ITUПубликации
Рекомендации

Международный союз электросвязи
Сектор стандартизации

Рекомендация

МСЭ-Т M.3387 (03/2024)

СЕРИЯ M: Управление электросвязью, включая СУЭ и техническое обслуживание сетей

Сеть управления электросвязью

Требования к управлению системами федеративного машинного обучения

РЕКОМЕНДАЦИИ МСЭ-Т СЕРИИ M

Управление электросвязью, включая СУЭ и техническое обслуживание сетей

Введение и общие принципы технического обслуживания и организации технического обслуживания	M.10–M.299
Международные системы передачи	M.300–M.559
Международные телефонные каналы	M.560–M.759
Системы сигнализации по общему каналу	M.760–M.799
Международные системы телеграфной и фототелеграфной передачи	M.800–M.899
Международные арендованные первичные и вторичные групповые тракты	M.900–M.999
Международные арендованные каналы	M.1000–M.1099
Системы и службы подвижной электросвязи	M.1100–M.1199
Международная телефонная сеть общего пользования	M.1200–M.1299
Международные системы передачи данных	M.1300–M.1399
Обозначения и обмен информацией	M.1400–M.1999
Международная сеть транспортировки сообщений	M.2000–M.2999
Сеть управления электросвязью	**M.3000–M.3599**
Цифровые сети с интеграцией служб	M.3600–M.3999
Системы сигнализации по общему каналу	M.4000–M.4999

Для получения более подробной информации просьба обращаться к Перечню Рекомендаций МСЭ-Т.

Рекомендация МСЭ-Т M.3387

Требования к управлению системами федеративного машинного обучения

Резюме

Рекомендация МСЭ-Т M.3387 применима при проектировании архитектуры, исследовании и разработке моделей федеративного машинного обучения (FMLM). В сфере больших данных и искусственного интеллекта (ИИ) встают трудные задачи по обеспечению конфиденциальности данных и информационной безопасности в связи со все возрастающим регуляторным давлением. Многие повседневные операции в системах и приложениях, работающих с большими данными, например слияние пользовательских данных из различных источников для построения модели машинного обучения, считаются незаконными в рамках действующей нормативно-правовой базы.

Федеративное машинное обучение (FML) призвано служить жизнеспособным решением для обеспечения распределенной работы с данными в приложениях машинного обучения. В рамках FML владельцы данных не обмениваются исходными данными напрямую и не позволяют ни одной из сторон восстанавливать путем вероятностного вывода личную информацию других сторон. В целях содействия построению и использованию FMLM и повышения качества услуг FML в Рекомендации МСЭ-Т M.3387 устанавливаются требования к управлению системами федеративного машинного обучения (FMLS), включая функциональную архитектуру FMLS, а также требования в области базового управления, управления моделями и управления данными.

Хронологическая справка[*]

Издание	Рекомендация	Утверждение	Исследовательская комиссия	Уникальный идентификатор
1.0	МСЭ-Т M.3387	11.03.2024 г.	2-я	11.1002/1000/15786

Ключевые слова

Услуга федеративного машинного обучения, система федеративного машинного обучения, требования к управлению.

[*] Для получения доступа к Рекомендации наберите в адресном поле вашего браузера URL http://handle.itu.int/, после которого укажите уникальный идентификатор Рекомендации.

ПРЕДИСЛОВИЕ

Международный союз электросвязи (МСЭ) является специализированным учреждением Организации Объединенных Наций в области электросвязи и информационно-коммуникационных технологий (ИКТ). Сектор стандартизации электросвязи МСЭ (МСЭ-Т) – постоянный орган МСЭ. МСЭ-Т отвечает за изучение технических, эксплуатационных и тарифных вопросов и за выпуск Рекомендаций по ним с целью стандартизации электросвязи на всемирной основе.

На Всемирной ассамблее по стандартизации электросвязи (ВАСЭ), которая проводится каждые четыре года, определяются темы для изучения исследовательскими комиссиями МСЭ-Т, которые, в свою очередь, вырабатывают Рекомендации по этим темам.

Утверждение Рекомендаций МСЭ-Т осуществляется в соответствии с процедурой, изложенной в Резолюции 1 ВАСЭ.

В некоторых областях информационных технологий, которые входят в компетенцию МСЭ-Т, необходимые стандарты разрабатываются на основе сотрудничества с ИСО и МЭК.

ПРИМЕЧАНИЕ

В настоящей Рекомендации термин "администрация" используется для краткости и обозначает как администрацию электросвязи, так и признанную эксплуатационную организацию.

Соблюдение положений данной Рекомендации осуществляется на добровольной основе. Однако данная Рекомендация может содержать некоторые обязательные положения (например, для обеспечения функциональной совместимости или возможности применения), и в таком случае соблюдение Рекомендации достигается при выполнении всех указанных положений. Для выражения требований используются слова "следует", "должен" ("shall") или некоторые другие обязывающие выражения, такие как "обязан" ("must"), а также их отрицательные формы. Употребление таких слов не означает, что от какой-либо стороны требуется соблюдение положений данной Рекомендации.

ПРАВА ИНТЕЛЛЕКТУАЛЬНОЙ СОБСТВЕННОСТИ

МСЭ обращает внимание на вероятность того, что практическое применение или выполнение настоящей Рекомендации может включать использование заявленного права интеллектуальной собственности. МСЭ не занимает какую бы то ни было позицию относительно подтверждения, действительности или применимости заявленных прав интеллектуальной собственности, независимо от того, доказываются ли такие права членами МСЭ или другими сторонами, не относящимися к процессу разработки Рекомендации.

На момент утверждения настоящей Рекомендации МСЭ не получил извещения об интеллектуальной собственности, защищенной патентами, которые могут потребоваться для выполнения настоящей Рекомендации. Однако те, кто будет применять Рекомендацию, должны иметь в виду, что вышесказанное может не отражать самую последнюю информацию, и поэтому им настоятельно рекомендуется обращаться к патентной базе данных БСЭ по адресу: http://www.itu.int/ITU-T/ipr/.

СОДЕРЖАНИЕ

Рекомендация МСЭ-Т M.3387

Требования к управлению системами федеративного машинного обучения

1 Сфера применения

В настоящей Рекомендации устанавливаются требования к управлению системами федеративного машинного обучения (FMLS). В сферу применения данной Рекомендации входят следующие аспекты:

– общая функциональная архитектура FMLS;

– требования в области базового управления, к которой относятся настройка свойств системы, разрешений ее узлов, управление запросами на получение услуги и т. д.;

– требования в области управления моделями, к которой относятся развертывание исходной модели, управление алгоритмами обучения, управление механизмами агрегирования и т. д.;

– требования в области управления данными, к которой относятся защищенное хранение, извлечение данных, передача ресурсов данных и т. д.;

– сценарий использования.

2 Справочные документы

Указанные ниже Рекомендации МСЭ-Т и другие справочные документы содержат положения, которые путем ссылок на них в данном тексте составляют положения настоящей Рекомендации. На момент публикации указанные издания были действующими. Все Рекомендации и другие источники могут подвергаться пересмотру; поэтому пользователям данной Рекомендации предлагается изучить возможность применения последнего издания Рекомендаций и других справочных документов, перечисленных ниже. Перечень действующих на настоящий момент Рекомендаций МСЭ-Т регулярно публикуется. Ссылка на документ, приведенный в настоящей Рекомендации, не придает ему как отдельному документу статус Рекомендации.

[IEEE 3652.1] IEEE standard 3652.1 (2020), *IEEE Guide for Architectural Framework and Application of Federated Machine Learning.*

3 Определения

3.1 Термины, определенные в других документах

В настоящей Рекомендации используются следующие термины, определенные в других документах.

3.1.1 шифрование (encryption) [b-ITU-T X.1367]: Криптографическое преобразование данных для получения зашифрованного текста.

3.1.2 механизм внутреннего вознаграждения (intrinsic incentive mechanism) [b-ITU-T Y.4205]: Механизм, обеспечивающий выработку внутреннего вознаграждения за участие в той или иной деятельности или вклад в нее, например ощущения самореализации, радости или вклада в достижение высшей цели.

3.2 Термины, определенные в настоящей Рекомендации

В настоящей Рекомендации определены следующие термины.

3.2.1 аудитор (auditor): Узел, отвечающий за контроль эффективности процесса федеративного машинного обучения для обеспечения его соответствия регуляторным требованиям.

3.2.2 координатор (coordinator): Узел, который отвечает за построение моделей федеративного машинного обучения среди разных владельцев данных и предоставляет такие модели клиентам федеративного машинного обучения.

3.2.3 владелец данных (data owner): Узел, который обладает правом собственности на набор данных, используемый в рамках федеративного машинного обучения, и выполняет обучение локальной модели, обеспечивая при этом конфиденциальность данных.

3.2.4 качество данных (data quality): Показатель, используемый для оценки достоверности и полезности набора данных.

3.2.5 набор данных (data set): Набор точек или примеров данных, которые используются для целей обучения, тестирования или оценки. Каждая точка данных в наборе данных представляет собой образец, включающий идентификатор данных, признаки данных (включают имена и значения) или метку класса (в случае контролируемого обучения).

3.2.6 федеративное машинное обучение (federated machine learning (FML)): Система машинного обучения, которая облегчает совместное построение моделей машинного обучения на нескольких распределенных обучающих узлах без раскрытия конфиденциальных данных, принадлежащих владельцам данных.

3.2.7 модель федеративного машинного обучения (federated machine learning model (FMLM)): Результат процесса обучения системы федеративного машинного обучения. Обученная модель используется для решения задач логического вывода информации по новым данным.

3.2.8 система управления федеративным машинным обучением (federated machine learning management system (FMLMS)): Система, обеспечивающая управление ресурсами узлов и услугами обучения моделей в рамках систем федеративного машинного обучения.

3.2.9 услуга федеративного машинного обучения (federated machine learning service): Услуга обучения моделей искусственного интеллекта, использующая метод федеративного машинного обучения и предоставляющая на выходе глобально обученную модель.

3.2.10 клиент федеративного машинного обучения (federated machine learning service client (FMLSC)): Прикладной логический объект, который инициирует запрос на получение услуги федеративного машинного обучения, а в ответ получает обученные модели федеративного машинного обучения.

3.2.11 система федеративного машинного обучения (federated machine learning system (FMLS)): Система с множеством узлов обучения, которые совместно создают и используют модели машинного обучения, не раскрывая принадлежащие участникам исходные и конфиденциальные данные.

3.2.12 исходные данные (raw data): Совокупность наборов данных, которые собирают, хранят и поддерживают владельцы данных. Исходные данные содержат конфиденциальную информацию пользователей и владельцев данных.

3.2.13 обучение (training): Процесс федеративного машинного обучения, в том числе локального обучения исходных данных и агрегирование промежуточных обновленных параметров для оптимизации производительности моделей федеративного машинного обучения.

4 Сокращения и акронимы

В настоящей Рекомендации используются следующие сокращения и акронимы.

AI	Artificial Intelligence	ИИ	Искусственный интеллект
AUC	Area Under the Curve		Площадь под кривой
CAP	Computing Access Point		Точка доступа к вычислительным ресурсам
CNN	Convolutional Neural Network		Сверточная нейронная сеть
CPU	Central Processing Unit	ЦП	Центральный процессор
FML	Federated Machine Learning		Федеративное машинное обучение
FMLM	Federated Machine Learning Model		Модель федеративного машинного обучения
FMLMS	Federated Machine Learning Management System		Система управления федеративным машинным обучением
FMLS	Federated Machine Learning System		Система федеративного машинного обучения

FMLSC	Federated Machine Learning Service Client	Клиент федеративного машинного обучения
GPU	Graphics Processing Unit	Графический процессор
ID	Identification	Идентификация
IoV	Internet of Vehicles	Интернет транспортных средств
MEC	Mobile Edge Computing	Мобильные периферийные вычисления
MSE	Mean Squared Error	Среднеквадратическая ошибка

5 Соглашения

В настоящей Рекомендации:

– Ключевое слово "**требуется**" означает требование, которому необходимо неукоснительно следовать и отклонение от которого не допускается, если будет сделано заявление о соответствии данному документу.

– Ключевое слово "**рекомендуется**" означает требование, которое рекомендуется, но не является абсолютно необходимым. Таким образом, для заявления о соответствии настоящему документу данное требование не является обязательным.

– Ключевые слова "**может факультативно**" означают необязательное требование, которое допустимо, но не имеет рекомендательного значения. Эти термины не означают, что вариант реализации поставщика должен обеспечивать выполнение соответствующей функции, активируемой по желанию оператора сети/поставщика услуг. Это означает лишь, что поставщик может факультативно предоставить данную функцию и по-прежнему заявлять о соответствии спецификации.

6 Обзор

Как отмечается в документе [IEEE 3652.1] *IEEE Guide for Architectural Framework and Application of Federated Machine Learning*, в сфере больших данных и искусственного интеллекта (ИИ) встают трудные задачи по обеспечению конфиденциальности данных и информационной безопасности в связи со все возрастающим регуляторным давлением. Многие повседневные операции в системах и приложениях, работающих с большими данными, например слияние пользовательских данных из различных источников для построения модели машинного обучения, считаются незаконными в рамках действующей нормативно-правовой базы. Федеративное машинное обучение (FML) призвано служить жизнеспособным решением для обеспечения распределенной работы с данными в приложениях машинного обучения. В рамках FML владельцы данных не обмениваются исходными данными напрямую и не позволяют ни одной из сторон восстанавливать путем вероятностного вывода личную информацию других сторон.

Международный стандарт [IEEE 3652.1] устанавливает основы архитектуры FML, способствующие совместной работе с участием многих сторон и облегчающие ее. Вместе с тем в различных бизнес-сценариях к услугам обучения моделей предъявляются разные требования. Поэтому система федеративного машинного обучения (FMLS) должна координировать работу обучающих узлов FML посредством управления свойствами системы, модели и данных для стабильного и защищенного предоставления услуги FML.

FMLS – это система, которая предоставляет услугу FML клиентам федеративного машинного обучения (FMLSC). Для обеспечения безопасного и эффективного предоставления услуги FML следует определить функции настройки свойств системы, настройки обучающего узла и управления запросами на получение услуги, а также другие функции управления. Эти функции предоставляет система управления федеративным машинным обучением (FMLMS).

Исходя из принципов работы FMLS, изложенных в [IEEE 3652.1], настоящая Рекомендация устанавливает требования к управлению FMLS, в том числе требования в области базового управления, управления моделями и управления данными.

7 Сценарий управления системой федеративного машинного обучения

Сценарий управления системой федеративного машинного обучения показан на рисунке 1. В FMLS обучающие узлы FML рассматриваются как имеющие различные роли, соответствующие функциям, которые они выполняют в рамках данной задачи FML, включая роли координатора, аудитора и владельца данных.

Координатор отвечает за координацию выполнения задачи FML в составе FMLS и вывод обученной модели федеративного машинного обучения (FMLM). Аудитор отвечает за контроль за всем процессом FML для обеспечения надежности и безопасности данных. Владелец данных отвечает за локальное обучение и обновление моделей. Более подробно функции этих ролей изложены в [IEEE 3652.1].

Интерфейсы, относящиеся к управлению FMLMS, показаны на рисунке 1. В основном это два интерфейса:

– **интерфейс I1**, который представляет собой интерфейс, расположенный между FMLMS и FMLSC и используемый для передачи требований к услуге федеративного машинного обучения от FMLSC и возвращения в FMLSC обученной FMLM;

– **интерфейс I2**, который представляет собой интерфейс, расположенный между FMLMS и FMLS и используемый для управления ресурсами узлов и задачами обучения.

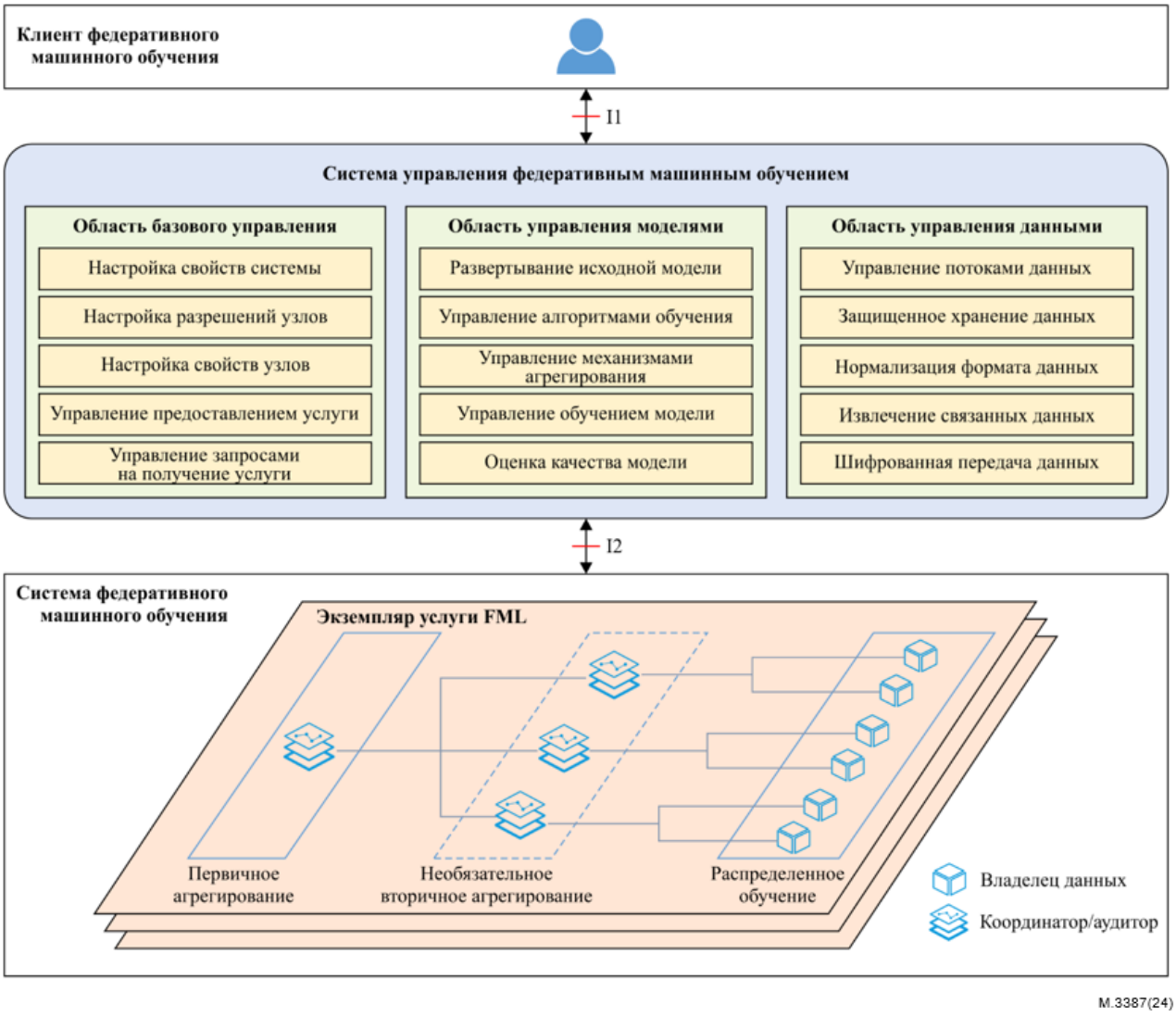

Рисунок 1 – Сценарий управления системой федеративного машинного обучения

Область базового управления, к которой относятся настройка свойств системы, настройка разрешений узлов, управление предоставлением услуги и управление запросами на получение услуги.

– **Настройка свойств системы** – инициализация и изменение свойств FMLS.

– **Настройка разрешений узлов** – управление разрешениями обучающих узлов FML в соответствии с правилами безопасности, установленными FMLMS.

– **Настройка свойств узлов** – инициализация и изменение свойств обучающих узлов FML.

– **Управление предоставлением услуги** – управление топологией предоставления услуги FML и оценка качества такой услуги.

– **Управление запросами на получение услуги** – классификация запросов на получение услуги FML, их обработка и ответ на них.

Область управления моделью, к которой относятся развертывание исходной модели, управление алгоритмами обучения, управление механизмами агрегирования, управление обучением модели и оценка качества модели.

– **Развертывание исходной модели** – развертывание исходной FMLM в координаторе в соответствии с требованиями к услуге FML.

– **Управление алгоритмами обучения** – выбор подходящих алгоритмов машинного обучения исходя из требований к услуге FML.

– **Управление механизмами агрегирования** – выбор или разработка надлежащих стратегий агрегирования исходя из требований к услуге FML, параметров ресурсного обеспечения и признаков данных.

– **Управление обучением модели** – управление и контроль за процессом обучения модели, включая передачу и обновление FMLM.

– **Оценка качества модели** – оценка качества FMLM по соответствующим показателям.

Область управления данными, к которой относятся управление потоками данных, защищенное хранение данных, нормализация формата данных, извлечение связанных данных и шифрованная передача данных.

– **Управление потоками данных** – управление потоками метаданных к исходным данным, а также потоками данных FMLM.

– **Защищенное хранение данных** – хранение метаданных к исходным данным с использованием различных методов шифрования.

– **Нормализация формата данных** – стандартизация формата метаданных к исходным данным, например табличной формы.

– **Извлечение связанных данных** – извлечение данных, связанных с задачей FML, в виде наборов данных для обучения FMLM.

– **Шифрованная передача данных** – выбор алгоритмов для шифрования передаваемых данных и каналов связи.

8 Требования в области базового управления

8.1 Требования к настройке свойств системы

Требуется, чтобы система управления федеративным машинным обучением (FMLMS) осуществляла настройку свойств FMLS, в том числе свойств задач и ресурсов, для поддержки функциональной реализации услуг FML.

8.1.1 Свойства задач

Требуется, чтобы FMLMS осуществляла настройку свойств задач в соответствии с запросами на получение услуги FML, поступившими от FMLSC.

– Тип задачи – категория задачи FML, например классификация изображений или генерация текста.

– Приоритет задачи – уровень приоритета задачи FML, например высокий, средний или низкий.

8.1.2 Свойства ресурсов

Требуется, чтобы FMLMS осуществляла настройку параметров ресурсного обеспечения задачи FML, включая вычислительные ресурсы, ресурсы связи и ресурсы хранения.

– Вычислительные ресурсы – полный объем вычислительных ресурсов, требуемых для выполнения задач FML, например число центральных процессоров (ЦП) и графических процессоров (ГП).

– Ресурсы связи – полный объем ресурсов связи, требуемых для выполнения задач FML, например мощность и полоса пропускания передатчика.

– Ресурсы хранения – полный объем ресурсов хранения, требуемых для выполнения задач FML, например объем свободного дискового пространства и выделенного хранилища.

8.2 Требования к настройке разрешений узлов

Требуется, чтобы FMLMS осуществляла выбор надлежащих обучающих узлов FML с достаточным объемом ресурсов для выполнения данной конкретной задачи FML, а затем выдавала разрешения на доступ и управляла ими исходя из уровня защищенности и надежности узлов.

8.3 Требования к настройке свойств узлов

Требуется, чтобы FMLMS поддерживала настройку свойств обучающих узлов FML, включая свойства ролей, вычислительных ресурсов, ресурсов связи и ресурсов хранения.

8.3.1 Свойства ролей

Требуется, чтобы FMLMS осуществляла настройку ролей обучающих узлов FML, к которым относятся роли координатора, аудитора и владельца данных. Конкретные функции этих трех ролей изложены в [IEEE 3652.1].

8.3.2 Свойства вычислительных ресурсов

Требуется, чтобы FMLMS обеспечивала поддержку обучающих узлов FML в части настройки их свойств, относящихся к вычислительным ресурсам, например числа ЦП и графических процессоров.

8.3.3 Свойства ресурсов связи

Требуется, чтобы FMLMS обеспечивала поддержку обучающих узлов FML в части настройки их свойств, относящихся к ресурсам связи, например мощности и полосы пропускания передатчика.

8.3.4 Свойства ресурсов хранения

Требуется, чтобы FMLMS обеспечивала поддержку обучающих узлов FML в части настройки их свойств, относящихся к хранению данных, например объема свободного дискового пространства и выделенного хранилища.

8.4 Требования к управлению предоставлением услуги

Требуется, чтобы FMLMS осуществляла управление качеством предоставления услуги FML, включая управление топологией предоставления услуги и оценку качества услуги.

8.4.1 Управление топологией предоставления услуги

Требуется, чтобы FMLMS осуществляла управление топологией предоставления услуги FML, включая генерацию и реконструкцию топологии.

– Генерация топологии предоставления услуги – генерация топологии предоставления услуги FML для данной конкретной задачи обучения модели FML, включая роли и связи между обучающими узлами FML, и передача информации о топологии на все обучающие узлы FML.

– Реконструкция топологии предоставления услуги – реконструкция топологии предоставления услуги FML в случае отказа узла предоставления услуги, исчерпания ресурсов и т. д.

8.4.2 Оценка качества услуги

Рекомендуется, чтобы FMLMS осуществляла оценку качества услуги FML, включая оценку качества работы сети и системы служебных стимулов.

– Оценка качества работы сети – оценка показателей качества работы системы федеративного машинного обучения (FMLS), таких как потребление ресурсов и временна́я задержка.

– Оценка системы служебных стимулов – комплексная оценка вклада обучающих узлов FML в предоставляемую услугу FML и настройка механизма внутреннего вознаграждения по размеру вклада, мотивирующего каждый узел обучения на участие в предоставлении услуги. Комплексная оценка вклада включает оценку потребления ресурсов, вклада в повышение качества модели и других параметров.

8.5 Требования к управлению запросами на получение услуги

Требуется, чтобы FMLMS осуществляла классификацию запросов на получение услуги и планирование их выполнения в условиях одновременного существования нескольких запросов.

– Классификация запросов на получение услуги – классификация запросов на получение услуги по важности, приоритету, требованиям к задержке и т. д.

– Планирование выполнения запросов – обработка запросов на получение услуги с планированием их выполнения тем или иным методом, например в соответствии с приоритетом.

9 Требования в области управления моделью

9.1 Требования к развертыванию исходной модели

Требуется, чтобы FMLMS поддерживала доставку FMLM и запуск процесса их обучения в соответствии с запросами на получение услуги FML, поступившими от FMLSC.

9.2 Требования к управлению алгоритмами обучения

Требуется, чтобы FMLMS осуществляла выбор алгоритмов машинного обучения и настройку их параметров в соответствии с требованиями к услуге FML.

– Выбор алгоритмов – выбор подходящих алгоритмов машинного обучения, таких как нейронные сети или дерево решений.

– Настройка параметров – настройка параметров и гиперпараметров алгоритма машинного обучения, таких как скорость обучения, размер батча и коэффициент регуляризации.

9.3 Требования к управлению механизмами агрегирования

Требуется, чтобы FMLMS осуществляла выбор механизмов агрегирования и настройку их параметров в соответствии с требованиями к услуге FML, параметрами ресурсного обеспечения и признаками данных.

– Выбор механизма – выбор или разработка надлежащего механизма агрегирования (например, синхронный, асинхронный, полусинхронный).

– Настройка параметров – настройка параметров механизма агрегирования (например, число раундов агрегирования, число кластеров, весовые коэффициенты (веса) в агрегировании модели).

9.4 Требования к управлению обучением модели

Требуется, чтобы FMLS обеспечивала поддержку владельцев данных в части проведения локального обучения FMLM и отправки обученной модели координатору.

Требуется, чтобы FMLS обеспечивала поддержку координатора в части управления процессом обучения FMLM, включая широковещательную рассылку, агрегирование и обновление.

Требуется, чтобы FMLS обеспечивала поддержку аудитора в части контроля за процессом обучения FMLM исходя из нормативных правил, например определяла надежность обучающего узла FML и измеряла вклад обучающих узлов FML.

9.5 Требования в области оценки качества модели

Требуется, чтобы FMLMS осуществляла оценку качества FMLM по соответствующим показателям качества.

ПРИМЕЧАНИЕ. – Показатели качества модели включают в себя точность, полноту и площадь под кривой (AUC) для классификационных моделей и среднеквадратическую ошибку (MSE) для регрессионных моделей [b-ITU-T Y.3179].

Рекомендуется, чтобы FMLMS корректировала процесс обучения по результатам оценки модели в целях повышения качества модели.

10 Требования в области управления данными

10.1 Требования к управлению потоками данных

Требуется, чтобы FMLS поддерживала генерацию, хранение, передачу и обновление метаданных к исходным данным.

ПРИМЕЧАНИЕ. – Факультативно владельцы данных могут обмениваться метаданными к исходным данным (например, идентификаторами данных (ID), характеристиками данных) в процессе FML.

Требуется, чтобы FMLS поддерживала передачу, агрегирование и обновление данных FMLM.

Требуется, чтобы FMLS поддерживала управление безопасностью данных модели в целях предотвращения утечки персональных данных во внешние и вредоносные узлы в процессе FML.

10.2 Требования в области защищенного хранения данных

Требуется, чтобы FMLMS поддерживала защищенное хранение метаданных к исходным данным и данных модели.

10.3 Требования к нормализации формата данных

Требуется, чтобы FMLMS собирала метаданные к исходным данным, представляющие их признаки, и предоставляла универсальный стандартный формат базы данных.

ПРИМЕЧАНИЕ. – Согласно [IEEE 3652.1] исходные данные FML обычно хранятся в стандартном формате базы данных, где каждая строка представляет образец данных, а каждый столбец – признак или метку этого образца. Набор атрибутов признака обычно представляется в виде собственных векторов (X_1, X_2, ..., X_n). В контролируемом обучении полный набор данных обучения состоит из признаков, представленных X, и меток, представленных Y.

10.4 Требования к извлечению связанных данных

Требуется, чтобы FMLS обеспечивала поддержку владельцев данных в части извлечения исходных данных, связанных с обучением модели, в виде наборов данных. В FMLS множество наборов данных пересекаются по идентификаторам образцов и атрибутам признаков. В зависимости от степени пересечения идентификаторов образцов или признаков выделяют следующие три случая.

– Горизонтальное FML – построение модели, в которой наборы данных существенно пересекаются в пространстве признаков, но не в пространстве идентификаторов. Координатор отвечает за установление соответствия между признаками у разных владельцев данных.

– Вертикальное FML – построение модели, в которой наборы данных существенно пересекаются в пространстве образцов, но не в пространстве признаков. Координатор отвечает за установление соответствия между образцами у разных владельцев данных.

– Федеративное трансферное обучение – построение модели, в которой наборы данных не имеют существенного пересечения ни в пространстве образцов, ни в пространстве признаков. Координатор отвечает за применение многократно используемых знаний в отношении различных доменов признаков.

10.5 Требования в области шифрованной передачи данных

Требуется, чтобы FMLS поддерживала технологии защиты конфиденциальности данных, такие как защищенные многосторонние вычисления, гомоморфное шифрование и дифференциальная конфиденциальность, для исключения вероятностного вывода исходных данных по данным модели другими обучающими узлами FML.

Рекомендуется, чтобы FMLS поддерживала технологии шифрования канала передачи для обеспечения безопасной среды передачи данных.

Дополнение I

Пример сценария использования FMLMS для управления обучением FMLM для обнаружения дорожных аномалий в интернете транспортных средств

(Данное Дополнение не является неотъемлемой частью настоящей Рекомендации.)

В настоящем Дополнении представлен типовой пример приложения и услуги с использованием системы управления федеративным машинным обучением (FMLMS) для управления системой федеративного машинного обучения (FMLS) в интернете транспортных средств (IoV). В нем также описаны функции FMLMS для предоставления услуги FML, рассматриваемые в настоящей Рекомендации.

I.1 Введение

Новые приложения на базе интеллектуальных устройств, таких как умные транспортные средства, предъявляют строгие требования к задержке и конфиденциальности. Это делает облачные вычисления непригодными для применения в данных сценариях, открывая дорогу для использования FML на базе мобильных периферийных вычислений (MEC). Суть FML на базе MEC состоит в распределенном обучении моделей на мобильных устройствах с ограниченными вычислительными ресурсами, емкостью хранилища, энергетическим ресурсом и пропускной способностью канала передачи; исходные данные при этом хранятся локально. Параметры модели доставляются на близлежащие точки доступа к вычислительным ресурсам (CAP) для дальнейшего агрегирования. В IoV обмен данными между транспортными средствами для совместного анализа улучшает впечатление от вождения и повысить качество обслуживания. Таким образом, путем разработки совместной вычислительной архитектуры FML на базе MEC содействует обнаружению дорожных аномалий с одновременной защитой конфиденциальности данных.

I.2 Совместная облачная, периферийная и терминальная архитектура FML

В IoV FMLS использует FML для обучения глобальной FMLM, которая применяется для решения задач обнаружения дорожных аномалий. В число компонентов FMLS входят владельцы данных, развернутые на умных транспортных средствах, координаторы, развернутые в точках доступа к вычислительным ресурсам (CAP), и агентский сервер FML (AS на рисунке I.1), развернутый на облачном сервере (CS на рисунке I.1). FMLMS развертывается в облаке и управляет процессом обучения FML. При этом применяются технологии MEC для обеспечения качества услуги FML посредством планирования задач и регионального агрегирования модели. Сценарий обучения FML в рамках совместной облачной, периферийной и терминальной архитектуры показан на рисунке I.1.

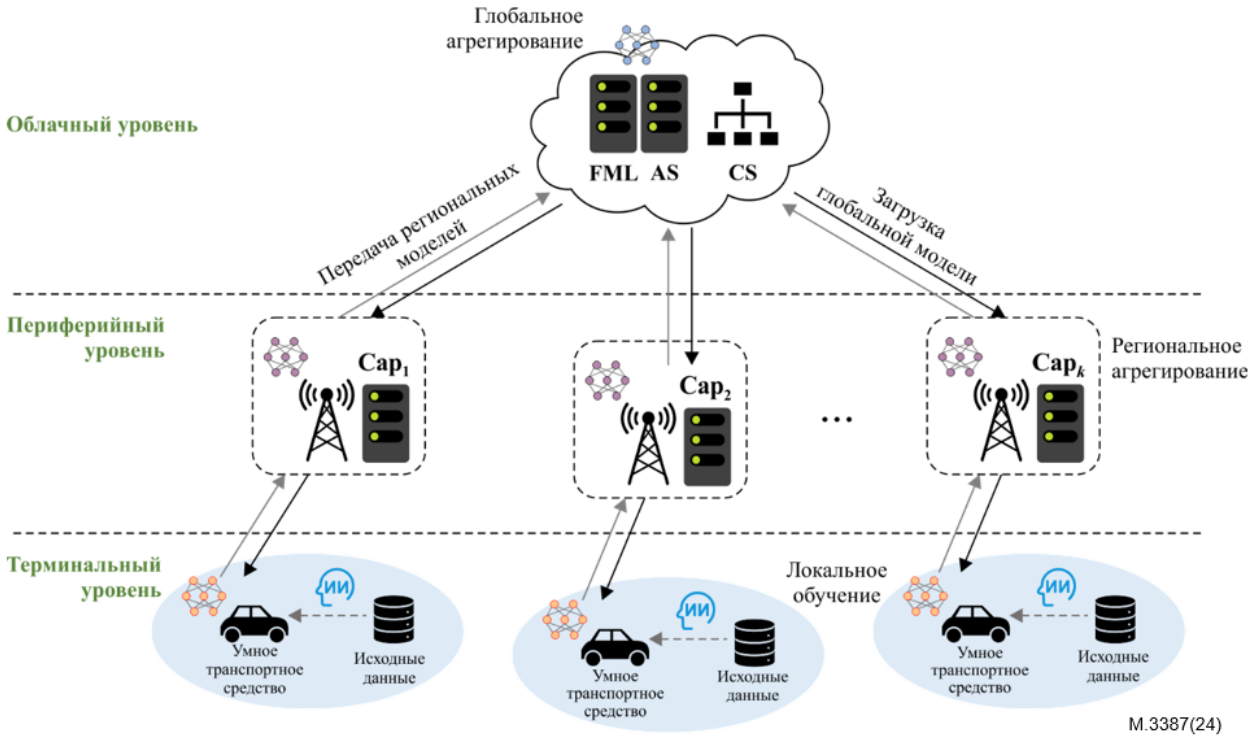

Рисунок I.1 – Сценарий FML с использованием совместной облачной, периферийной и терминальной вычислительной архитектуры в IoV

Как видно из рисунка I.1, функции обучающих узлов FML на каждом уровне распределены следующим образом.

– Терминальный уровень. Владельцы данных, в роли которых выступают умные транспортные средства, обучают локальную FMLM на локально сгенерированных данных и передают параметры получившейся модели в CAP для регионального (вторичного) агрегирования.

– Периферийный уровень. CAP, которые в автомобильных применениях обычно представляют собой придорожные блоки, отвечают, каждая в отдельности, за сбор параметров локальных моделей у владельцев данных в конкретном регионе. Затем они обновляют региональные FMLM посредством регионального (вторичного) агрегирования и передают параметры обновленной модели на агентский сервер. Кроме того, CAP собирают метаданные к исходным данным и информацию о состоянии устройств, а затем передают все это в систему управления федеративным машинным обучением (FMLMS).

– Облачный уровень. В облаке развертывается FMLMS для управления рабочим процессом FML. Одновременно агентский сервер агрегирует параметры всех региональных моделей для обучения глобальной FMLM посредством глобального (первичного) агрегирования. Кроме того, FMLMS осуществляет оценку качества работы FMLS.

I.3 Процесс управления FMLS при обучении модели для обнаружения дорожных аномалий

В сценариях FML с использованием совместной облачной, периферийной и терминальной вычислительной архитектуры FMLMS развертывается в облаке для управления сетевыми ресурсами и качеством услуги FML. Она отвечает за обеспечение эффективного, устойчивого и защищенного предоставления услуги FML в IoV. Приведенный ниже пример иллюстрирует обучение модели для обнаружения дорожных аномалий умными транспортными средствами.

Шаг 1. Все потенциальные владельцы данных (то есть умные транспортные средства) обращаются к сети FML. Клиент федеративного машинного обучения (FMLSC) направляет в FMLMS запрос на получение услуги обучения модели для обнаружения дорожных аномалий через интерфейс 1.

Шаг 2. FMLMS настраивает атрибуты задачи FML в соответствии с запросом на получение услуги, идентифицировав ее как задачу распознавания изображений и назначив ей приоритет. Исходя из имеющихся сетевых ресурсов FMLMS выбирает стратегию стимулирования, побуждающую большее число владельцев данных присоединиться к выполнению задачи обучения модели для обнаружения дорожных аномалий.

На основании выработанной стратегии стимулирования FMLMS определяет состав обучающих узлов FML для выполнения данной задачи FML. Затем FMLMS назначает роли всем обучающим узлам FML. Далее FMLMS определяет свойства соответствующих ресурсов для всех обучающих узлов FML, в том числе свойства ролей, вычислительных ресурсов, ресурсов связи и ресурсов хранения.

Исходя из ролей, назначенных обучающим узлам FML, FMLMS генерирует топологию предоставления услуги FML, включая связи между обучающими узлами и соответствующие конфигурации атрибутов.

После этого FMLMS определяет алгоритм обучения и механизм агрегирования для задачи обнаружения дорожных аномалий, например сверточную нейронную сеть (CNN) и асинхронный механизм агрегирования.

Шаг 3. FMLMS передает в FMLS через интерфейс 2 исходную CNN-модель, свойства системы, разрешения узлов, свойства узлов, топологию предоставления услуги, алгоритм обучения и механизм агрегирования. Затем FMLMS передает в FMLS через интерфейс 2 параметры, относящиеся к алгоритмам конфиденциальности и шифрования канала передачи, обеспечивая тем самым шифрование данных модели и канала передачи для обеспечения конфиденциальности и безопасности FMLM.

Шаг 4. FMLS развертывает исходную CNN-модель на всех обучающих узлах FML и требует от владельцев данных предварительно обработать свои исходные данные в стандартизированном формате. После этого все владельцы данных получают данные изображений в соответствии с запросом на получение услуги для обучения локальной FMLM.

Шаг 5. Каждый владелец данных передает обученную локальную модель в близлежащую CAP. Затем CAP на основе собранных данных локальных моделей генерируют региональные модели и отправляют их на агентский сервер для глобального агрегирования.

Шаг 6. FMLMS осуществляет контроль за качеством работы FMLS и качеством FMLM. FMLMS может корректировать топологию предоставления услуги в соответствии с данными о качестве работы FMLS и определять, следует ли завершить выполнение задачи обучения исходя из текущего качества глобальной FMLM. Когда точность модели достигает той, которая установлена в требованиях к услуге FML, FMLMS передает глобально обученную FMLM в FMLSC через интерфейс 1. Если точность модели не отвечает требованиям к услуге, выполнение задачи обучения продолжается.

Библиография

[b-ITU-T X.1367] Рекомендация МСЭ-Т X.1367 (2020 г.), *Стандартный формат журналов регистрации ошибок в интернете вещей (IoT) для операций, связанных с инцидентами безопасности*.

[b-ITU-T Y.3179] Рекомендация МСЭ-Т Y.3179 (2021 г.), *Архитектурная основа модели машинного обучения в будущих сетях, включая IMT-2020*.

[b-ITU-T Y.4205] Рекомендация МСЭ-Т Y.4205 (2019 г.), *Требования к связанным с IoT краудсорсинговым системам и эталонная модель таких систем*.

СЕРИИ РЕКОМЕНДАЦИЙ МСЭ-Т

Серия A	Организация работы МСЭ-Т
Серия D	Принципы тарификации и учета и экономические и стратегические вопросы международной электросвязи/ИКТ
Серия E	Общая эксплуатация сети, телефонная служба, функционирование служб и человеческие факторы
Серия F	Нетелефонные службы электросвязи
Серия G	Системы и среда передачи, цифровые системы и сети
Серия H	Аудиовизуальные и мультимедийные системы
Серия I	Цифровая сеть с интеграцией служб
Серия J	Кабельные сети и передача сигналов телевизионных и звуковых программ и других мультимедийных сигналов
Серия K	Защита от помех
Серия L	Окружающая среда и ИКТ, изменение климата, электронные отходы, энергоэффективность; конструкция, прокладка и защита кабелей и других элементов линейно-кабельных сооружений
Серия M	**Управление электросвязью, включая СУЭ и техническое обслуживание сетей**
Серия N	Техническое обслуживание: международные каналы передачи звуковых и телевизионных программ
Серия O	Требования к измерительной аппаратуре
Серия P	Качество телефонной передачи, телефонные установки, сети местных линий
Серия Q	Коммутация и сигнализация, а также соответствующие измерения и испытания
Серия R	Телеграфная передача
Серия S	Оконечное оборудование для телеграфных служб
Серия T	Оконечное оборудование для телематических служб
Серия U	Телеграфная коммутация
Серия V	Передача данных по телефонной сети
Серия X	Сети передачи данных, взаимосвязь открытых систем и безопасность
Серия Y	Глобальная информационная инфраструктура, аспекты протокола Интернет, сети последующих поколений, интернет вещей и "умные" города
Серия Z	Языки и общие аспекты программного обеспечения для систем электросвязи

Отпечатано в Швейцарии
Женева, 2024 г.