

# Методология структурного синтеза хранилища гетерогенных данных технической диагностики высоковольтного оборудования объектов электроэнергетики

А. И. Хальясмаа, С. В. Поршнева

**Аннотация**—В работе рассматривается методология структурного синтеза хранилища гетерогенных данных (ХГД), накапливаемых в процессе технической диагностики (ТД) высоковольтного электрооборудования (ВЭО) объектов электроэнергетики, которая базируется на онтологии данной предметной области. Необходимость создания методологии обусловлена тем, что модели данных, традиционно используемых для хранения диагностической информации (ДИ), собираемой в процессе ТД ВЭО (темпоральные, реляционные и файловые) не обеспечивают их семантической согласованности. В результате имеют место очевидные проблемы в процессе причинно-следственной интерпретации результатов технической диагностики. Для обеспечения семантической согласованности данных ТД, гетерогенных по своей природе, проведена формализация области технической диагностики высоковольтного электрооборудования в виде системы сущностей и отношений, а также набора операторов семантического картирования и онтологической нормализации. Разработана структура представления диагностических данных, обеспечивающая переход от измеряемых параметров к диагностическим фактам и эксплуатационным рекомендациям. Предложенная методология обеспечивает воспроизводимость, масштабируемость и эволюционную устойчивость структуры хранения диагностических данных (ДД) и может быть использована при построении цифровых моделей технического состояния оборудования и интеграции ДД в корпоративные системы управления активами.

**Ключевые слова**—техническая диагностика, высоковольтное оборудование, хранилище гетерогенных данных, онтология предметной области, структурный синтез.

## 1. ВВЕДЕНИЕ

В процессе эксплуатации объектов электроэнергетики формируются значительные объемы разнородной ДИ, для хранения которой, традиционно, используются темпоральные, реляционные и файловые модели данных (см. рис. 1). Например, для инфракрасного (ИК) контроля трансформатора тока необходима

термограмма (снимок в ИК спектре из файловой части хранилища данных), данные о токовой нагрузке из АСУ ТП или SCADA-системе (темпоральная модель), метеорологические данные, которые могут быть представлены как в метаданных снимка, так и отдельно в темпоральной модели [1]. В ряде случаев может потребоваться история эксплуатации трансформатора (даты ввода в эксплуатацию, ремонтов и т.д.) из реляционной базы данных (БД), а также данные инспекционных обходов из автоматизированных информационных систем (АИС).

Следует отметить, что известные в электроэнергетике информационные модели (ИМ), включая общую ИМ (*Common Information Model – CIM*) [2], ориентированы преимущественно на описание конфигурации и режимных параметров энергосистемы. Они не обеспечивают формализованного представления агрегированных оценок технического состояния ВЭО. ГОСТ Р 71853–2024 [3] задает рамочные требования к архитектуре систем удаленного мониторинга и ТД ВЭО, но не формализует ИМ.

В этой связи разработка методологии структурного синтеза ХГД ТД ВЭО, обеспечивающей согласование ДИ, формируемой различными источниками, представляемой с помощью перечисленных выше моделей данных, является актуальной задачей.



Рис. 1: Базовые модели хранения данных ТД ВЭО

## II. МЕТОДОЛОГИЯ СТРУКТУРНОГО СИНТЕЗА ХГД

### А. Базовые принципы методологии

В рамках данной работы задача структурного синтеза ХГД рассматривается как задача выявления

Статья получена 14 марта 2026.

А. И. Хальясмаа, канд. тех. наук, доц., зав. лабораторией, a.i.khaliassmaa@urfu.ru

С. В. Поршнева, д-р. тех. наук, проф., проф., s.v.porshnev@urfu.ru

Федеральное государственное автономное образовательное учреждение высшего образования «Уральский федеральный университет имени первого Президента России Б.Н. Ельцина», Екатеринбург.

инвариантных компонентов предметной области ТД ВЭО, к которым были отнесены следующие сущности: «единица оборудования» (объект диагностики); «диагностическое событие» (факт проведения обследования); «измерение» (количественное значение диагностического параметра, зафиксированное в определенных условиях); «наблюдение» (интерпретируемое проявление состояния объекта); «дефект» (результат обобщения диагностических признаков); «контекст эксплуатации», включающий режимные параметры и внешние воздействия; «происхождение данных» (свойство, обеспечивающее установление соответствия между диагностическим фактом и первичным артефактом (протоколом, отчетом, изображением)). Выявленные сущности присутствуют во всех диагностических процедурах, однако, в существующей практике ТД ВЭО, зачастую представлены фрагментарно и не имеют единой формализованной структуры связей между собой.

Мы называем единым ХГД, для описания структуры которого используются единые онтологическое и концептуально-логическое представления данных ТД ВЭО, а также реализован соответствующий централизованный механизм доступа к ним. При этом физически данные ТД ВЭО, описываемые различными типами моделей, хранятся независимо друг от друга, а их единство обеспечивается за счет согласования и унификации доступа к данным за счет семантического картирования и использования нормализованного слоя диагностических фактов. Отметим, что выбранный подход не требует перестроения или замены концептуальных и логических моделей данных существующих систем оперативного и режимного учета, таких как SCADA, АСУ ТП, АСКУЭ и других корпоративных информационных систем (ИС).

Требования к ХГД ТД ВЭО были сформулированы на основе следующих основополагающих принципов.

**1. Принцип системной иерархичности**, в соответствии с которым ХГД рассматривается как многоуровневая система, структура которой представлена на рис. 3.



Рис. 2: Системная иерархия уровней представления данных в ХГД ТД ВЭО

При этом каждый уровень выполняет самостоятельную функцию, взаимодействуя только с соседними уровнями. Это исключает жесткую связанность структуры данных с конкретной

физической реализацией и обеспечивает устойчивость системы при изменениях на нижележащих уровнях.

**2. Принцип семантической совместимости**, в соответствии с которым структура ХГД обеспечивает совместимость гетерогенных данных ТД ВЭО не только на уровне форматов и структур хранения, но и на уровнях смыслов, за счет использования онтологической модели (ОМ) предметной области и семантического согласования (картирования), и их интерпретации в едином понятийном пространстве.

**3. Принцип динамичности онтологии**, в соответствии с которым ОМ предметной области допускает при необходимости, ее модификацию, что обеспечивает эволюционную устойчивость структуры данных ТД ВЭО и сохранение интерпретируемости накопленной ДИ с течением времени (в условиях постоянного совершенствования нормативной базы ТД ВЭО, состава оборудования и методов ТД).

**4. Принцип адаптивности и масштабируемости**, в соответствии с которым методология структурного синтеза ХГД ТД ВЭО при появлении новых источников, параметров и методов диагностики должна обеспечивать адаптацию структуры ХГД, а также масштабирование системы при росте объемов и разнообразия ДИ за счет расширения онтологической и концептуально-логической моделей.

**5. Принцип причинно-следственной интерпретируемости диагностических данных**, в соответствии с которым методология структурного синтеза ХГД должна обеспечивать представление данных ТД ВЭО с учетом причинно-следственных связей между режимными воздействиями, процессами деградации ВЭО и наблюдаемыми диагностическими проявлениями.

**6. Принцип совместимости ХГД с корпоративными системами**, в соответствии с которым методология структурного синтеза ХГД ТД ВЭО должна обеспечивать согласование структуры данных и ОМ с ИМ и классификаторами корпоративных систем управления активами с целью интеграции результатов диагностики в ERP- и EAM-системы.

**7. Принцип интеграции со стандартами**, в соответствии с которым методология структурного синтеза ХГД ТД ВЭО должна обеспечивать согласованность структуры ДД с международными и национальными стандартами электроэнергетики и управления активами без жесткой привязки к конкретным версиям стандартов.

### В. Формализация задачи синтеза ХГД

Обозначим:  $D$  – множество гетерогенных данных ТД ВЭО, поступающих из различных ИС;  $O = (C, R)$  – ОМ предметной области ТД, где  $C$  – множество классов,  $R$  – множество отношений между классами.

Тогда задача структурного синтеза ХГД ТД ВЭО данных состоит в построении отображения

$$M: D \rightarrow O, \quad (1)$$

которое обеспечивает сопоставление ДД с классами ОМ и разработку концептуальной, логической и физической моделей ХГД.

Результат отображения  $M$  – множество

нормализованных диагностических фактов  $F$ , где каждый диагностический факт представляет интерпретированное состояние объекта ТД, сформированное на основе множества  $D$  и семантического контекста элементов данного множества. Критерием адекватности решения обсуждаемой задачи наличие соответствия между каждым элементом множества  $F$  и соответствующим элементом множества  $D$ :

$$\forall f_i \in F \Rightarrow \exists d_i \in D. \quad (2)$$

Таким образом, задача структурного синтеза ХГД ТД ВЭО состоит в построении такого отображения  $M$  (1), которое соответствует (2).

### С. Информационная модель методологии структурного синтеза ХГД

Информационная модель методологии структурного синтеза ХГД ТД ВЭО, обеспечивающая решение обсуждаемой задачи, представлена на рис. 3. Данная модель состоит из следующих компонентов.

**Блок «Входные основания»** включающий: сущность «Источники диагностических данных», содержащую перечни источников ДД ТД ВЭО, а также соответствующие им модели данных (темпоральные, реляционные и файловые) данных; сущность «Нормативная база и терминология», содержащую нормативную базу ТД ВЭО и соответствующую терминологию, действующие ограничения при интерпретации диагностических параметров, состав диагностических признаков, методы контроля и формы представления результатов.

**Блок «Модели»**, включающий: сущность «М1. Онтологические модели предметной области», обеспечивающую использование инвариантных сущностей, состояний и отношений, используемых в ТД

ВЭО; сущность «М2. Концептуальная модель диагностических данных», обеспечивающую проекцию ОМ на задачи хранения и анализа ДИ и определение состава сущностей, атрибутов и связей, релевантных предметной области; сущность «М3. Логическая модель диагностических фактов», обеспечивающую формализацию структуры ХГД и представления ДД без привязки к конкретной физической реализации и служит основанием для последующего извлечения и анализа ДИ.

**Блок «Методы»**, включающий: сущность «П1. Понятийная проекция», обеспечивающую отображение нормативно закрепленных понятий и терминов предметной области технической диагностики в ОМ, формируя единое понятийное пространство; сущность «П2. Структурная формализация», обеспечивающую преобразование концептуальных представлений диагностических данных в логическую модель диагностических фактов, пригодную для формального описания и дальнейшей обработки; сущность «П3. Семантическое картирование и нормализация», обеспечивающую согласование разнородных данных, поступающих из различных источников, с онтологической и логической моделями, обеспечивая формирование нормализованных диагностических фактов с сохранением их трассируемости к первичным данным; сущность «П4. Управление изменениями (версионность)», обеспечивающую учет динамики текущего состояния предметной области и обеспечивает эволюционную устойчивость методологии за счет поддержки версионности онтологических представлений и актуализации правил семантического картирования.

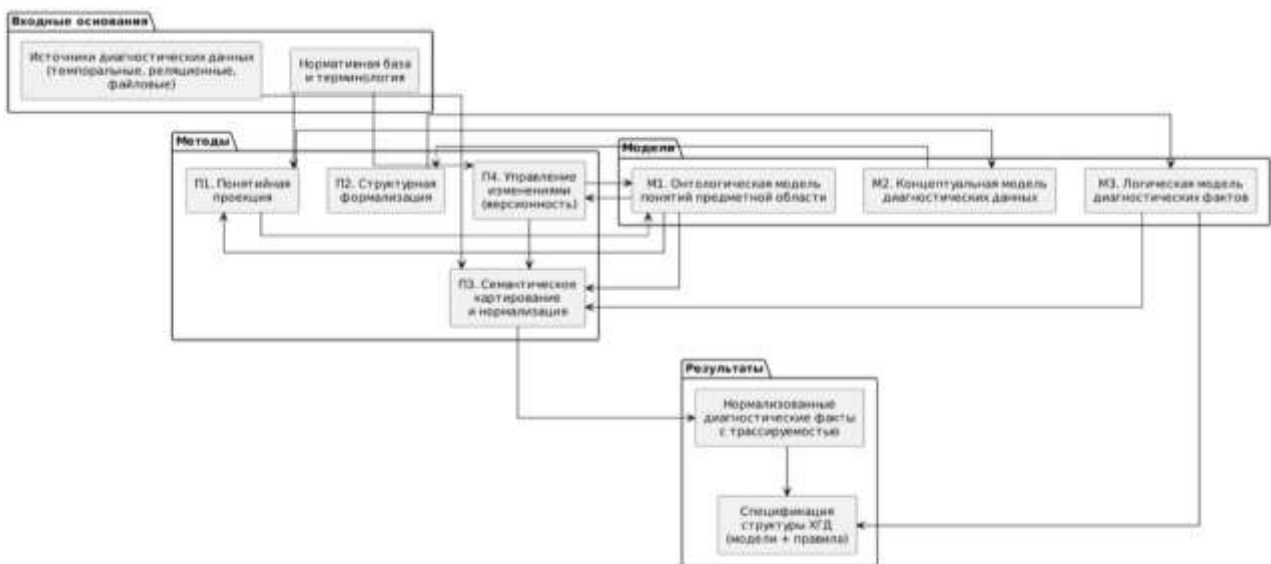


Рис. 3: Информационная модель методологии структурного синтеза ХГД ТД ВЭО

**Блок «Результаты»**, содержащий следующие сущности: «Нормализованные диагностические факты с трассируемостью», обеспечивающую применение метода семантического картирования за счет формирования нормализованных диагностических фактов с обеспечением трассируемости ДИ к исходным

данным, что позволяет воспроизводимо извлекать диагностическую информацию независимо от формы ее первоначального хранения; «Спецификация структуры ХГД (модели + правила)», обеспечивающую согласованность сформированных моделей и правил со спецификацией структуры ХГД,

включающей описание онтологической, концептуальной и логической моделей, а также правила доступа и интерпретации диагностических данных.

#### D. Этапы реализации метода структурного синтеза ХГД.

Практическая реализация методологии синтеза ХГД, как следует из информационной модели методологии структурного синтеза ХГД ТД ВЭО, представленной на рис. 3, реализуется выполнением следующей последовательности действий.

1. **Создание онтологической модели предметной области** на основе результатов анализа нормативной базы ТД, а также существующих информационных моделей ВЭО.

2. **Структурная формализация диагностических данных** на основе анализа онтологической модели с целью определения состава сущностей, атрибутов и связей, используемых далее для представления ДИ, выбора используемой далее концептуальной модели данных.

3. **Семантическое картирование источников данных** с целью установления соответствия между ДД, представленными различными ИС и сущностями онтологической модели (отображение ДД в единое понятийное пространство предметной области).

4. **Нормализация** (на основе результатов семантического картирования) **диагностических фактов** с целью формирования нормализованных диагностических фактов, обеспечивающих согласование ДИ и исходных данных.

5. **Формирование логической структуры ХГД**, обеспечивающего интеграцию диагностических параметров, признаков, фактов и выводов в виде согласованной системы представления данных.

### III. Оценка применимости существующих моделей для структурного синтеза ХГД

#### A. Критерии оценивания

Для оценки применимости известных подходов в задаче структурного синтеза ХГД ТД ВЭО были использованы следующие критерии.

**Системная иерархичность (К1)** – наличие возможности разделения онтологического, концептуального, логического и физического уровней представления данных с локализацией изменений (принцип системной иерархичности).

**Семантическая совместимость (К2)** – наличие возможности согласования данных, предоставленных

различными источниками ДИ, не только на уровне форматов и схем хранения, но и на уровне смыслов (принцип семантической совместимости).

**Масштабируемость (К3)** – наличие возможности расширения структуры ХГД ТД ВЭО в случае добавлении новых источников ДИ, диагностируемых параметров и объектов, не требующей внесения каких-либо изменений в структурные модели ранее полученной ДИ (принцип адаптивности и масштабируемости).

**Динамичность и эволюционная устойчивость (К4)** – наличие возможности адаптации структуры ХГД ТД ВЭО к изменениям нормативной базы и методов ТД при сохранении интерпретируемости ретроспективной информации (принцип динамической онтологии).

**Интеграция со стандартами (К5)** – наличие возможности согласования ДИ с действующими национальными и международными стандартами без жесткой привязки к конкретным версиям данных документов (принцип интеграции со стандартами).

**Смысловая нормализация гетерогенности (К6)** – наличие возможности приведения темпоральных, реляционных и иерархических данных к единому понятийному представлению с учетом контекстов моделей и терминов ТД (принципы совместимости ХГД ТД ВЭО с корпоративными ИС и причинно-следственной интерпретируемости диагностических данных).

#### B. Анализ подходов интеграции разнородных данных

Традиционно, для решения проблемы интеграции разнородных данных в единую структуру хранения используются: **реляционная модель Э. Кодда** [4] (модель М1); **модель «сущность–связь» П. Чена** [5] (модель М2); **модели Инмона** [6] и **Кимбалла** [7] (модели М3, М4, соответственно); **объектно-ориентированные модели данных** [8] (модель М5); **онтологические модели** [9, 10] (модель М6).

Оценки применимости известных моделей данных в задаче структурного синтеза ХГД ТД ВЭО на основе сформулированных критериев представлены в Табл. 1, из которой видно, что классические модели данных ориентированы на структурную согласованность и аналитическую обработку, однако не обеспечивают формализованного согласования смыслового содержания ДИ из разнородных источников.

Табл. 1: Оценки применимости существующих моделей данных в задаче структурного синтеза ХГД ТД ВЭО

Модель	Требования					
	К1	К2	К3	К4	К5	К6
М1	Нет	Нет	Низкая	Низкая	Нет	Нет
М2	Нет	Средняя*	Низкая	Низкая	Средняя	Средняя
М3	Средняя	Средняя*	Высокая	Средняя	Средняя	Средняя
М4	Средняя	Средняя*	Высокая	Средняя	Средняя	Средняя
М5	Высокая	Высокая	Средняя	Средняя	Средняя	Высокая
М6	Высокая	Высокая	Высокая	Высокая**	Высокая	Высокая

\* на уровне форматов.

\*\* в случае использования динамической онтологии.

В то время как онтологически ориентированные модели, напротив, обеспечивают многоуровневое описание предметной области, формализацию понятий и связей, а также возможность эволюционного расширения структуры без нарушения интерпретируемости данных [11, 12], поэтому онтологический подход был использован в задаче синтеза ХГД ТД ВЭО.

#### IV. СТРУКТУРА ХГД НА ПРИМЕРЕ ДИАГНОСТИКИ ОТКРЫТОГО РАСПРЕДЕЛИТЕЛЬНОГО УСТРОЙСТВА

##### A. Описание объекта ТД

Рассмотрим пример ХГД ТД распределительного устройства (ОРУ) электрической станции/подстанции напряжением 110–220 кВ, особенность которой состоит в высокой концентрации электрооборудования, разнообразии режимов эксплуатации и применения различных методов ТД. ХГД разработано в соответствии с предложенной методологией структурного синтеза ХГД ТД ВЭО. Структура синтезированного ХГД обеспечивает хранение ДД, перечень которых установлен ГОСТ 20911-89 [13], а также эксплуатационных данных (паспортные сведения, режимные параметры, схемные состояния и метеорологические условия, определяющие эксплуатационный контекст ТД).

##### B. Структура ХГД ТД ВЭО ОРУ 110–220 кВ

Рассмотрим концептуальную модель синтезированного ХГД ТД ВЭО ОРУ 110–220 кВ, представленную на рис. 4, из которого видно, что в ней используются следующие сущности.

**OSG (open switch gear)**, обеспечивает хранение информации об ОРУ как целостном объекте диагностики.

**OSGCell (open switch gear cell)**, обеспечивает хранение информации о внутренней структуре ОРУ, состоящей из ячеек, в которых размещается ВЭО.

**EquipmentUnit**, (единица оборудования) обеспечивает хранение информации о конкретной единице диагностируемого ВЭО, обеспечивая поддержку иерархической структуры ВЭО, используемой в цифровых ИМ на основе CIM в соответствии с серией стандартов ГОСТ Р 58651 [14].

**DiagnosticObject** (объект диагностики), обеспечивает хранение результатов ТД ВЭО при рассмотрении объекта диагностики, а также его составных частей и функциональных элементов, с различных точек зрения с сохранением при этом целостности модели данных ТД.

**DataArtifact** (артефакт данных), обеспечивает хранение информации об исходных данных для диагностики ВЭО, включая метку времени, источник использованной информации, системы и тип проведения обследования. При использовании данной сущности обеспечивается выполнение требований ГОСТ Р 71853–2024 [3], предусматривающих сохранение меток времени и источников данных в системах мониторинга и управления. При этом не требуется введение отдельной сущности

диагностического события. Сущность *DataArtifact* позволяет обеспечить трассируемость ДИ от агрегированных диагностических фактов до первичных артефактов (снимков, протоколов испытаний, записей информационных систем). Это обеспечивает возможность воспроизводимой интерпретации результатов диагностики и согласованного анализа данных различных источников.

**ControlMethod** (метод контроля) обеспечивает хранение перечня методов контроля: инфракрасный контроль, хроматографический анализ растворенных газов, физико-химический анализ масла, электрические измерения, визуальный контроль и др. Сущность используется для указания происхождения диагностических параметров.

**DiagnosticParameter** (диагностический параметр) обеспечивает хранение перечня диагностических параметров, формируемых методами контроля. В соответствии с ГОСТ 20911–89 [13] диагностический параметр определяется как параметр объекта, используемый при его диагностировании.

**RecordedParameter** (записанный параметр) обеспечивает хранение фактических значений параметров; связывает конкретные данные (измерения, термограммы и прочее) с соответствующим диагностическим параметром, что обеспечивает разделение между нормативным описанием параметров и их наблюдаемыми реализациями.

**DiagnosticFeature** (диагностический признак) обеспечивает хранение справочника диагностических признаков, формируемые на основе параметров в рамках определенного метода анализа. Признаки отражают результат аналитической обработки данных и служат основой для дальнейшего вывода диагностических фактов. В соответствии с ГОСТ Р ИСО 13372–2013 [15] диагностический признак трактуется как параметр сигнала, несущий информацию о техническом состоянии объекта.

Разделение сущностей *DiagnosticParameter* и *DiagnosticFeature* обусловлено необходимостью различения уровня измеряемых данных и уровня их интерпретации. Один и тот же диагностический параметр может использоваться для формирования различных признаков. Например, термограмма, представляет собой диагностический параметр – измерения, полученные в ходе ИК-контроля. При этом максимальная температура в определенной области (такой как контактное соединение), извлеченная из термограммы, является диагностическим признаком, используемым далее при формировании диагностических фактов. Кроме того, такое разделение соответствует нормативно-технической базе ТД ВЭО [13, 15], разделяющей термины «параметр» и «признак».

**AnalysisMethod** (метод анализа) обеспечивает хранение перечня методов анализа.

**RecordedFeature** (записанный признак) обеспечивает хранение значений признаков. Разделение на типовые сущности и их зафиксированные проявления применяется для параметров и признаков, поскольку они представляют измеряемые или вычисляемые

величины, имеющие конкретные значения во времени.

**DiagnosticFact** (диагностический факт) обеспечивает хранение установленных диагностических фактов, которые отражают выявленные состояния, отклонения или события, имеющие диагностическое значение.

**FactType**, **ConclusionType**, **RecommendationType** обеспечивают хранение типов фактов, выводов и рекомендаций, соответственно, для упрощения поиска

нужной специалисту информации в ХГД, а также группирования результатов запросов к ХГД.

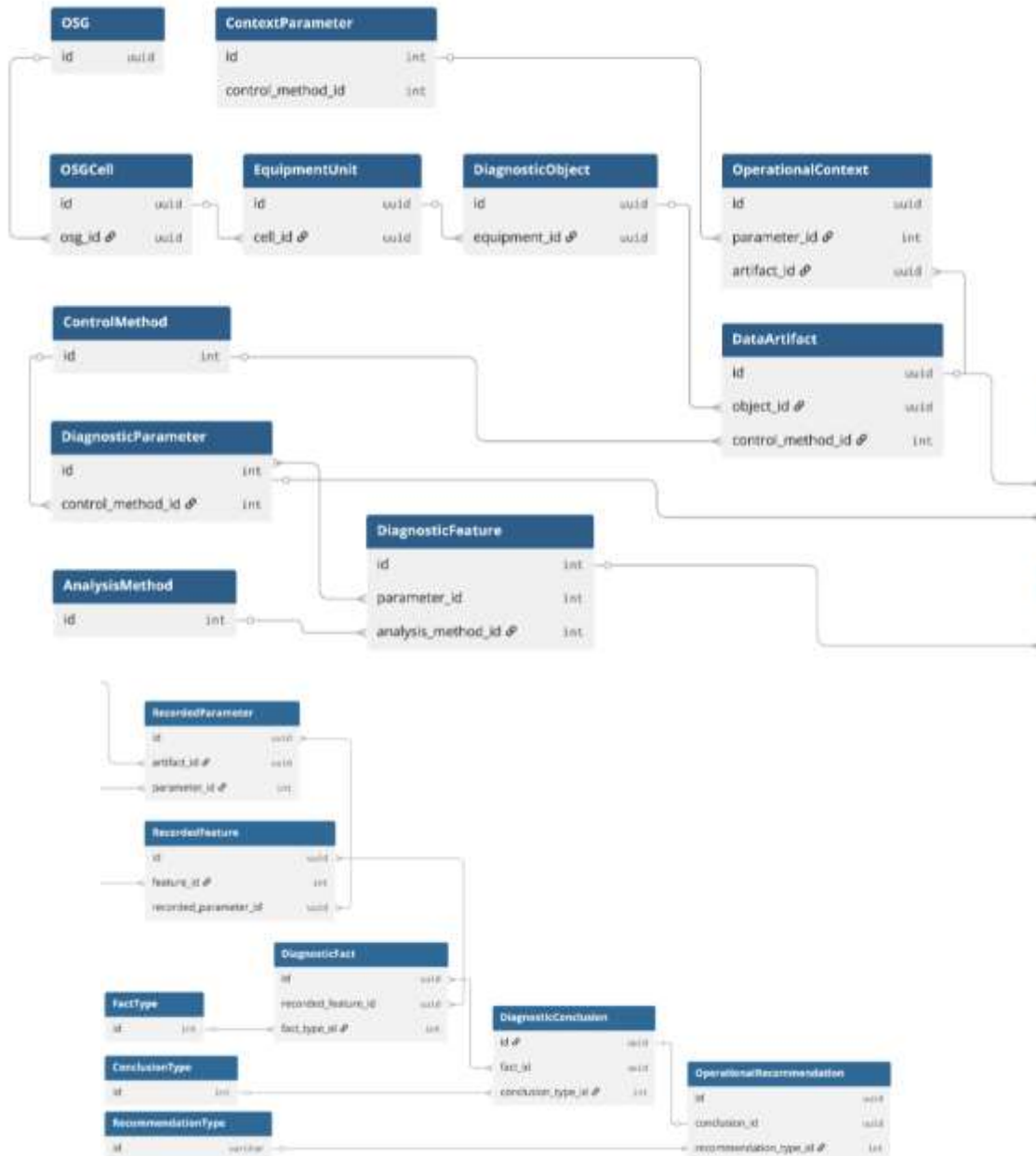


Рис. 4: Концептуальная модель ХГД системы технической диагностики ОРУ типовой электрической станции или подстанции

**ContextParameter** (семантика параметров условий эксплуатации ВЭО) обеспечивает хранение семантических описаний условий эксплуатации и процесса ТД.

**OperationalContext** (контент параметров условий эксплуатации в момент получения ДИ) обеспечивает хранение значений параметров условий эксплуатации в момент получения ДИ, между условиями эксплуатации и выявленными артефактами технического состояния

объекта ТД, с целью контекстно-зависимой интерпретации диагностических результатов, обеспечивающую корректную интерпретацию диагностических данных и предотвращение ошибочные выводы, обусловленных изменениями условий эксплуатации оборудования.

Включение эксплуатационного контекста диагностики позволяет учитывать влияние режимных параметров и внешних факторов на диагностические

проявления. Это обеспечивает корректную интерпретацию диагностических данных и предотвращает ошибочные выводы, связанные с изменениями условий эксплуатации оборудования.

**Diagnostic Conclusion** (диагностический вывод) обеспечивает хранение выводов, формируемых на основе диагностических фактов. Каждый вывод связан с одним или несколькими диагностическими фактами.

**Operational Recommendation** (эксплуатационная рекомендация) обеспечивает хранение рекомендаций, формируемых на основе диагностических выводов, что является завершающим элементом концептуальной цепочки.

Представленная структура отражает иерархию интерпретации диагностической информации. На нижнем уровне располагаются измеряемые диагностические параметры, далее формируются диагностические признаки, на основе которых устанавливаются диагностические факты, характеризующие техническое состояние оборудования. Диагностические факты служат основанием для формирования диагностических выводов и эксплуатационных рекомендаций.

Таким образом, предложенная концептуальная схема ХГД соответствует структурной модели процесса ТД и обеспечивает формализованное представление диагностической информации на различных уровнях интерпретации. При этом связующим звеном между онтологическим описанием предметной области и логической структурой хранения диагностических данных является онтологическая модель предметной области.

Отметим, что в процессе проектирования ХГД были реализованы основные принципы концептуального проектирования БД: разделение уровней представления данных, различие между типовыми сущностями и их зафиксированными проявлениями, вынесение категориальных значений в справочные сущности и минимальность концептуального уровня. Схема полностью основана на ОМ диагностического процесса и нормативных определениях диагностических параметров и признаков, зафиксированных в ГОСТ 20911–89 [13] и ГОСТ Р ИСО 13372–2013 [15], а также учитывает требования к трассируемости диагностических данных, сформулированные в ГОСТ Р 71853–2024 [3] и серии стандартов ГОСТ Р 58651 [14].

В рамках разработанной методологии онтологический уровень рассматривается как первичный, определяющий инвариантную структуру предметной области, тогда как концептуальная и логическая модели формируются как его проекция. Соответственно, выбор модели данных на логическом и физическом уровнях определяется требованиями прикладной задачи и условиями эксплуатации системы.

Использование RDF-хранилищ или графовых БД является возможной альтернативой, однако в рассматриваемой задаче основная нагрузка на ХГД связана с обработкой запросов для формирования наборов гетерогенных диагностических данных.

Реляционная модель в этих условиях обеспечивает

эффективное выполнение сложных аналитических запросов (фильтрация, агрегация, объединение данных) и транзакционную целостность. Дополнительным преимуществом является совместимость с существующей инфраструктурой хранения данных, что позволяет реализовать ХГД без модификации существующих систем (рис. 1) – путем хранения *URI* на артефакты данных в существующих хранилищах.

В связи с этим реляционная модель выбрана в качестве базовой для реализации ХГД при сохранении онтологического уровня как механизма семантического согласования.

При переходе к логической модели для обеспечения 4-й нормальной формы, в ней исключаются связи «многие-ко-многим» между параметрами, признаками, фактами и выводами. Для этого вводятся дополнительные сущности, такие как *FeatureParameter*, *RecordedFeatureParameter*, *FactFeature* и *ConclusionFact*. Данные сущности обеспечивают:

- связывание диагностических признаков с набором параметров, используемых при их формировании;
- фиксацию соответствия между наблюдаемыми признаками и установленными диагностическими фактами;
- возможность формирования диагностических выводов на основе совокупности фактов, но не единичных наблюдений.

Физический уровень хранения данных в системе ХГД ТД ВЭО реализует логическую модель в рамках реляционной системы управления БД и определяет способ размещения, индексирования и доступа к исходным данным. В отличие от концептуального и логического уровней, физическая модель не вводит новых сущностей, а задает параметры хранения и организацию доступа к уже определенным отношениям.

Данные инфраструктурного уровня, представленные сущностями *OSG*, *OSGCell*, *EquipmentUnit*, *DiagnosticObject*, размещаются в отдельных таблицах с кластеризацией по первичным ключам и индексированием по внешним ключам, обеспечивающим иерархическую навигацию по объектам ОПУ. Это позволяет эффективно выполнять выборки по структуре ОПУ, ячейкам и единицам оборудования.

Диагностические данные, представленные сущностями *DataArtifact*, *RecordedParameter* и *RecordedFeature*, хранятся в виде временных записей с обязательной фиксацией метки времени.

Числовые значения параметров и признаков хранятся в типизированных столбцах с поддержкой числовой агрегации, тогда как текстовые и бинарные данные (включая ссылки на внешние файлы протоколов и спектральных данных) представлены в виде строковых идентификаторов или *BLOB*-полей. Сущность *DataArtifact* обеспечивает хранение ссылок (*Uniform Resource Identifier – URI*), устанавливающих связь с внешними файловыми хранилищами (например, хранилищами спектров, изображений, протоколов испытаний). Таким образом, физический уровень хранения реализует гибридную архитектуру:

метаданные и числовые параметры размещаются в системе управления БД, а крупные бинарные объекты могут храниться во внешнем репозитории.

Ассоциативные таблицы (*FeatureParameter*, *FactFeature*, *ConclusionFact* и др.) физически реализуются как компактные таблицы связей с индексированием по обоим внешним ключам, что обеспечивает эффективное разрешение связей типа «многие-ко-многим» и ускоряет построение диагностических цепочек «параметр – признак – факт – вывод – рекомендация».

Следует отметить, что использование сущности *DataArtifact* с атрибутом *URI* обеспечивает возможность интеграции ХГД с существующими на типовом предприятии моделями данных без необходимости изменения их физической структуры и схем хранения. Описанные выше концептуальная и логическая модели ХГД ТД ВЭО определяют физическую модель реляционной БД ХГД, связанной с первичными источниками данных через метаданные и *URI*. Атрибут *URI* сущности *DataArtifact* выполняет функцию унифицированного механизма адресации внешних ресурсов (паспортов электрооборудования, протоколов испытаний, файлов измерений, записей в реляционных БД АИС инспекционных обходов), что позволяет использовать уже функционирующие корпоративные системы хранения данных без их модификации.

#### V. ОЦЕНКА ВЫПОЛНЕНИЯ ТРЕБОВАНИЙ К ХГД В РАССМАТРИВАЕМОЙ ЗАДАЧЕ

Соответствие между перечисленными выше требованиями, предъявляемыми к ХГД ТД ВЭО ОРУ 110–220 кВ, и особенностями его архитектуры представлены в табл. 2.

Из табл. 2 видно, что архитектура ХГД ТД ВЭО соответствует совокупности системных требований, предъявляемых к промышленным информационным моделям, включая иерархическую структурированность уровней представления данных, семантическую согласованность разнородных источников, масштабируемость структуры, эволюционную устойчивость к изменениям нормативной и методической базы, интеграционную совместимость со стандартами и смысловую нормализацию гетерогенных данных.

Табл. 2: Соответствие между установленными требованиями к ХГД и особенностями архитектуры синтезированного ХГД, обеспечивающих данных выполнение

Требование	Механизм обеспечения требования
Системная иерархичность	Разграничение онтологического, концептуального, логического и физического уровней. Изменения локализуются на соответствующем уровне без необходимости структурной переработки всей системы.
Семантическая совместимость	Онтология обеспечивает согласование данных на уровне понятий предметной области, независимо от форматов хранения и исходных моделей данных.
Масштаб-	Логическая структура допускает

ируемость	добавление новых объектов, параметров и методов диагностики без изменения ранее сформированных данных.
Динамичность и эволюционная устойчивость	Структура модели позволяет вводить новые методы анализа, классификаторы и интерпретационные правила при сохранении интерпретируемости ретроспективной информации и непротиворечивости данных.
Интеграция со стандартами	Типологические сущности и классификаторы обеспечивают сопоставление данных с действующими национальными и международными стандартами без жесткой привязки к их конкретным версиям.
Смысловая нормализация гетерогенности	Единое понятийное пространство формируется на основе онтологической модели, временной привязки артефактов диагностики и использования унифицированных идентификаторов и <i>URI</i> , что позволяет интегрировать разнородные данные без изменения их физической структуры.

Интеграции ХГД и внешних источников данных представлена на рис. 5. ХГД обеспечивает объединение диагностических и эксплуатационных данных, формируемых в различных информационных системах и представленных в реляционной, темпоральной и файловой моделях.



Рис. 5: Схема интеграции ХГД и внешних источников данных на физическом уровне

При этом физическая реализация ХГД ТД ВЭО обеспечивает структурную согласованность данных с их семантической и логической моделями за счет разделения хранения метаданных и артефактов данных. Предложенный подход обеспечивает согласованное представление диагностической информации, формируемой различными источниками.

Следует отметить, что за счет наличия временных меток и связей с объектом диагностики обеспечивается накопление диагностических фактов во времени, что позволяет анализировать динамику изменения технического состояния и рассматривать замкнутый контур обратной связи. Временная метка (*DataArtifact.timestamp*) через цепочку связей доступна для сущностей *DiagnosticFact*, *DiagnosticFeature* и *DiagnosticParameter*, так что путем запроса к ХГД можно получить всю последовательность диагностических фактов, выводов и рекомендаций во

времени по единице оборудования. Это позволяет учитывать результаты предыдущих диагностических фактов при формировании последующих диагностических выводов.

Для иллюстрации выполнения сформулированных требований рассмотрим несколько типовых примеров.

**Системная иерархичность.** При переходе предприятия на новую систему хранения диагностических протоколов (из локального файлового хранилища к распределенному объектному) изменяется только физический уровень хранения данных, тогда как логическая и концептуальная модели остаются неизменными, поскольку доступ к диагностическим артефактам осуществляется через сущность *DataArtifact*, содержащую универсальный идентификатор ресурса (*URI*).

**Семантическая совместимость.** Результаты диагностического обследования трансформатора тока могут фиксироваться в протоколе испытаний в виде текстового документа, где оборудование идентифицируется по диспетчерскому наименованию, тогда как в АИС обходов оборудования используется внутренний идентификатор оборудования согласно требованиям ИМ *CIM* [14]. В архитектуре ХГД различные формы идентификации сопоставляются с единой сущностью *EquipmentUnit*, связанной с конкретной ячейкой ОРУ и объектом диагностики, имеющей в качестве атрибутов как диспетчерское наименование, так и идентификатор согласно требованиям *CIM*. Это позволяет интерпретировать данные диагностического протокола, записи обходов и данные *SCADA* как относящиеся к одному и тому же объекту оборудования, обеспечивая их семантическую согласованность независимо от формата хранения и способа идентификации.

**Масштабируемость.** При добавлении в информационную систему предприятия нового энергообъекта (например, ОРУ-500 кВ) в ХГД добавляются новые экземпляры сущностей *OSG*, *OSGCell* и *EquipmentUnit*, соответствующие структуре оборудования данного объекта. При этом структура таблиц диагностических параметров, признаков, фактов и выводов остается неизменной, поскольку модель данных изначально допускает хранение информации о произвольном количестве объектов диагностики и добавление новых признаков и параметров.

**Динамичность и эволюционная устойчивость.** При внедрении нового метода диагностики – акустической диагностики частичных разрядов – в систему добавляются, например, новый диагностический параметр *AcousticEmission* и связанный с ним диагностический признак *AcousticEmissionLevel*. При этом существующие таблицы диагностических параметров, признаков и фактов не требуют изменения структуры, а новые данные интегрируются через существующие сущности модели.

**Интеграция со стандартами.** Как уже отмечено, идентификаторы сущностей *EquipmentUnit* и *OSGCell* могут быть назначены согласно *CIM*. Это позволяет связывать диагностические данные ХГД с данными

корпоративных систем управления активами и эксплуатационных систем предприятия. Кроме того, как показано выше, методология учитывает терминологию стандартов [13] и [15].

**Смысловая нормализация гетерогенности.** Например, при диагностике контактного соединения разъединителя одновременно используются три типа данных: термограмма (файловая модель данных), токовая нагрузка из *SCADA* (темпоральная модель данных) и паспортные характеристики оборудования из корпоративной базы данных (реляционная модель). В ХГД ТД ВЭО эти данные сопоставляются с единым диагностическим объектом и используются совместно для формирования диагностического факта о перегреве контактного соединения.

Подтверждение соответствия программной реализации ХГД ТД ВЭО установленным требованиям (см. табл. 2) является предметом последующих публикаций.

## VI. ЗАКЛЮЧЕНИЕ

Предложен метод решения задачи структурного синтеза ХГД ТД, создаваемых системами ТД, функционирующих на различных физических принципах, в процессе технической диагностики электрооборудования объектов электроэнергетики, основанный на использовании онтологически ориентированной методологии, предусматривающей формализацию предметной области ТД, и семантическое картирование диагностических данных систем и представленной в различных моделях данных.

Разработана структура представления диагностической информации, включающая иерархию сущностей диагностического процесса: диагностические параметры, диагностические признаки, диагностические факты, диагностические выводы и эксплуатационные рекомендации. Такая структура обеспечивает переход от измеряемых параметров к интерпретации технического состояния оборудования и формированию управленческих решений.

Предложенный подход может быть использован при проектировании систем мониторинга и диагностики электрооборудования, а также создании цифровых моделей технического состояния оборудования и интеграции диагностических данных в корпоративные системы управления активами.

## БИБЛИОГРАФИЯ

- [1] Хальясмаа А. И., Матренин П. В. Новая архитектура программного обеспечения визуализации, трансформации и анализа снимков оборудования // International Journal of Open Information Technologies, 2025, Т. 13, № 5, С. 34–40.
- [2] Kim H. J., Jeong C. M., Sohn J.-M. [et al.] A comprehensive review of practical issues for interoperability using the common information model in smart grids // Energies, 2020, vol. 13 (6), art. 1435.
- [3] ГОСТ Р 71853–2024. Единая энергетическая система и изолированно работающие энергосистемы. Система удаленного мониторинга и диагностики оборудования объектов электроэнергетики. Общие требования. М.: Российский институт стандартизации, 2024.
- [4] Codd E. F. A relational model of data for large shared data banks // Communications of the ACM, 1970, vol. 13 (6), pp. 377–387.

- [5] Chen P. P. The entity–relationship model: Toward a unified view of data // *ACM Transactions on Database Systems*, 1976, vol. 1 (1), pp. 9–36.
- [6] Inmon W. H. *Building the data warehouse*. 4th ed. New York: John Wiley & Sons, 2005.
- [7] Kimball R., Ross M. *The data warehouse toolkit: The definitive guide to dimensional modeling*. 3rd ed. Hoboken: John Wiley & Sons, 2013. 608 p.
- [8] Kim W. Object-oriented databases: Definition and research directions // *IEEE Transactions on Knowledge and Data Engineering*, 1990, vol. 2 (3), pp. 327–341.
- [9] Karabulut E., Pileggi S. F., Groth P., Degeler V. Ontologies in digital twins: A systematic literature review // *Future Generation Computer Systems*, 2024, vol. 153, pp. 442–456.
- [10] Sinha P. K., Gajbe S. B., Debnath S., Sahoo S., Chakraborty K., Mahato S. S. A review of data mining ontologies // *Data Technologies and Applications*, 2022, vol. 56 (2), pp. 172–204.
- [11] Андрюшкевич С. К., Ковалев С. П., Нефедов Е. И. Разработка цифрового двойника энергетической системы на основе онтологической модели // *Автоматизация в промышленности*, 2020, № 1, С. 51–56.
- [12] Понкин И. В., Куприяновский В. П., Редькина А. И. К вопросу о содержании понятия и особенностях онтологии энергетического интернета и его правового и технологического образов // *International Journal of Open Information Technologies*, 2018, vol. 7 (8), pp. 87–93.
- [13] ГОСТ 20911–89. Техническая диагностика. Термины и определения. Утв. постановлением Госкомитета СССР по качеству 26.12.1989 № 4143. Введ. 01.01.1991.
- [14] ГОСТ Р 58651 (серия). Информационная модель электроэнергетики. Основные положения. М.: Стандартинформ, 2019.
- [15] ГОСТ Р ИСО 13372–2013. Мониторинг состояния и диагностика машин. Термины и определения. М.: Стандартинформ, 2014.

# Ontology-Based Methodology for the Structural Synthesis of a Heterogeneous Data Warehouse for Diagnostics of High-Voltage Equipment at Power Facilities

A. I. Khalyasmaa, S. V. Porshnev

**Abstract**—This paper considers a methodology for the structural synthesis of a heterogeneous data warehouse (HDW) used to store data accumulated during the technical diagnostics of high-voltage equipment at power facilities. The need for such a methodology arises from the fact that data models traditionally used for storing diagnostic information (temporal, relational, and file-based) do not ensure semantic consistency of the data. As a result, significant difficulties arise in the cause-and-effect interpretation of technical diagnostic results. To ensure semantic consistency of diagnostic data that are heterogeneous in their physical nature, the domain of technical diagnostics of high-voltage electrical equipment is formalized as a system of entities and relationships, together with a set of semantic mapping and ontological normalization operators. A structure for representing diagnostic data is developed that enables the transition from measured parameters to diagnostic facts and operational recommendations. The proposed methodology ensures reproducibility, scalability, and evolutionary stability of the diagnostic data storage structure and can be applied in the construction of digital models of equipment technical condition and in the integration of diagnostic data into corporate asset management systems.

**Keywords**— *technical diagnostics, high-voltage equipment, heterogeneous data warehouse, domain ontology, structural synthesis*

## REFERENCES

- [1] A. I. Khalyasmaa and P. V. Matrenin, “Novaya arkhitektura programmogo obespecheniya vizualizatsii, transformatsii i analiza snimkov oborudovaniya [New software architecture for visualization, transformation, and analysis of equipment images],” *International Journal of Open Information Technologies*, vol. 13 (5), pp. 34–40, 2025.
- [2] H. J. Kim, C. M. Jeong, J.-M. Sohn [et al.], “A comprehensive review of practical issues for interoperability using the common information model in smart grids,” *Energies*, vol. 13 (6), art. 1435, 2020.
- [3] GOST R 71853–2024. Unified energy system and isolated power systems. Remote monitoring and diagnostics system for power industry equipment. General requirements. Moscow: Russian Institute for Standardization, 2024.
- [4] E. F. Codd, “A relational model of data for large shared data banks,” *Communications of the ACM*, vol. 13 (6), pp. 377–387, 1970.
- [5] P. P. Chen, “The entity–relationship model: Toward a unified view of data,” *ACM Transactions on Database Systems*, vol. 1 (1), pp. 9–36, 1976.
- [6] W. H. Inmon, *Building the Data Warehouse*, 4th ed. New York: John Wiley & Sons, 2005.
- [7] R. Kimball and M. Ross, *The Data Warehouse Toolkit: The Definitive Guide to Dimensional Modeling*, 3rd ed. Hoboken: John Wiley & Sons, 2013.
- [8] W. Kim, “Object-oriented databases: Definition and research directions,” *IEEE Transactions on Knowledge and Data Engineering*, vol. 2 (3), pp. 327–341, 1990.
- [9] E. Karabulut, S. F. Pileggi, P. Groth, and V. Degeler, “Ontologies in digital twins: A systematic literature review,” *Future Generation Computer Systems*, vol. 153, pp. 442–456, 2024.
- [10] P. K. Sinha, S. B. Gajbe, S. Debnath, S. Sahoo, K. Chakraborty, and S. S. Mahato, “A review of data mining ontologies,” *Data Technologies and Applications*, vol. 56 (2), pp. 172–204, 2022.
- [11] S. K. Andryushkevich, S. P. Kovalev, and E. I. Nefedov, “Razrabotka tsifrovogo dvoynika energeticheskoy sistemy na osnove ontologicheskoy modeli [Development of a digital twin of an energy system based on an ontological model],” *Avtomatizatsiya v promyshlennosti*, no. 1, pp. 51–56, 2020.
- [12] I. V. Ponkin, V. P. Kupriyanovskiy, and A. I. Redkina, “K voprosu o sodержanii ponyatiya i osobennostyakh ontologii energeticheskogo interneta i ego pravovogo i tekhnologicheskogo obrazov [On the concept and features of the ontology of the energy Internet and its legal and technological aspects],” *International Journal of Open Information Technologies*, vol. 7 (8), pp. 87–93, 2018.
- [13] GOST 20911–89. Technical diagnostics. Terms and definitions. Approved 26.12.1989, entered into force 01.01.1991.
- [14] GOST R 58651 (series). Information model of electric power systems. Basic provisions. Moscow: Standartinform, 2019.
- [15] GOST R ISO 13372–2013. Condition monitoring and diagnostics of machines. Terms and definitions. Moscow: Standartinform, 2014.