A fully-annotated, open-design dataset of autonomous and piloted high-speed flight', *IEEE Robotics and Automation Letters*, 9(4), pp. 3799–3806. doi: 10.1109/LRA.2024.3371288.
3. Bennett, M.K., Younes, N., Joyce, K. (2020) 'Automating drone image processing to map coral reef substrates using google earth engine', *Drones*, 4(3), art. 50. doi: 10.3390/drones4030050.
4. Zhuo, X., Koch, T., Kurz, F., Fraundorfer, F., Reinartz, P. (2017) 'Automatic UAV image geo-registration by matching UAV images to georeferenced image data', *Remote Sensing*, 9(4), art. 376. doi: 10.3390/rs9040376.
5. Podvesovskii, A.G., Korostelyov, D.A., Lupachev, E.A., Belyakov, N.V. (2022) 'Building a repository of generalized computational experiments based on the ontological approach', *Ontology of Designing*, 12(1), pp. 41–56 (in Russ.). doi: 10.18287/2223-9537-2022-12-1-41-56.
6. Wiemann, T., Igelbrink, F., Pütz, S., Hertzberg, J. (2019) 'A file structure and reference data set for high resolution hyperspectral 3D point clouds', *IFAC-PapersOnLine*, 52(8), pp. 403–408. doi: 10.1016/j.ifacol.2019.08.101.
7. Morast, E. 'A deep learning based approach to object recognition from LiDAR data along Swedish railroads', *TRITA, TRITA-ABE-MBT-22448*, available at: <https://www.diva-portal.org/smash/get/diva2:1678019/FULLTEXT01.pdf> (accessed April 15, 2024).
8. Wang, S., Li, G., Yao, X., Zeng, Y., Pang, L., Zhang, L. (2019) 'A distributed storage and access approach for massive remote sensing data in MongoDB', *ISPRS Int. J. Geo-Inf.*, 8(12), art. 533. doi: 10.3390/ijgi8120533.
9. Béjar-Martos, J.A., Rueda-Ruiz, A.J., Ogayar-Angueta, C.J., Segura-Sánchez, R.J., López-Ruiz, A. (2022) 'Strategies for the storage of large LiDAR datasets – a performance comparison', *Remote Sensing*, 14(11), art. 2623. doi: 10.3390/rs14112623.
10. Allahham, MHD S., Al-Sa'd, M.F., Al-Ali, A. et al. (2019) 'DroneRF dataset: A dataset of drones for RF-based detection, classification and identification', *Data in Brief*, 26, art. 104313. doi: 10.1016/j.dib.2019.104313.
11. Zakharova, A.A., Kutakhov, V.P., Meshcheryakov, R.V., Podvesovskii, A.G., Smolin A.L. (2023) 'Modeling cargo transportation tasks in an unmanned air transportation system', *Aerospace Instrument-Making*, (3), pp. 3–15 (in Russ.).
12. Zakharova, A., Podvesovskii, A. (2023) 'Model for optimization of heterogeneous cargo transportation using UAVs, taking into account the priority of delivery tasks', in *SIST. Proc. ADOP*, 362, pp. 257–268. doi: 10.1007/978-981-99-4165-0_24.
13. Morrell, W.C., Birkel, G.W., Forrer, M. et al. (2017) 'The experiment data depot: A web-based software tool for biological experimental data storage, sharing, and visualization', *ACS Synth. Biol.*, 6(12), pp. 2248–2259. doi: 10.1021/acssynbio.7b00204.

Авторы

Коростелев Дмитрий Александрович ^{1, 2},

к.т.н., доцент,
nigm85@mail.ru

Подвесовский Александр Георгиевич ^{1, 2},

к.т.н., доцент, профессор,
apodv@tu-bryansk.ru

Захарова Алена Александровна ¹,

д.т.н., доцент, главный научный сотрудник,
zaawmail@gmail.com

¹ Институт проблем управления
им. В.А. Трапезникова РАН,
г. Москва, 117997, Россия

² Брянский государственный технический
университет, г. Брянск, 241035, Россия

Authors

Dmitry A. Korostelyov ^{1, 2},

Cand. of Sci. (Engineering),
Associate Professor, nigm85@mail.ru

Aleksandr G. Podvesovskii ^{1, 2},

Cand. of Sci. (Engineering), Associate Professor,
Professor, apodv@tu-bryansk.ru

Alena A. Zakharova ¹,

Dr.Sci. (Engineering), Associate Professor,
Chief Researcher, zaawmail@gmail.com

¹ V.A. Trapeznikov Institute of Control Sciences
of Russian Academy of Sciences,
Moscow, 117997, Russian Federation

² Bryansk State Technical University,
Bryansk, 241035, Russian Federation