
ФЕДЕРАЛЬНОЕ АГЕНТСТВО
ПО ТЕХНИЧЕСКОМУ РЕГУЛИРОВАНИЮ И МЕТРОЛОГИИ



ПРЕДВАРИТЕЛЬНЫЙ
НАЦИОНАЛЬНЫЙ
СТАНДАРТ
РОССИЙСКОЙ
ФЕДЕРАЦИИ

ПНСТ
943—
2024

Искусственный интеллект

**СТРУКТУРА АРХИТЕКТУРЫ
СИСТЕМ МАШИННОГО ОБУЧЕНИЯ
В БУДУЩИХ СЕТЯХ, ВКЛЮЧАЯ IMT-2020**

[ITU-T Y.3172 (2019), NEQ]

Издание официальное

Москва
Российский институт стандартизации
2024

Предисловие

1 РАЗРАБОТАН Научно-образовательным центром компетенций в области цифровой экономики Федерального государственного бюджетного образовательного учреждения высшего образования «Московский государственный университет имени М.В. Ломоносова» (МГУ имени М.В. Ломоносова) и Обществом с ограниченной ответственностью «Институт развития информационного общества» (ИРИО)

2 ВНЕСЕН Техническим комитетом по стандартизации ТК 164 «Искусственный интеллект»

3 УТВЕРЖДЕН И ВВЕДЕН В ДЕЙСТВИЕ Приказом Федерального агентства по техническому регулированию и метрологии от 28 октября 2024 г. № 79-пнст

4 Настоящий стандарт разработан с учетом основных нормативных положений международного документа ITU-T Y.3172 (2019) «Структура архитектуры систем машинного обучения в будущих телекоммуникационных сетях, включая IMT-2020» (ITU-T Y.3172 (2019) «Architectural framework for machine learning in future networks including IMT-2020», NEQ)

Правила применения настоящего стандарта и проведения его мониторинга установлены в ГОСТ Р 1.16—2011 (разделы 5 и 6).

Федеральное агентство по техническому регулированию и метрологии собирает сведения о практическом применении настоящего стандарта. Данные сведения, а также замечания и предложения по содержанию стандарта можно направить не позднее чем за 4 мес до истечения срока его действия разработчику настоящего стандарта по адресу: 119991 Москва, Ленинские горы, д. 1, и/или в Федеральное агентство по техническому регулированию и метрологии по адресу: 123112 Москва, Пресненская набережная, д. 10, стр. 2.

В случае отмены настоящего стандарта соответствующая информация будет опубликована в ежемесячном информационном указателе «Национальные стандарты» и также будет размещена на официальном сайте Федерального агентства по техническому регулированию и метрологии в сети Интернет (www.rst.gov.ru)

© Оформление. ФГБУ «Институт стандартизации», 2024

Настоящий стандарт не может быть полностью или частично воспроизведен, тиражирован и распространен в качестве официального издания без разрешения Федерального агентства по техническому регулированию и метрологии

Содержание

1 Область применения	1
2 Нормативные ссылки	1
3 Термины и определения	1
4 Сокращения	2
5 Соглашения по терминологии	3
6 Общие положения	4
7 Высокоуровневые архитектурные требования	4
8 Структура высокоуровневой архитектуры	12
9 Требования обеспечения безопасности	18
Приложение А (справочное) Реализация высокоуровневой архитектуры в базисных сетях для машинного обучения, зависящих от конкретной технологии.	19
Приложение Б (справочное) Соотнесение архитектурных компонентов и поддерживающих аспектов с требованиями раздела 8.	20
Библиография	22

Федеральное агентство
по техническому регулированию
и метрологии

Федеральное агентство
по техническому регулированию
и метрологии

Введение

Настоящий стандарт определяет структуру архитектуры системы машинного обучения в будущих телекоммуникационных сетях, включая IMT-2020. Представлен набор архитектурных требований и конкретных архитектурных компонентов, необходимых для удовлетворения этих требований. Эти компоненты включают, в частности, конвейер машинного обучения, а также функции управления и оркестровки машинного обучения, но не ограничиваются ими. Также описана интеграция этих компонентов в будущие телекоммуникационные сети, включая IMT-2020, и рекомендации по применению данной структуры архитектуры системы в различных базисных сетях для машинного обучения, зависящих от конкретной технологии.

Искусственный интеллект

СТРУКТУРА АРХИТЕКТУРЫ СИСТЕМ МАШИННОГО ОБУЧЕНИЯ В БУДУЩИХ СЕТЯХ,
ВКЛЮЧАЯ IMT-2020

Artificial intelligence.

Architectural framework for machine learning in future networks including IMT-2020

Срок действия — с 2025—01—01
до 2028—01—01

1 Область применения

Настоящий стандарт устанавливает структуру архитектуры систем машинного обучения, применяемых в будущих телекоммуникационных сетях, в том числе в сетях IMT-2020.

Представлен набор архитектурных требований, необходимых для разработки конкретных архитектурных компонентов, обеспечивающих соответствие описанным требованиям. Настоящий стандарт содержит описание структуры архитектуры системы, необходимой для интеграции подобных компонентов в будущие телекоммуникационные сети, в том числе сети IMT-2020. Дополнительно представлено руководство по применению этой структуры архитектуры системы в различных базисных сетях для машинного обучения, зависящих от конкретной технологии.

2 Нормативные ссылки

В настоящем стандарте использована нормативная ссылка на следующий стандарт:
ГОСТ Р 71476 Искусственный интеллект. Концепции и терминология искусственного интеллекта

Примечание — При пользовании настоящим стандартом целесообразно проверить действие ссылочных стандартов в информационной системе общего пользования — на официальном сайте Федерального агентства по техническому регулированию и метрологии в сети Интернет или по ежегодному информационному указателю «Национальные стандарты», который опубликован по состоянию на 1 января текущего года, и по выпускам ежемесячного информационного указателя «Национальные стандарты» за текущий год. Если заменен ссылочный стандарт, на который дана недатированная ссылка, то рекомендуется использовать действующую версию этого стандарта с учетом всех внесенных в данную версию изменений. Если заменен ссылочный стандарт, на который дана датированная ссылка, то рекомендуется использовать версию этого стандарта с указанным выше годом утверждения (принятия). Если после утверждения настоящего стандарта в ссылочный стандарт, на который дана датированная ссылка, внесено изменение, затрагивающее положение, на которое дана ссылка, то это положение рекомендуется применять без учета данного изменения. Если ссылочный стандарт отменен без замены, то положение, в котором дана ссылка на него, рекомендуется применять в части, не затрагивающей эту ссылку.

3 Термины и определения

В настоящем стандарте применены термины по ГОСТ Р 71476, а также следующие термины с соответствующими определениями:

3.1 базисная сеть для машинного обучения (machine learning underlay network): Телекоммуникационная сеть и связанные с ней сетевые функции, которые взаимодействуют с соответствующими функциями надстройки для машинного обучения.

Примечание — Сеть IMT-2020 является примером базисной сети для машинного обучения.

3.2 конвейер машинного обучения (machine learning pipeline): Набор логических узлов, каждый из которых обладает определенными функциональными возможностями, которые могут быть объединены для формирования приложения машинного обучения в телекоммуникационной сети.

Примечание — Узлы конвейера машинного обучения — это объекты, которые управляются стандартным образом и могут быть размещены в различных сетевых функциях.

3.3 надстройка машинного обучения (machine learning overlay): Слабосвязанная модель развёртывания функциональных возможностей машинного обучения, интеграция и управление которыми с сетевыми функциями стандартизированы.

Примечание — Надстройка машинного обучения направлена на минимизацию взаимозависимостей между функциями машинного обучения и сетевыми функциями благодаря использованию стандартных интерфейсов, обеспечивая тем самым параллельное развитие и тех, и других.

3.4 оркестратор функций машинного обучения; MLFO (machine learning function orchestrator; MLFO): Логический узел с функциями управления и оркестровки узлов в конвейерах машинного обучения.

3.5

оркестровка (orchestration): Тип композиции, где один определенный элемент используется для наблюдения за другими элементами и управления ими.

Примечание — Элемент, который управляет оркестровкой, сам не является частью оркестровки (экземпляра композиции).

[ГОСТ Р ИСО/МЭК 18384-1—2017, пункт 2.16]

3.6 полигон машинного обучения (machine learning sandbox): Среда, в которой можно обучать, тестировать модели машинного обучения и оценивать их влияние на сеть.

Примечание — Полигон машинного обучения предназначен для предотвращения воздействия приложения машинного обучения на сеть или для ограничения использования определенных функций машинного обучения.

3.7 сетевая функция (network function): В контексте IMT-2020 — функция обработки в сети.

Примечания

1 См. [1].

2 Сетевые функции включают в себя функциональные возможности сетевых узлов, например управление сеансом, управление мобильностью и транспортные функции, функциональное поведение и интерфейсы которых определены, но не ограничиваются ими.

3 Сетевые функции могут быть реализованы на выделенном оборудовании или в виде виртуализированных программных функций.

4 Сетевые функции не рассматриваются как ресурсы, скорее любые сетевые функции могут быть реализованы с использованием ресурсов.

4 Сокращения

В настоящем стандарте применены следующие сокращения:

MO	— машинное обучение;
AF	— функция приложения (application function);
AN	— сеть доступа (access network);
API	— интерфейс программирования приложений (application programming interface);
AR/VR	— дополненная реальность/виртуальная реальность (augmented reality/virtual reality);
CN	— базовая сеть (core network);
FMC	— конвергенция стационарной и мобильной связи (fixed mobile convergence);

GPU	— графический процессор (graphic processor unit);
IMT-2020	— международные мобильные телекоммуникации-2020 (International Mobile Telecommunications-2020);
IoT	— интернет вещей (internet of things);
KPI	— ключевой показатель эффективности (key performance indicator);
MEC	— периферийные вычисления мультисервисного доступа (multi-access edge computing);
mIoT	— массивный интернет вещей (massive Internet of things);
MLFO	— оркестратор функций машинного обучения (machine learning function orchestrator);
MnS	— служба управления (management service);
MPP	— прогнозирование закономерностей мобильности (mobility pattern prediction);
NF	— сетевая функция (network function);
NOP	— сетевой оператор (network operator);
OAM	— эксплуатация, администрирование и техническое обслуживание (operations, administration and maintenance);
QoS	— качество обслуживания (quality of service);
RCA	— анализ первопричин (root cause analysis);
RRC	— управление радиоресурсами (radio resource control);
SBA	— сервис-ориентированная архитектура (service-based architecture);
SMF	— функция управления сеансом (session management function);
SON	— самооптимизирующаяся сеть (self-optimizing network);
UE	— пользовательское оборудование (user equipment);
UPF	— функция плоскости пользователя (user plane function);
VoLTE	— голосовая связь в сетях LTE (Voice over Long-Term Evolution);
V2X	— система обмена информацией между транспортным средством и его окружением (vehicle-to-everything).

5 Соглашения по терминологии

В настоящем стандарте приняты следующие соглашения.

Ключевые слова «требуется, чтобы» означают требование, которое должно строго соблюдаться и отклонение от которого не допускается, если будет сделано заявление о соответствии настоящему стандарту.

Ключевое слово «рекомендуется» означает требование, которое рекомендуется, но не является абсолютно необходимым. Таким образом это требование не является обязательным для заявления о соответствии настоящему стандарту.

Ключевое слово «опционально» указывает на необязательное требование, которое допустимо, но не предполагает какого-либо свойства рекомендации. Этот термин не подразумевает, что реализация поставщика должна предоставлять такую возможность и эта функция может быть дополнительно включена сетевым оператором/поставщиком услуг. Скорее это означает, что поставщик может дополнительно предоставить функцию и по-прежнему заявлять о соответствии спецификации.

В основной части настоящего стандарта и приложениях к нему иногда встречаются такие слова, как «должен», «не должен», «следует» и «может», и в этом случае их следует интерпретировать соответственно как «требуется», «запрещено», «рекомендуется» и «опционально». Появление таких фраз или ключевых слов в приложении или в материалах, явно помеченных как информативные, следует интерпретировать как не имеющее нормативного характера.

Конвейер МО — в настоящем стандарте символ, показанный на рисунке 1, обозначает подмножество (включая собственное подмножество) узлов в конвейере МО. Когда данный символ используется в рисунке, он обозначает подмножество узлов конвейера МО, которые явно не показаны на этом рисунке.



Рисунок 1 — Символ, используемый для обозначения подмножества узлов в конвейере МО

Точки выхода и входа сервиса — в настоящем стандарте узел и его точки выхода и входа сервиса обозначаются символом, приведенным на рисунке 2. Точка выхода сервиса отображается зеленым прямоугольником в верхней части узла, а точка входа сервиса соответствует красному прямоугольнику в верхней части узла.



Рисунок 2 — Символ, используемый для обозначения узла с его точками выхода и входа сервиса

6 Общие положения

МО позволяет научить вычислительные системы получать знания из данных без необходимости явного программирования для реализации сложных задач, таких как обнаружение характеристик или прогнозирование поведения. Поскольку МО становится важной технической тенденцией в отрасли, сетевые операторы и другие заинтересованные стороны ищут экономически эффективные способы включения МО в будущие телекоммуникационные сети, включая IMT-2020.

Хотя преимущества такой интеграции обсуждались во многих случаях использования [например, устранение неполадок в сети, прогнозирование сетевого трафика, настройка оптимизации трафика, аудит информационной безопасности (см. [2], [3])], такая интеграция сталкивается со множеством проблем. Проблемами являются:

- а) гетерогенный характер функций МО и уникальные характеристики будущих коммуникационных технологий предъявляют разнообразный набор требований к интеграции;
- б) дорожные карты развития этих функций МО и сетей связи не согласованы;
- в) стоимость интеграции с точки зрения влияния на архитектуру является важным фактором;
- г) разные механизмы управления функциональными возможностями МО и сетевыми функциями (см. [1]) нарушат управление операциями сетей связи.

Для решения этих проблем предназначена структура архитектуры системы для интеграции МО с будущими телекоммуникационными сетями, включая IMT-2020.

Основываясь на высокоуровневых архитектурных требованиях, такая структура сначала предоставляет общий словарь и номенклатуру для функций МО и их взаимосвязей с сетями связи. Эта структура определяет контрольные точки для будущих телекоммуникационных сетей, включая IMT-2020, которые обеспечивают слабосвязанную интеграцию с функциями МО. Механизмы управления МО также описаны в структуре с учетом существующих принципов управления, которые были определены в сетях IMT-2020.

Эта структура обеспечивает стандартный метод интеграции функций МО в будущих телекоммуникационных сетях, включая IMT-2020.

7 Высокоуровневые архитектурные требования

В этом разделе представлены высокоуровневые требования к проектированию структуры архитектуры системы по разделу 8.

Данные высокоуровневые архитектурные требования классифицируются следующим образом:

- средства корреляции данных между уровнями и гетерогенными технологиями: этот набор требований адресован проблеме, приведенной в разделе 6, перечисление а), входящие в него требования представлены в 7.1.

Примечание — Уровни могут соответствовать различным сегментам сети, где контрольные точки между ними определены в соответствующих спецификациях сетевой архитектуры. Например, [4] определяет контрольную точку RP-ап между AN и CN, так что AN и CN могут соответствовать разным уровням сети IMT-2020;

- средства развертывания: этот набор требований адресован проблеме, приведенной в разделе 6, перечисление б), входящие в него требования представлены в 7.2;

- требования, относящиеся к интерфейсам между архитектурными компонентами: этот набор требований адресован проблемам, приведенным в разделе 6, перечисления б) и в), входящие в него требования представлены в 7.3;

- требования, связанные с декларативными спецификациями, используемыми для спецификации приложений МО: этот набор требований адресован проблемам, приведенным в разделе 6, перечисления в) и г), входящие в него требования представлены в 7.4;

- требования, связанные с управлением архитектурными компонентами: этот набор требований адресован проблеме, приведенной в разделе 6, перечисление г), входящие в него требования представлены в 7.5.

7.1 Средства корреляции

В таблице 1 представлены требования к средствам корреляции данных между уровнями и гетерогенными технологиями.

Таблица 1 — Высокоуровневые требования. Средства корреляции

REQ-ML-COR-001	Рекомендуется использовать архитектуру МО для корреляции данных, поступающих из нескольких источников
Описание	В будущих телекоммуникационных сетях источники данных могут быть гетерогенными, интегрированными с различными NF (сетевые функции) и могут генерировать данные в разных форматах. Эти разнообразные «перспективы» могут дать ценную информацию при корреляционном анализе. Необходимы архитектурные компоненты, позволяющие функциям МО собирать и сопоставлять данные из различных источников в сети
Примечания	1 Например, анализ данных UE (пользовательское оборудование), AN (сеть доступа), CN (базовой сети) и AF (функция приложения) необходим для прогнозирования потенциальных проблем, связанных с QoS (качество обслуживания) в сквозных пользовательских потоках. 2 Другими примерами таких источников данных являются функциональные возможности SON (самооптимизирующаяся сеть), которые отслеживают и коррелируют сетевые аварийные сигналы и KPI (ключевые показатели эффективности) (см. [5]), а затем предпринимают соответствующие действия для устранения аварийных сигналов или улучшения качества KPI сети или дают рекомендации по проектированию сети без вмешательства человека
REQ-ML-COR-002	а) Архитектура МО требуется для поддержки многочисленных технологий будущих телекоммуникационных сетей, включая IMT-2020, для обеспечения сквозного взаимодействия с пользователем. б) Рекомендуется, чтобы архитектура МО взаимодействовала с внешними функциональными объектами, не относящимися к IMT-2020, для обеспечения сквозного взаимодействия с пользователем
Описание	В будущих телекоммуникационных сетях будет сосуществовать множество технологий, например лицензионные и нелицензионные беспроводные технологии, технологии FMC (конвергенция стационарной и мобильной связи), унаследованные и будущие технологии. Появление сегментации сети (см. [6]) является одним из примеров, в котором важны вертикальные технологии (сети, адаптированные для предоставления гибких решений для различных рыночных сценариев) и их интеграция в будущие телекоммуникационные сети. Взаимодействие архитектуры МО с такими функциональными объектами поможет в достижении KPI, как, например, указано в [5]. Таким образом, важно, чтобы архитектура МО была способна интегрироваться с несколькими базовыми технологиями и даже поддерживать функции приложений

Окончание таблицы 1

Примечания	<p>1 V2X (система обмена информацией между транспортным средством и его окружением) является примером вертикального приложения, которое может получить выгоду от поддержки сегментации сети.</p> <p>2 Различные технологии сетей связи (например, 3G, 4G, 5G) могут рассматриваться как примеры базовых технологий. Приложения для организации развлечений в автомобиле, обработки данных с дронов, развлекательные приложения с использованием AR/VR (дополненная/виртуальная реальность) являются примерами функций приложений.</p> <p>3 Примерами сетевых функций, которые не управляются напрямую NOP (сетевой оператор), являются датчики, силовые цепи и различные модули IoT (интернет вещей). В некоторых случаях данные таких сетевых функций используются для управления и мониторинга как сетевых функций, развернутых в сети, так и внешних функциональных объектов сети. Эти данные также могут затем использоваться NOP в различных типах оптимизации сетевых параметров для достижения увеличения покрытия, пропускной способности и качества</p>
REQ-ML-COR-003	Архитектура MO требуется для поддержки распределенной реализации функций MO на нескольких уровнях
Описание	Распределенная реализация помогает использовать данные с разных уровней для обогащения (при необходимости) локально доступных данных с одного уровня. Таким образом, функциональные возможности MO могут быть многоуровневыми. Независимые экземпляры функций MO могут создаваться на нескольких уровнях
Примечания	<p>1 В некоторых вариантах использования источник данных может размещаться в AN, а предварительно обработанные данные использоваться моделью MO, которая находится в CN [7].</p> <p>2 Функциональные возможности MO могут использоваться для MPP (прогнозирование закономерностей мобильности) или конфигурации сегмента сети [1] с использованием входных данных из сети на нескольких уровнях (от источников в AN или CN)</p>

7.2 Средства развертывания

В таблице 2 представлены требования к средствам развертывания функций MO.

Т а б л и ц а 2 — Высокоуровневые требования. Средства развертывания

REQ-ML-DEP-001	Рекомендуется использовать архитектуру MO для определения точек взаимодействия между функциями MO и NF (сетевые функции) базисных сетей для MO, зависящих от конкретной технологии, вне зависимости от функций приложения MO
Описание	Спецификация приложений MO в будущих телекоммуникационных сетях будет связана с существующими (или новыми) сетевыми службами или сетевыми функциями. На основе спецификации приложений MO определяются размещение и характеристики функций MO. Источником данных и целью результатов MO являются точки тесной интеграции с NF базисной сети для MO, зависящими от конкретной технологии. Помимо этого, функции MO могут быть общими и не иметь тесной интеграции с NF, зависящими от конкретной технологии. Требуется четко определенные контрольные точки между функциями MO и NF базисной сети для MO, зависящей от конкретной технологии
Примечания	<p>1 Результатом MO могут быть политики или конфигурации, применяемые в сети, а целью результатов MO могут быть функции в сети для применения результатов MO. Такое применение результатов MO может контролироваться политикой сетевого оператора.</p> <p>2 В некоторых вариантах использования источник данных и классификация трафика (на основе результатов MO) могут быть размещены в плоскости пользователя сети, например UPF (функция плоскости пользователя) [4]. Их можно рассматривать как точки тесной интеграции между функциями MO и базисной сетью для MO [например, AN (сеть доступа) или CN (базовая сеть)]. Другая функциональность MO (например, модель MO) не имеет таких зависимостей интерфейса от базисных сетей для MO</p>

Окончание таблицы 2

REQ-ML-DEP-002	Архитектура МО требуется для поддержки отдельного или комбинированного развертывания функций МО в различных NF базисных сетях для МО
Описание	В будущих телекоммуникационных сетях функции управления и оркестровки [6] будут соответствующим образом оптимизировать расположение и производительность NF. Более того, ограничения, применимые к функциям МО, могут быть уникальными
Примечания	1 Для МО может потребоваться GPU (графический процессор) и изолированная среда для исключения влияния на другие функции сети. 2 Например, в зависимости от времени задержки, доступности данных и других особенностей приложений МО, функциональные возможности МО могут иметь источник данных и проводить обучение в CN или в AN
REQ-ML-DEP-003	Архитектура МО требуется для поддержки гибкого размещения функций МО (в координации с функциями управления и оркестровки [1], [8]) в NF базисных сетях для МО
Описание	При выделении ресурсов для функций МО требуется учитывать различные ограничения (например, ограничения ресурсов NF, ограничения задержки, специфические для конкретного приложения МО, доступность данных, зависящих от конкретного приложения МО, возможности преобразования данных, ограничения производительности, возможности обучения и характеристики модели). Архитектура МО должна обеспечивать возможность гибкого размещения функций МО в сети, оптимального для производительности приложений МО и на основе ограничений, определенных в декларативных спецификациях приложений МО. Ограничения для онлайн-обучения и прогнозирования для приложений МО в реальном времени, которые зафиксированы в спецификации, являются входными данными для размещения функций МО в сети, которые могут обеспечить оптимальную производительность для конкретного варианта использования
Примечания	1 Если модель МО включает нейронную сеть, то желательно ее разместить в системе из графических процессоров. 2 В зависимости от требований приложений МО функциональные возможности МО, обеспечивающие чувствительные к задержке краткосрочные прогнозы, могут размещаться ближе к периферии. На размещение также может влиять фактор доступности данных. Это можно скоординировать с разделением функций МО, упомянутым в REQ-ML-DEP-002. 3 В некоторых вариантах использования классификация данных плоскости пользователя может выполняться с использованием моделей МО. Поскольку такая классификация является приложением, чувствительным к задержке, модель МО может размещаться в транспортной сети, а обучение выполняться в CN
REQ-ML-DEP-004	Архитектуре МО требуется поддерживать подключение и отключение новых источников данных или целевых объектов конфигурации в работающую среду МО
Описание	Некоторые продвинутое приложения в будущих телекоммуникационных сетях, например mIoT (массивный интернет вещей), требуют обработки неструктурированных данных из огромного количества источников данных, которые могут быть готовы к использованию по принципу «включай и работай». Одним из таких вариантов использования является анализ зарегистрированных данных для обнаружения аномалий в сетях
Примечания	1 Масштабирование функций МО в зависимости от типа и объема входящих данных является примером обработки новых источников данных. 2 Конфигурация функций МО в сети должна обеспечивать подключение и отключение новых источников данных или целевых конфигураций на основе метаданных об источниках данных

7.3 Связь с интерфейсом

В таблице 3 приведены требования, связанные с интерфейсом.

Т а б л и ц а 3 — Требования высокого уровня, связанные с интерфейсом

REQ-ML-INT-001	Рекомендуется, чтобы архитектура МО поддерживала интерфейс для передачи обученных моделей между функциями МО нескольких уровней
Описание	Обучение моделей имеет определенные специфические потребности, например доступность определенных типов процессоров, доступность определенных видов обучающих данных. После завершения обучения модель МО должна быть отправлена в базисную сеть для МО, зависящую от конкретной технологии, и в которой размещена модель МО. Обучение модели может проводиться отдельно от эксплуатируемой сети. Таким образом, отправка обученных моделей на несколько уровней является важным требованием
Примечания	1 UE (пользовательское оборудование), AN (сеть доступа) и CN (базовая сеть) можно рассматривать как примеры сетевых функций базисных сетей для МО, зависящих от конкретной технологии, в которых может размещаться обученная модель. 2 CN и AN можно рассматривать как примеры нескольких уровней. 3 В зависимости от доступности данных, обучение может проводиться в CN или AN, а сама обученная модель (например классификатор) размещается в транспортных сетях
REQ-ML-INT-002	Архитектуре МО требуется поддерживать интерфейс передачи данных для обучения или тестирования моделей между функциями МО нескольких уровней
Описание	Определенные уровни с доступными данными могут не иметь возможности для обучения. В таких случаях может возникнуть необходимость в отправке данных для обучения или тестирования на уровни, имеющие такие возможности
Примечания	1 AN можно рассматривать как примеры сетей с ограниченными ресурсами, тогда как CN можно рассматривать как уровень, на котором может быть доступен ресурс для масштабирования. 2 Предварительная обработка данных может быть выполнена в транспортной сети; предварительно обработанные данные далее отправляются для обучения в модель, расположенную в CN
REQ-ML-INT-003	а) При определении интерфейса с базисными сетями для МО в архитектуре МО рекомендуется использовать существующие стандартные протоколы везде, где это возможно, с требуемыми расширениями всюду, где это необходимо. б) Рекомендуется, чтобы архитектура МО поддерживала специальные интерфейсы или API (интерфейсы программирования приложений) для взаимодействия с зависящими от конкретной технологии NF (сетевые функции), для получения данных от них или для их конфигурирования. в) Рекомендуется, чтобы архитектура МО поддерживала логические интерфейсы между функциональными возможностями МО, которые могут быть размещены на нескольких уровнях, а реализация таких логических интерфейсов осуществлялась в соответствии со сценариями развертывания
Описание	Источникам данных и целям результатов МО могут потребоваться специальные интерфейсы или интерфейсы программирования приложений API с базисными сетями для МО для извлечения данных или применения конфигураций. В некоторых вариантах использования может потребоваться тесная связь на этапе интеграции между источником и целью функциональных возможностей МО и NF. В некоторых случаях может потребоваться расширение таких интерфейсов для достижения функциональных возможностей МО в конкретном варианте использования. Функциональные возможности МО могут использовать интерфейсы, предоставляемые базисной сетью для МО, в качестве источника данных или цели конфигураций. В этом смысле такие сетевые API могут выступать как реализация интерфейса к источнику и цели функциональных возможностей МО

Окончание таблицы 3

Примечания	<p>1 В пункте а) некоторые интерфейсы могут быть реализованы путем повторного использования существующих протоколов (таких как RRC (управление радиоресурсами) [4], MEC (периферийные вычисления мультисервисного доступа) [9], MnS (служба управления) [10]).</p> <p>2 Например, в пункте б) источник, работающий с UE, может использовать специальные API для извлечения данных из клиентского приложения, обеспечивающего голосовую связь в сетях LTE (VoLTE) [11].</p> <p>3 Например, в пункте в), например в случае, когда источник работает в AN, но требует измерений от UE, AN должна настроить UE для этих измерений с помощью протокола RRC</p>
REQ-ML-INT-004	В архитектуре MO рекомендуется поддерживать совместное использование данных между функциональными возможностями MO на основе механизмов распределенного межуровневого совместного использования данных
Описание	Межуровневое совместное использование данных необходимо для принятия коррелированных решений MO в будущих телекоммуникационных сетях
Примечание	Такие концепции, как озера данных, появляются в будущих облаках и также могут быть использованы в облаках сетевых операторов

7.4 Декларативные спецификации

В таблице 4 приведены требования, касающиеся декларативных спецификаций.

Таблица 4 — Требования высокого уровня. Декларативная спецификация

REQ-ML-SPEC-001	Архитектуре MO требуется поддерживать стандартный метод представления приложений MO, который может быть преобразован в функциональные возможности MO в NF (сетевые функции) базисных сетей для MO, зависящих от конкретных технологий
Описание	Характерной чертой будущих телекоммуникационных сетей является автоматизация с использованием декларативной спецификации и соответствующего преобразования в необходимые конфигурации. При распространении данного метода на MO необходимо поддерживать декларативную спецификацию приложений MO и соответствующие преобразования в конфигурации функциональных возможностей MO
Примечание	Интерпретация декларативной спецификации позволяет конфигурировать функциональные возможности MO в сети путем преобразования спецификации в конфигурацию, которая может быть размещена сетевыми функциями
REQ-ML-SPEC-002	Архитектуре MO требуется поддерживать приложения MO для спецификации источников данных, репозиториях моделей MO, целевых результатов моделей MO и ограничений на сетевые ресурсы. Архитектуре MO требуется поддерживать требования к ограничениям по времени, предъявляемым к приложениям MO
Описание	Разделение между технологически независимой частью приложения MO и его развертыванием, зависящим от конкретной технологии, учитывается во время разработки сервисов будущих телекоммуникационных сетей. Декларативные спецификации для приложений MO обеспечивают это разделение. Различные приложения MO имеют различные временные ограничения. Эти ограничения формируют важные входные данные для функций управления и оркестровки при определении местоположения, объединения в цепочку и мониторинга функциональных возможностей MO

Окончание таблицы 4

Примечания	<p>1 Например, в пункте а) таблицы 3 описание, созданное на метаязыке, может отражать требования оператора сети к приложению МО. Функции управления и оркестровки могут преобразовать их в конфигурацию, которая может быть реализована в сети.</p> <p>2 В пункте б) таблицы 3 в самом узком масштабе применение МО в сетевых функциях формирования луча, планирования, адаптации канала связи будет иметь критерии задержки порядка микросекунд, тогда как для сетевых функций транспортных и базовых сетей критерии задержки составляют несколько миллисекунд. Наименее требовательными с точки зрения задержки являются функции на уровне управления, например обнаружение аномалий и «дыр» в покрытии, для которых допустима задержка на минуты, часы или дни</p>
REQ-ML-SPEC-003	Архитектура МО требуется для поддержки гибкого разделения функциональных возможностей МО на основе спецификаций приложений МО
Описание	<p>Спецификация приложения МО является важным вкладом в развертывание МО в будущих телекоммуникационных сетях, включая IMT-2020. Но сетевые возможности могут изменяться (аппаратное обеспечение может быть добавлено или удалено), NF могут быть запланированы или динамически (повторно) настроены с помощью функций управления и оркестровки. Эти динамические изменения могут потребовать изменения в разделении и местоположении функциональных возможностей МО (например, может быть принято решение о совместном размещении определенных функций МО на основе изменений в пропускной способности канала, или может быть принято решение о добавлении нового источника данных на основе применения функций управления и оркестровки). Таким образом, необходима комбинация:</p> <p>а) исходных данных из спецификации;</p> <p>б) требований к функциям МО, которые допускают отдельное и комбинированное развертывание;</p> <p>в) координации с базовой функцией управления и оркестровки</p>
Примечание	Новый источник данных может быть реализован на основе решений о масштабировании в сети
REQ-ML-SPEC-004	Архитектура МО требуется для поддержки спецификаций приложений МО третьими сторонами для определения источников данных, репозиториях моделей МО, целевых результатов моделей МО и ограничений на сетевые ресурсы
Описание	Сторонние [1] поставщики услуг могут предлагать инновационные услуги поверх будущих телекоммуникационных сетей. Данные услуги могут включать в себя новые алгоритмы МО. Сотрудничество между сторонними поставщиками и сетевыми операторами может привести к появлению новых источников данных или возможностей для агрегирования. Декларативный процесс в архитектуре должен расширять возможности для включения сторонних поставщиков, способных включить эти функциональные возможности в спецификацию, чтобы конечные пользователи могли воспользоваться инновационными услугами, предлагаемыми сторонними поставщиками
Примечания	<p>1 В некоторых случаях модели МО могут предоставляться в виде приложений сторонних производителей.</p> <p>2 Приложение для смартфона, которое взаимодействует с датчиками на пользовательском оборудовании, является примером стороннего источника данных.</p> <p>3 Сторонний поставщик услуг передачи голосовых сообщений с использованием интернет-протокола (Voice over IP) хочет оптимизировать качество звонков по сети, запустив приложение МО, конфигурирующее параметры сети. Спецификация такого приложения МО является примером спецификации стороннего поставщика</p>

7.5 Управление функциональными возможностями МО

В таблице 5 приведены требования, касающиеся управления функциональными возможностями МО.

Таблица 5 — Требования высокого уровня. Управление функциональными возможностями МО

REQ-ML-MNG-001	Архитектуре МО требуется поддерживать выбор модели МО во время настройки функциональных возможностей МО
Описание	Достижения в области МО предполагают, что в будущих телекоммуникационных сетях будут использоваться модели МО с различными характеристиками (например, с использованием различных методов оптимизации и весовых коэффициентов), подходящие для различных проблемных пространств и характеристик данных. В будущих телекоммуникационных сетях новые источники данных могут добавляться динамически. Чтобы распространить приложения МО на новые и гетерогенные источники данных, выбор модели МО должен осуществляться динамически, на основе данных, предоставленных источниками
Примечание	Подключение новых сетевых функций (например, новой UPF (функция в плоскости пользователя)) к эксплуатируемой сети может служить примером динамического подключения новых источников данных
REQ-ML-MNG-002	Архитектуре МО требуется поддерживать обучение и обновление модели МО, предотвращая при этом воздействие на сеть
Описание	При обучении модели МО учитывается несколько факторов: использование специального оборудования для повышения скорости, доступность данных, оптимизация параметров, избежание предвзятости, распределение обучения и т. д. Кроме того, в будущих телекоммуникационных сетях следует избегать сбоев в обслуживании во время обучения модели и ее обновления
Примечание	—
REQ-ML-MNG-003	Архитектуре МО требуется поддерживать возможности мониторинга сетевых операций на основе эффекта МО и обновления моделей МО и/или политик без ущерба для сети
Описание	Необходимо отслеживать эффекты воздействия МО на сеть. Различные KPI постоянно подвергаются измерению, поэтому влияние МО на них, а также на сами функции МО необходимо постоянно отслеживать и при необходимости корректировать. Функциональные возможности МО необходимо обучить для будущего распознавания и обработки подобных скорректированных сценариев
Примечание	Примером процесса, в котором осуществляется мониторинг и обновление в сети, является непрерывное совершенствование рабочих процессов автоматизированного восстановления после сбоев. RCA (анализ первопричин) процесса восстановления после сбоя предоставляется для настройки функций управления и оркестровки, а также оценивается эффект, производимый МО в сети, для последующей оптимизации самих функций МО
REQ-ML-MNG-004	Архитектуре МО требуется поддерживать функциональность оркестровки для управления всеми функциональными возможностями МО в сети
Описание	Производительность функциональных возможностей МО в сети отслеживается, и когда производительность падает ниже заданного порога, функциональные возможности МО перенастраиваются для ее улучшения. Разнообразные источники данных (такие как CN (базовая сеть), AN (сеть доступа)) приводят к тому, что могут применяться различные методы обучения, включая распределенное обучение. Сложные модели могут быть обучены с использованием разнообразных данных. Производительность таких моделей можно определить и сравнить в изолированной среде. Основываясь на таких сравнениях, операторы сетей могут затем выбрать соответствующие функциональные возможности МО (на основе внутренних политик) для конкретных приложений МО

Окончание таблицы 5

Примечания	1 Оценка включает в себя оценку производительности сети наряду с производительностью алгоритмов МО. 2 Повторный выбор модели МО является примером перенастройки, выполненной для повышения производительности функциональных возможностей МО
REQ-ML-MNG-005	Архитектуре МО требуется поддерживать гибкое объединение функциональных возможностей МО в цепочку, включая многоуровневую цепочку. Требуется гибкое объединение функциональных возможностей МО в цепочку в зависимости от хостинга и позиционирования в разных NF (сетевые функции) и на разных уровнях. Это делается для возможности использования гибридных или распределенных функций МО
Описание	Объединение функциональных возможностей МО в цепочку может быть использовано для создания комплексных функциональных возможностей МО. Функции управления и оркестровки предоставляют NOP (сетевые операторы) возможности для быстрого проектирования, разработки и развертывания сетевых сервисов в базисных сетях для МО, зависящих от конкретных технологий. Аналогичным образом функции МО в сети нуждаются в механизмах, включающих гибкое объединение в цепочку, чтобы соответствовать инновациям в базисных сетях для МО. По мере того, как сетевые сервисы базисных сетей для МО быстро развиваются и развертываются, то же самое происходит и с функциональными возможностями МО поверх них благодаря использованию гибких методов объединения в цепочку. Это требование направлено на обеспечение функциональным возможностям МО способности адаптироваться к динамическому созданию сервисов и оркестровке в базисных сетях для МО
Примечание	—

8 Структура высокоуровневой архитектуры

Структура высокоуровневой архитектуры, описанная в данном разделе, поддерживает требования, перечисленные в разделе 7.

В этом разделе приводится описание высокоуровневой архитектуры и ее компонентов, а также даются рекомендации по реализации высокоуровневой архитектуры в различных базисных сетях для МО, зависящих от конкретной технологии.

Такой подход к проектированию архитектуры обеспечивает возможность анализа как решений МО, инвариантных относительно базисных сетей для МО, зависящих от конкретных технологий, так и специфических вопросов интеграции МО в такие базисные сети для МО. Пример реализации этой структуры архитектуры системы в сетях, зависящих от конкретной технологии, представлен в приложении А.

8.1 Высокоуровневые архитектурные компоненты

В данном разделе определены высокоуровневые архитектурные компоненты, которые являются существенными частями структуры архитектуры системы. Интеграция таких компонентов в архитектуру сети путем взаимодействия с NF (сетевые функции), наряду с определением местоположения функциональных возможностей МО в такой сети, образует структуру архитектуры системы.

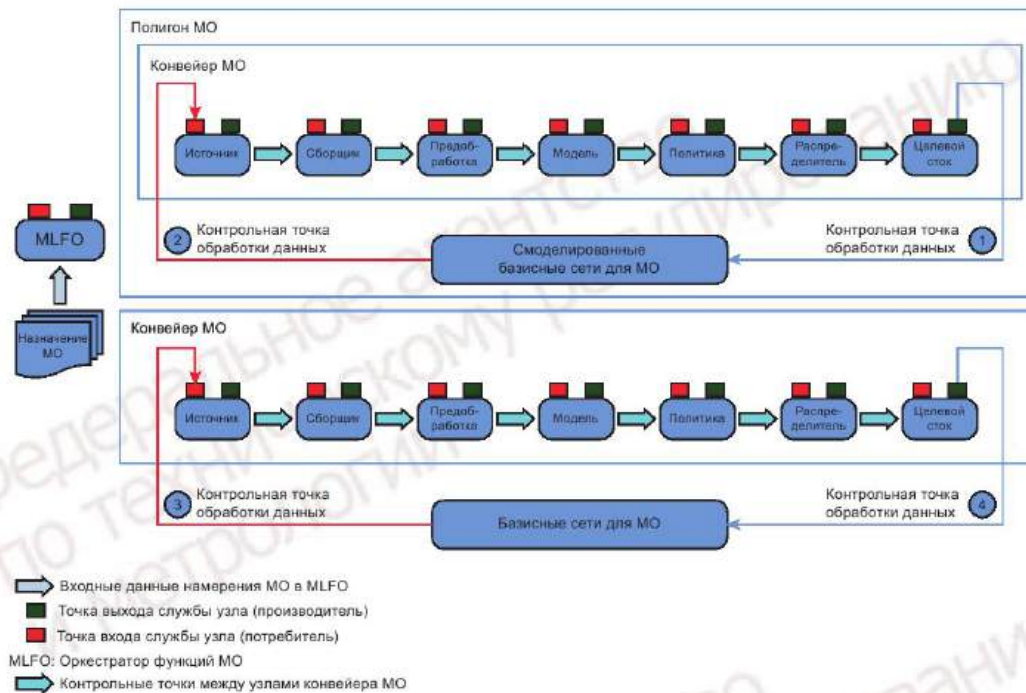


Рисунок 3 — Компоненты высокоуровневой архитектуры

В 8.1.1—8.1.3 описаны архитектурные компоненты.

8.1.1 Конвейер машинного обучения

Как определено в разделе 3, конвейер МО — это набор логических узлов, каждый из которых обладает определенными функциональными возможностями, которые могут быть объединены для формирования приложения МО в телекоммуникационной сети.

Примечание — Конвейер МО может быть развернут в моделируемых или эксплуатируемых базисных сетях для МО.

Узлы конвейера МО:

- источник: этот узел является источником данных, которые могут быть использованы в качестве входных данных для конвейера МО.

Примечание — Потенциальные узлы источника включают UE (пользовательское оборудование), SMF (функция управления сеансом) [4] и AF (функция приложения) [4];

- сборщик: этот узел отвечает за сбор данных из одного или нескольких источников.

Примечание — Узел сборщика может иметь возможность настраивать узлы источников. Например, протокол RRC (управление радиоресурсами) [10] может использоваться для настройки UE, действующего в качестве узла источника. Узел сборщика может использовать специфичные для поставщика протоколы OAM (эксплуатация, администрирование и техническое обслуживание) для настройки SMF (функция управления сеансом), действуя в качестве узла источника. Такие конфигурации могут использоваться для управления характером данных, их детализацией и периодичностью при их генерации из источника;

- предобработка: этот узел отвечает за очистку данных, агрегирование данных или выполнение любой другой предварительной обработки, необходимой для приведения данных в форму, подходящую для использования моделью МО;

- модель: это модель МО в форме, подходящей для использования в конвейере МО.

Примечание — Примером является алгоритм МО, реализованный в программном обеспечении как NF (сетевая функция) [4];

- политика: этот узел позволяет применять политики к результатам работы модели из соответствующего узла.

Примечание — Этот узел можно использовать, например для минимизации воздействия, когда результаты МО применяются в эксплуатируемых базисных сетях для МО. Оператор сети может ввести специальные правила для обеспечения работоспособности сети, например серьезные обновления могут производиться только в ночное время или при низком трафике данных в сети;

- распределитель: этот узел отвечает за идентификацию целевых стоков и распределение результатов из узла модели по соответствующим узлам целевых стоков.

Примечание — Данный узел может иметь возможность настраивать узлы целевых стоков. Например, протокол RRC может использоваться для настройки UE, действующего в качестве узла целевого стока;

- целевой сток: этот узел является целью вывода МО, над которым он выполняет действие.

Примечания

1 Например, UE, действующее в качестве узла целевого стока, может регулировать периодичность измерения канала на основе результатов МО.

2 «Точка выхода службы» узла — это точка, в которой можно получить доступ к службам, предоставляемым данным узлом (роль производителя). «Точка входа службы» узла — это точка, в которой могут быть использованы службы других узлов (роль потребителя).

8.1.2 Оркестратор функций машинного обучения

Как определено в разделе 3, MLFO (оркестратор функций машинного обучения) — это логический узел с функциями управления и оркестровки узлов в конвейерах МО на основе назначений МО и/или динамических условий сети.

Примечания

1 MLFO выбирает и повторно выбирает модель МО, основываясь, например, на ее производительности.

2 Размещение узлов конвейера МО, основанное на соответствующих сетевых возможностях и ограничениях приложения МО, является ответственностью MLFO.

3 MLFO также предоставляет функциональность соединения узлов в цепочку для формирования конвейера МО. Например, соединение в цепочку может использоваться для связи узла источника, реализованного в сети доступа, с узлом сборщика и узлом предобработки, реализованными в базовой сети. Сама цепочка декларируется в спецификации приложения МО, а ее конкретная технологическая реализация в сети осуществляется MLFO. MLFO определяет соединение узлов в цепочку с учетом ограничений (например, временных ограничений для прогнозирования).

4 На рисунке 4 показано назначение МО, чтобы подчеркнуть его важную роль в качестве входных данных для MLFO. Назначение МО — это декларативное описание, которое используется для спецификации приложения МО. Назначение МО не определяет какие-либо сетевые функции со специфическими технологиями, которые будут использоваться в приложении МО, а обеспечивает основу для отображения приложений МО в различные сетевые функции со специфическими технологиями. Назначение МО для определения приложений МО может использовать метаязык, специфичный для МО.

8.1.3 Полигон МО

Полигон МО — это изолированная среда, которая позволяет размещать отдельные конвейеры МО для их обучения, тестирования и оценки перед развертыванием в эксплуатируемой сети. Для обучения или тестирования полигон МО может использовать данные, сгенерированные моделируемыми и/или эксплуатируемыми базисными сетями для МО.

Примечание — Для обучения, тестирования и оценки моделей МО на полигоне МО могут пригодиться такие методы, как МО с учителем.

В дополнение к описанию вышеупомянутых высокоуровневых архитектурных компонентов следует обратить внимание на следующие вспомогательные архитектурные аспекты:

- SBA (сервис-ориентированная архитектура) [12], может использоваться для сопряжения функциональных возможностей МО и базисных сетей для МО. Аналогично для конвейера МО на полигоне данная архитектура может использоваться для сопряжения функциональных возможностей МО и моделируемых базисных сетей для МО. Это обеспечивает единую контрольную точку как для надстройки МО для NF (сетевые функции), так и для моделируемых базисных сетей для МО. SBA также может использоваться как между самими узлами конвейера МО, так и для управления функциональными возможностями МО с помощью MLFO;

- контрольные точки оперирования данными определяются между конвейером МО и моделируемыми или эксплуатируемыми базисными сетями для МО. Используя эти контрольные точки, любое

воздействие на базисные сети для МО локализуется в источнике данных и целевой конфигурации (в результате выполнения конвейера МО).

Примечания

1 Расширения существующих протоколов могут использоваться в контрольных точках оперирования данными для минимизации воздействия архитектуры на базисные сети для МО.

2 В случае, когда сетевые функции в базисных сетях для МО не совместимы с SBA, в контрольных точках оперирования данными также следует использовать протоколы, не поддерживающие требования SBA.

На рисунке 3 стрелками 2 и 3 показаны пути к конвейеру МО для данных, сгенерированных из моделируемых базисных сетей для МО и базисных сетей для МО, соответственно. Стрелки 1 и 4 показывают пути из конвейера МО для настройки целевого объекта на основе результатов МО в моделируемых базисных сетях для МО и базисных сетях для МО соответственно.

В приложении Б приведено соответствие архитектурных компонентов и вспомогательных архитектурных аспектов требованиям высокого уровня по разделу 7.

8.2 Высокоуровневая архитектура

Высокоуровневая архитектура, представленная на рисунке 4, основана на требованиях высокого уровня по разделу 7 и опирается на архитектурные компоненты, описанные в 8.1.

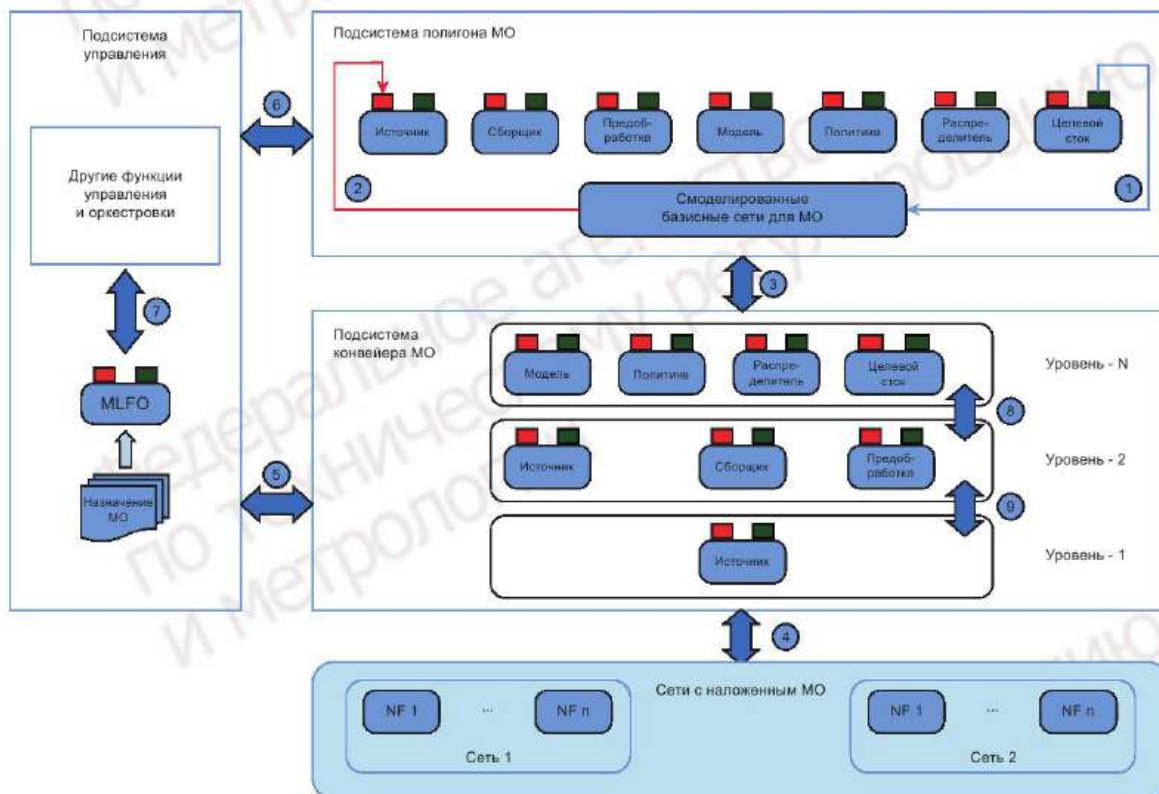


Рисунок 4 — Высокоуровневая архитектура

На рисунке 4 изображены следующие контрольные точки:

- 1, 2 — контрольные точки оперирования данными между моделируемыми базисными сетями для МО и конвейером МО в подсистеме полигона МО;
- 3 — контрольная точка между подсистемой полигона МО и подсистемой конвейера МО;
- 4 — контрольная точка между подсистемой конвейера МО и базисными сетями для МО.

Примечание — Данная контрольная точка обозначает контрольные точки оперирования данными, представленные на рисунке 3 в виде стрелок 3 и 4;

- 5, 6 — контрольные точки между подсистемой управления, подсистемой конвейера МО и подсистемой полигона МО соответственно;
- 7 — контрольная точка между MLFO и другими функциями управления и оркестровки подсистемы управления;
- 8, 9 — контрольные точки между узлами конвейера МО, расположенными на разных уровнях.

В следующих подразделах описаны три подсистемы высокоуровневой архитектуры, представленной на рисунке 4.

8.2.1 Подсистема управления

Данная подсистема включает в себя MLFO и другие функции управления и оркестровки, например, определенные в [6].

Подсистема управления позволяет расширить механизмы управления и оркестровки узлов конвейера МО, используемые для будущих телекоммуникационных сетей, включая IMT-2020. Это обеспечивает единый подход в управлении функциональными возможностями МО и сетевыми функциями.

MLFO работает в координации с другими функциями подсистемы управления с целью контроля узлов конвейера МО.

Примечания

1 Контрольные точки между функциями подсистемы управления и NF (сетевые функции) в базисных сетях для МО соответствуют контрольным точкам, определенным в соответствующих спецификациях, например, [6]. Данные контрольные точки выходят за рамки данного стандарта и потому не показаны на рисунке 4.

2 Взаимодействие между MLFO и другими функциями подсистемы управления может осуществляться с использованием SBA.

8.2.2 Подсистема конвейера МО

Конвейер МО представляет собой логический конвейер, который может быть настроен на существующие сетевые инфраструктуры. Службы MLFO используются для реализации и настройки узлов. При интеграции надстройки конвейера МО на базисные сети для МО, зависящие от конкретной технологии, может потребоваться расширение или определение специальных протоколов и API (интерфейсы программирования приложений).

Кроме того, необходимо отметить следующее:

а) SBA может использоваться для обеспечения связи как между сетевыми функциями (NF) и узлами конвейера МО, так и между самими узлами конвейера МО. Например, узел источника предоставляет доступ к интерфейсам для потребления данных от NF (сетевые функции) и производства данных для узла коллектора. Другой пример — узел с целевым стоком предоставляет доступ к интерфейсам для потребления результатов МО от узла с распределителем и производит конфигурации для NF, с которыми имеет связь.

Примечание — Вследствие гетерогенности сетевых функций и базисных сетей для МО SBA может не поддерживаться сетевыми функциями в базисных сетях для МО. В таких случаях между узлами конвейера МО и сетевыми функциями используются протоколы и API, специфичные для таких базисных сетей для МО;

б) размещение и построение цепочек узлов конвейера МО контролируется оркестратором MLFO; на осуществление этого контроля могут повлиять такие факторы, как:

- входные данные от назначения МО к MLFO, которые могут накладывать ограничения на размещение узлов конвейера МО.

Примечание — Примером ограничений по расположению узла модели МО является требование расположить узел модели на сетевом вычислительном ресурсе, предоставляющем определенный тип возможностей для ускорения;

- обратная связь, получаемая MLFO от функций управления и оркестровки, относящихся к базисным сетям для МО, или от узлов конвейера МО, может предоставить входные данные для размещения и построения цепочек узлов конвейера МО.

Примечание — Отвязка местоположения узлов конвейера МО от их функциональных возможностей, за исключением случаев ограничения производительности, достигается с помощью механизмов размещения и построения цепочек;

3) развертывание конвейера МО в будущих телекоммуникационных сетях, включая IMT-2020, может охватывать разные уровни. В этом случае для обеспечения многоуровневого распределения

конвейера МО используются контрольные точки между узлами, расположенные на разных уровнях конвейера МО.

Примечание — Как показано на рисунке 4, конвейер МО может быть распределен по нескольким уровням, например при конкретном развертывании конвейер может быть распределен между UE (пользовательское оборудование), AN (сеть доступа) и CN (базовая сеть). В зависимости от конкретного приложения МО возможны различные способы распределения узлов конвейера МО.

8.2.3 Подсистема полигона МО

Подсистема полигона МО позволяет конвейерам МО адаптироваться к динамическим сетевым средам, таким как будущие телекоммуникационные сети, включая IMT-2020, в которых возможно изменение ряда условий (например, условия радиointерфейса, расположение пользовательского оборудования, возможности и ресурсы сети). Подсистема полигона МО включает в себя конвейер(ы) МО и моделируемые базисные сети для МО и управляется MLFO в соответствии со спецификациями назначения МО. Подсистема полигона МО позволяет сетевым операторам изучать влияние результатов МО перед их развертыванием в эксплуатируемых базисных сетях для МО. Обратная связь от функций подсистемы управления поступает в подсистему полигона МО для того, чтобы конвейеры МО в подсистеме полигона МО могли адаптироваться к динамически меняющимся сетевым средам.

Необходимо отметить следующие моменты:

- контрольная точка между подсистемой конвейеров МО и подсистемой полигона МО позволяет конвейерам МО взаимодействовать с подсистемой полигона МО для обучения и обновления моделей МО;
- данные из базисных сетей для МО и/или моделируемых базисных сетей для МО могут использоваться для обучения моделей МО в подсистеме полигона МО;
- управление узлами конвейера МО в подсистеме полигона МО также контролируется MLFO. Это позволяет MLFO обучать и выбирать модель(и) МО для конкретного приложения МО.

8.3 Общие рекомендации по реализации высокоуровневой архитектуры

Общие рекомендации по реализации высокоуровневой архитектуры в различных базисных сетях для МО, зависящих от конкретных технологий, сводятся к следующему:

- реализация узлов конвейера МО: приложение МО описывается с помощью назначения МО. Поток информации в приложении МО может быть представлен с помощью цепочки узлов в конвейере МО. Данные из различных узлов с источниками, например, поступающие из различных базовых сетей для МО, должны быть собраны в узле со сборщиком и предварительно обработаны в узле с преобразованием до их передачи в узел с моделью МО. Затем результаты модели МО используются для применения в узле политики, которые будут реализованы в узле с целевым стоком. Приложение МО может быть создано путем реализации узлов конвейера МО с определенными ролями (например, источник, сборщик, целевой сток) и привязки этих узлов к сетевым функциям в базисной сети для МО, зависящей от конкретной технологии, исходя из соответствующих требований приложения МО и возможностей сетевых функций в базисной сети для МО. Реализация выполняется MLFO в координации с другими функциями управления и оркестровки;

- взаимодействие приложения МО с сетевыми функциями в базисной сети для МО: для приложения МО существует две точки специфического взаимодействия с сетевыми функциями в базисной сети для МО — узлы с источниками и целевыми стоками. Узлы с источниками могут иметь как основанный на SBA (сервис-ориентированная архитектура) интерфейс к сетевым функциям (производящим данные), так и не поддерживающий SBA интерфейс к сетевым функциям, зависящим от конкретной технологии. Подобным образом узлы с целевыми стоками могут иметь как основанный на SBA интерфейс к связанным с ним сетевым функциям (обеспечивающим исполнение выходных политик МО), так и не поддерживающий SBA интерфейс к сетевым функциям, зависящим от конкретной технологии;

- управление конвейером МО: осуществляется MLFO в координации с другими функциями управления и оркестровки;

- обучение и оценка моделей МО на полигоне МО: управляется MLFO независимо от базисных сетей для МО. Контрольная точка между подсистемой полигона МО и подсистемой конвейера МО используется для передачи обученных моделей МО, данных для обучения и обновлений моделей МО между подсистемой полигона МО и базисными сетями для МО, которые взаимодействуют с реальным конвейером МО.

В приложении А приведены примеры применения вышеуказанных рекомендаций при реализации высокоуровневой архитектуры в базисных сетях для МО, зависящих от конкретной технологии.

9 Требования обеспечения безопасности

Настоящий стандарт описывает структуру архитектуры системы МО, которую предполагается применять в будущих телекоммуникационных сетях, включая IMT-2020; в этой связи следует применять общие требования и механизмы сетевой безопасности в сетях на базе интернет-протокола IP [13], [14].

Требуется предотвращение несанкционированного доступа и утечки данных из конвейера МО с использованием механизмов аутентификации и авторизации, защиты от внешних атак и т. д., независимо от того, являются действия злоумышленными или нет.

Приложение А (справочное)

Реализация высокоуровневой архитектуры в базисных сетях для машинного обучения, зависящих от конкретной технологии

На рисунке А.1 приведен пример реализации высокоуровневой архитектуры в сети IMT-2020 (см. [4], [6]).

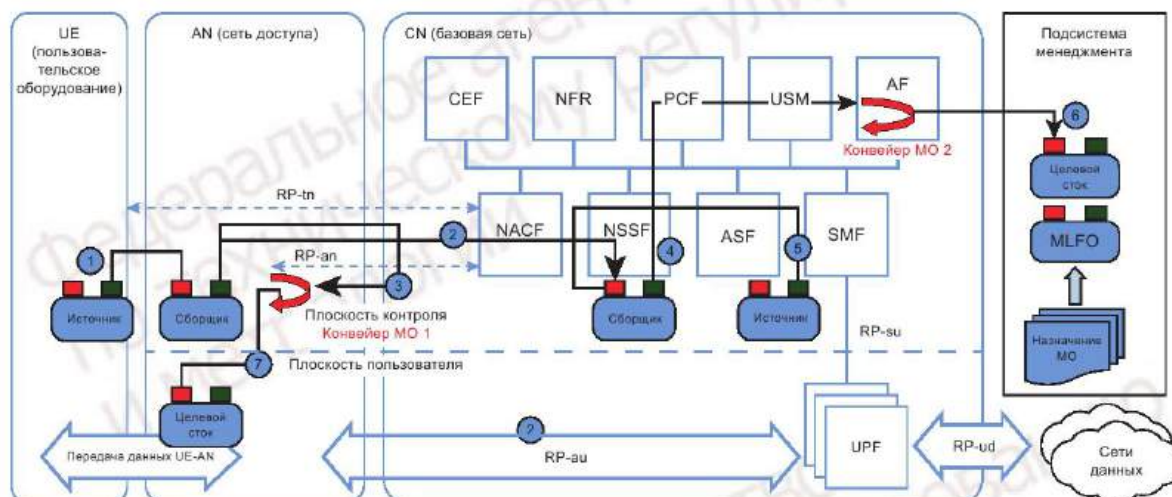


Рисунок А.1 — Пример реализации высокоуровневой архитектуры в сети IMT-2020

Данный пример реализации представлен следующим образом: конвейер МО демонстрирует расположение узлов конвейера МО в местах размещения узлов, например, в CN (базовая сеть), AN (сеть доступа), UE (пользовательское оборудование) или функциях управления. Например, конвейер МО, представленный стрелками 1→2→4→конвейер МО 2, использует входные данные от пользовательского оборудования для прогнозирования в CN (например, приложения МО на основе MPP (прогнозирование закономерностей мобильности)).

Примечание — Для простоты понимания подсистема полигона не показана на рисунке А.1, но ее функциональность применима и к рисунку А.1.

Описывают различные варианты реализации применительно к требованиям, указанным в разделе 7:

а) экземпляр реализации, поддерживающий требования 7.1 и 7.2:

1) стрелки 5 → 4 → конвейер МО 2 → 6 — данный конвейер МО использует входные данные от CN и, если необходимо, комбинацию входных данных от пользовательского оборудования, чтобы сделать прогнозы в CN и применить их к функциям управления. Такое применение результатов МО может, в свою очередь, повлиять на конфигурации различных уровней (например, на решения SON (самооптимизирующаяся сеть), принимаемые в CN, или на решения о распределении ресурсов, принимаемые в сети по замкнутому циклу);

2) стрелки 1 → 3 → конвейер МО 1 → 7 — данный конвейер МО использует входные данные от пользовательского оборудования и размещает модель МО в AN для принятия решений, чувствительных к задержке, которые применяются в самой AN;

б) экземпляр реализации, поддерживающий требования 7.3:

1) стрелка 1 — возможна реализация с помощью RRC (управление радиоресурсами);

2) стрелки 2, 3, 4, 5, 7 — возможна реализация в виде расширения контрольных точек в CN (см. [4]);

3) стрелка 6 — возможна реализация путем повторного использования контрольных точек, определенных в [6];

в) экземпляр реализации, поддерживающий требования 7.4.

Пользовательское оборудование является устройством с ограниченными ресурсами, поэтому на пользовательском оборудовании задается только источник. Данное ограничение указывается в назначении МО;

г) экземпляр реализации, поддерживающий требования 7.5:

MLFO размещает сборщики в AN и CN на основе спецификаций приложений МО в назначении МО. Для приложений, чувствительных к задержкам в AN, используется конвейер МО 1. Конвейер МО 2 используется для приложений, устойчивых к задержкам. Объединение в цепочку осуществляется в соответствии с требованиями, указанными в назначении МО.

**Приложение Б
(справочное)**

**Соотнесение архитектурных компонентов и поддерживающих аспектов
с требованиями раздела 8**

В таблице Б.1 приведено соотнесение архитектурных компонентов и вспомогательных аспектов с требованиями раздела 8.

Т а б л и ц а Б.1 — Соотнесение архитектурных компонентов и вспомогательных аспектов с требованиями

Архитектурные компоненты и вспомогательные аспекты	Требования	Объяснение соотнесения
Конвейер МО	См. 7.1 и 7.2	Конвейер МО предоставляет общую терминологию для МО в будущих телекоммуникационных сетях, включая IMT-2020. Определение узлов конвейера МО позволяет следить за эволюцией МО в отдельности от базисных сетей для МО. Задавание таких узлов является важной частью развертывания надстройки МО в различных базисных сетях для МО. Функции этих узлов определяются независимо от их расположения в сети, что делает возможным гибкое размещение таких узлов в сети. Сегментация, слияние и объединение в цепочку узлов конвейера МО позволяет разворачивать сложные функции из этих базовых узлов
MLFO	См. 7.2, 7.4 и 7.5	MLFO оркеструет узлы конвейера МО и управляет ими. Он также отвечает за оптимальное размещение в сети и объединение узлов конвейера МО в цепочку. MLFO реализует эти функции в координации с функциями управления и оркестровки. Декларативные спецификации приложений МО поддерживаются MLFO путем преобразования их в развертывания, специфичные для базисной сети для МО. Он также выбирает и перебирает модели, исходя из потребностей приложений МО и других ограничений, определенных в назначении МО
Полигон МО	См. 7.5	Данная подсистема создает возможность обучения, оценки и мониторинга производительности моделей МО перед их развертыванием в реальной сети. Она взаимодействует с базисными сетями для МО для передачи данных или обученных моделей. Кроме того, в ней используются моделируемые базисные сети для МО, которые могут генерировать данные, необходимые для обучения моделей МО
Назначение МО	См. 7.4	Данный вспомогательный аспект архитектуры, написанный на метаязыке, дает определение приложению МО, его потребностям с точки зрения ввода данных, вывода конфигурации, любых других ограничений или характеристик. Он используется в качестве входных данных для функциональных возможностей MLFO. Это исходные данные для MLFO для создания развертываемого, соответствующего требованиям конвейера МО в конкретных базисных сетях для МО. Стандартизация данного архитектурного аспекта позволяет сторонним поставщикам решений интегрироваться с конвейером МО

Окончание таблицы Б.1

Архитектурные компоненты и вспомогательные аспекты	Требования	Объяснение соотнесения
SBA	См. 7.3	<p>Данный вспомогательный аспект архитектуры используется для взаимодействия между сетевыми функциями базисной сети для МО и конвейером МО. Он обеспечивает необходимое слабосвязанное взаимодействие между надстройкой МО и базисными сетями для МО. SBA также используется для взаимодействия между моделируемыми базисными сетями для МО и конвейером МО. Она используется для обучения конвейера МО и настройки моделируемых сетей МО с помощью конвейера МО</p>

Библиография

- | | | |
|------|-----------------|--|
| [1] | ITU-T Y.3100 | Термины и определения для сетей IMT-2000 |
| [2] | ITU-T Y.3650 | Общие принципы сетевого взаимодействия на основе больших данных |
| [3] | ITU-T Y.3170 | Требования к системам машинного обучения для обеспечения качества обслуживания в сети IMT-2020 |
| [4] | ITU-T Y.3104 | Архитектура сети IMT-2020 |
| [5] | 3GPP TS 28.554 | Управление и оркестровка сетей 5G; ключевые показатели эффективности 5G (KPI) (Выпуск 15) |
| [6] | ITU-T Y.3111 | Общие принципы управления сетью и оркестровки IMT-2020, |
| [7] | ITU-T Y.3102 | Общие принципы сети IMT-2020 |
| [8] | ITU-T Y.3110 | Требования к управлению сетью и оркестровке IMT-2020 |
| [9] | ETSI MEC 003 | Вычисления на мобильных периферийных устройствах (MEC); общие принципы и эталонная архитектура |
| [10] | 3GPP TS 23.501 | Системная архитектура для системы 5G (Выпуск 15) |
| [11] | GSMA IR.92 | Профиль IMS для голосовой связи и SMS |
| [12] | ETSI TS 129 500 | Система 5G; техническая реализация архитектуры на основе обслуживания; этап 3 (3GPP TS 29.500, версия 15.0.0, выпуск 15) |
| [13] | ITU-T Y.2701 | Требования к безопасности для сетей последующих поколений версии 1 |
| [14] | ITU-T Y.3101 | Требования сети IMT-2020 |

УДК 004.8:006.354

ОКС 35.020

Ключевые слова: искусственный интеллект, структура архитектуры, ИИТ-2020, машинное обучение

Федеральное агентство
по техническому регулированию
и метрологии

Федеральное агентство
по техническому регулированию
и метрологии

Федеральное агентство
по техническому регулированию
и метрологии

Федеральное агентство
по техническому регулированию
и метрологии

Федеральное агентство
по техническому регулированию
и метрологии

Технический редактор *И.Е. Черепкова*
Корректор *О.В. Лазарева*
Компьютерная верстка *М.В. Малеевой*

Сдано в набор 01.11.2024. Подписано в печать 19.11.2024. Формат 60×84%. Гарнитура Ариал.
Усл. печ. л. 3,26. Уч.-изд. л. 2,70.

Подготовлено на основе электронной версии, предоставленной разработчиком стандарта

Создано в единичном исполнении в ФГБУ «Институт стандартизации»
для комплектования Федерального информационного фонда стандартов,
117418 Москва, Нахимовский пр-т, д. 31, к. 2.
www.gostinfo.ru info@gostinfo.ru