Обзорная статья
УДК 004.75:519.687.1+004.85
ББК 32.972.5+32.818.1
Д 58
DOI: 10.53598/2410-3225-2022-4-311-60-65

Анализ реализации федеративного обучения в граничных вычислениях с помощью FL-протокола

(*Рецензирована*)

Виталий Анатольевич Довгаль

Майкопский государственный технологический университет,
Адыгейский государственный университет, *Майкоп*, *Россия*, *urmia@mail.ru*

Аннотация. *Целью данного исследования являются обзор и анализ применения сравнительно новой методики машинного обучения в распределенной системе сбора и обработки информации – искусственно-интеллектуальных граничных вычислений, хорошо зарекомендовавшей себя для реализации роя дронов, выполняющих самостоятельный полет с точки зрения передачи собранных данных и обучения модели искусственного интеллекта. Особое внимание уделено одной из парадигм этой методики – федеративного машинного обучения.*

Ключевые слова: *беспроводные сети, Интернет вещей, беспилотные летательные аппараты, федеративное обучение, граничные вычисления, устройства Интернета вещей с элементами искусственного интеллекта, граничный искусственный интеллект*

Review Paper

Analysis of the implementation of federated learning in boundary computing using the FL protocol

Vitaliy A. Dovgal

Maikop State University of Technology, *Adyghe State University*,
Maikop, *Russia*, *urmia@mail.ru*

Abstract. *The purpose of this study is to review and analyze the application of a relatively new machine learning technique in a distributed information collection and processing system – artificially intelligent boundary computing, which has proven itself well for the implementation of a swarm of drones performing independent flight in terms of transmitting collected data and training an artificial intelligence model. Special attention is paid to one of the paradigms of this technique – federated machine learning.*

Keywords: *wireless networks, Internet of Things, unmanned aerial vehicles, federated learning, edge computing, intelligent edge, edge AI*

Введение

Граничные вычисления (Edge Computing, EC) – новая архитектура, расширяющая возможности облачных вычислений (Cloud Computing, CC), расположенная ближе к источникам данных. В сочетании с глубоким обучением (Deep Learning, DL) является многообещающей технологией и широко используется во множестве приложений. В качестве примера можно привести группу беспилотных летательных аппаратов (БПЛА), которые все чаще и чаще используются для выполнения некоторого задания (например, поисковой операции или мониторинговой миссии) в составе организованной группы, которую можно назвать роем (или стаей) [1]. Информация об окружающей обстановке регистрируется датчиками, установленными на дронах, а затем собирается в единый информационный массив с целью дальнейшей обработки информации и принятия какого-либо решения. Однако в обычных архитектурах DL с включенным EC ис-

точникам данных часто приходится отправлять данные и обмениваться ими с третьими сторонами, пограничными или облачными серверами, для обучения своих моделей. Эта архитектура часто непрактична из-за высоких требований к пропускной способности, легализации и уязвимостей конфиденциальности.

Для уменьшения проблем, связанных со снижением пропускной способности каналов связи, обеспечения конфиденциальности данных и их легализации используется несколько методик машинного обучения совокупности объектов, реализующих модель искусственного интеллекта. Одной из таких методик является концепция федеративного обучения (Federated Learning, FL), которую компания Google предлагает как альтернативный подход и многообещающее решение для распределенных клиентов, таких как стая беспилотных летательных аппаратов [2].

Целью данной статьи является проведение анализа методики федеративного обучения интеллектуальной модели для обработки данных на границе вычислительной системы, которой является распределенная вычислительная система, использующая искусственно-интеллектуальные граничные вычисления, как комбинации туманных вычислений и искусственного интеллекта [3].

1. Основы граничных вычислений и федеративного обучения

Федеративное обучение – это концепция распределенного машинного обучения, обеспечивающая решение проблем затрат на связь, конфиденциальность данных во время процесса обучения и их легализации, при котором модели обучаются на конечных устройствах с централизованным контролем без совместного использования локальных наборов данных [4]. Выделенный пограничный сервер или облачный сервер осуществляет агрегирование устройств, периодически собирая обученные параметры для создания и обновления лучшей и более точной модели. Уточненная модель отправляется обратно на граничные устройства для локального обучения.

Как правило, процесс федеративного обучения состоит из пяти этапов:

1. Сначала сервер FL определяет ML-модель (machine learning model) для обучения в локальной базе данных клиентов;

2. Подмножество текущих клиентов выбирается случайным образом или с использованием алгоритмов выбора клиентов, таких как Federated Client Selection (FedCS) [5];

3. Начальная или обновленная глобальная модель отправляется сервером с помощью многоадресной рассылки выбранным клиентам, которые загружают текущие параметры глобальной модели и обучают модель локально;

4. Каждый клиент в подмножестве отправляет обновления на сервер;

5. FL-сервер получает обновления и агрегирует их, используя алгоритмы агрегации (например, FedAvg [6]), для создания новой глобальной модели без доступа к каким-либо данным клиентов.

FL-сервер организует процесс обучения и передает обновления глобальной модели выбранным клиентам в каждой итерации цикла, который повторяется до тех пор, пока не будет достигнут желаемый уровень точности.

На рисунке 1 показана схема федеративного обучения в контексте пограничных вычислений, на которой представлено три типа FL-структур [7]:

а) пограничная структура FL включает в себя группу устройств, находящихся в непосредственной близости, что позволяет рассчитывать глобальную модель обучения на пограничном сервере. Для агрегирования локальные модели после локального обучения затем отправляются на пограничный сервер, расположенный рядом с пограничными устройствами. Пограничный сервер агрегирует и обновляет модель, а затем транслирует ее на конечные устройства. Пограничные серверы часто ограничены в ресурсах, что снижает их вычислительную эффективность;

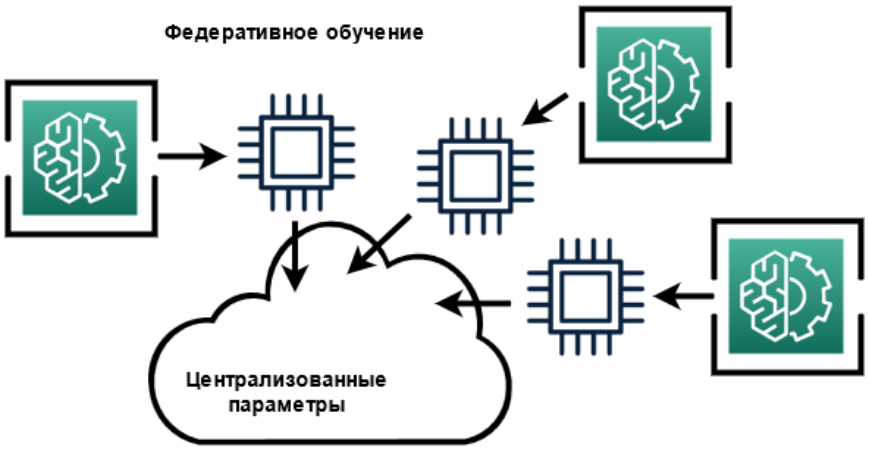

Рис. 1. Схема федеративного машинного обучения
Fig. 1. Scheme of federated machine learning

б) облачная структура FL рассчитывает параметры глобальной модели обучения для граничных географически распределенных систем: клиенты в пределах диапазона пограничного сервера будут сотрудничать в обучении модели DL, а пограничный сервер будет сервером параметров для FL с поддержкой edge;

в) иерархическая (облачная с поддержкой клиента) – структура, в которой сервер параметров обычно расположен рядом с конечным пользователем, что сокращает задержку связи.

Преимущества федеративного обучения как методики заключения модели в защищенную среду и ее последующего обучения без перемещения данных куда-либо, по сравнению с традиционным централизованным обучением ML:

а) время и пропускная способность, необходимые для обучения и вывода, значительно сокращаются, так как методика использует локальные данные, нечасто отправляемые на удаленный сервер (таким образом, обновленная модель может быть использована для прогнозирования на устройстве пользователя, для чего FL обеспечивает конфиденциальность и безопасность пользователя, поскольку данные остаются на персональном устройстве);

б) совместное обучение с использованием FL является простым и потребляет меньше энергии, поскольку модели обучаются на периферийных устройствах (типа беспилотных летающих аппаратов), что делает пограничные вычисления подходящей средой для обучения модели в мобильных пограничных сетях.

Недостатки этих систем обусловлены их звездообразной топологией, в которой отсутствует центральная структура и возникает необходимость согласования передаваемых данных для совместного их использования. Кроме того, узким местом в сети федеративного обучения является низкая скорость коммуникации, обусловленная взаимодействиями между центральным сервером и клиентами, а каждый сервер может поддерживать только ограниченное число клиентов, что приводит к снижению производительности обучения с течением времени.

2. Обзор и анализ реализации федеративного обучения в граничных вычислениях с помощью FL-протокола

Для общего представления об архитектуре системы федеративного обучения в граничных вычислениях необходимо проанализировать сетевой протокол, использующийся на системном уровне и повышающий общую производительность системы. Одним из наиболее подходящих для этих целей можно назвать протокол связи FL, имеющий дело с общим процессом обучения [8]. Протокол учитывает состояние связи между сервером и устройствами, что обеспечивает безопасность связи, зависящую от неста-

бильного подключения устройства и его доступности. Описываемый протокол реализуется на FL-сервере, являющимся распределенным сервисом на основе облака, и конечными устройствами (типа телефонов). Сервер, получив сообщение от устройства о его готовности выполнить FL-операцию для заданной совокупности федеративного обучения, выполняет запрошенную операцию. Проблемы, связанные с обучением или приложением, идентифицируются по глобально уникальному имени в FL-популяции. Задачи, связанные с совокупностью федеративного обучения, включают обучение с использованием предоставленных *n*-мерных параметров или тестирование обученных моделей на основе локальных данных. В заданном временном окне сервер обычно выбирает из потенциально десятков тысяч доступных ему устройств подмножество из нескольких сотен устройств, которые и используются для обработки конкретной FL-задачи. Взаимодействие между устройствами и сервером называется раундом, на протяжении которого устройства остаются подключенными к серверу. Сервер определяет, какие вычисления следует выполнять на выбранных устройствах, для чего он использует FL-план, содержащий график тензорного потока и инструкции о том, как его выполнить.

Как только раунд настроен, сервер отправляет каждому участнику контрольную FL-точку с текущими параметрами глобальной модели и любой другой необходимой информацией. Каждый участник отправляет контрольную FL-точку на сервер на основе своего локального набора данных и глобального состояния.

Сервер параметров федеративного обучения обновляет свое глобальное состояние, и процесс повторяется. Рисунок 2 иллюстрирует протокол связи, используемый для разработки глобальной одноэлементной модели в каждом раунде обучения, который состоит из трех этапов.

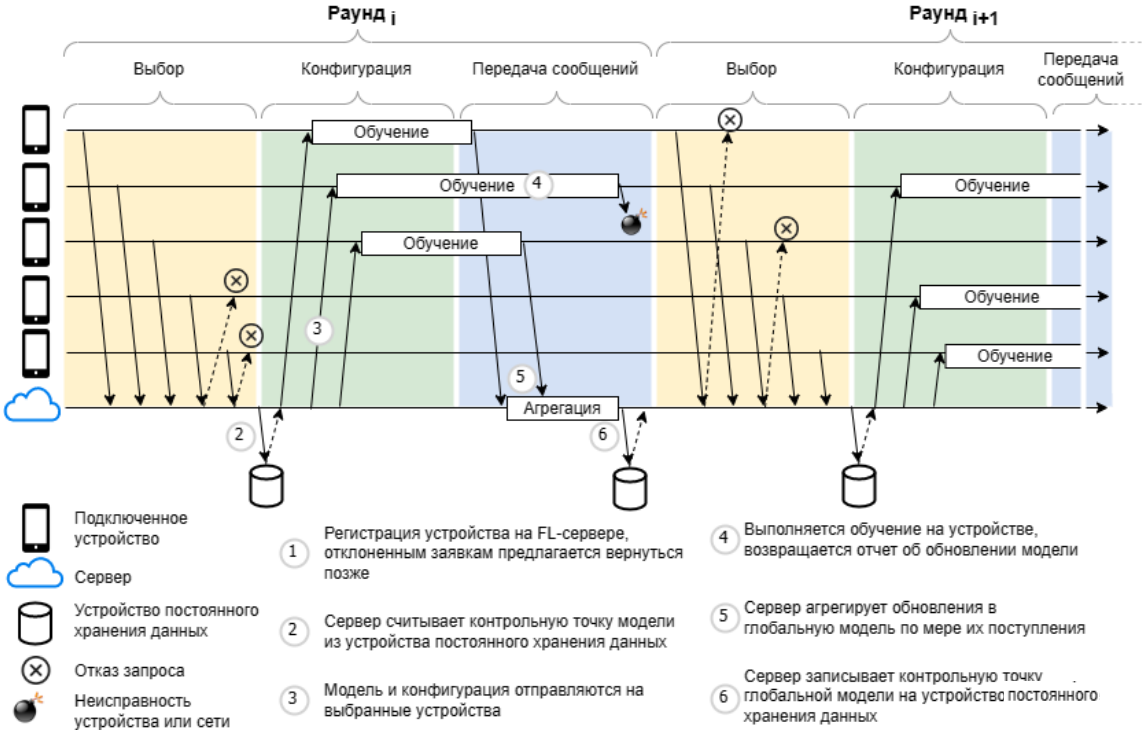

Рис. 2. Протокол федеративного машинного обучения (FL)

Fig. 2. Protocol of federated learning (FL)

1. Выбор: устройства, соответствующие критериям приемлемости, регулярно подключаются к серверу через двунаправленные потоки связи. Доступность клиентов отслеживается через поток: является ли устройство доступным или включено в процесс многоступенчатой коммуникации. Кроме того, сервер FL-параметров выбирает подмножество активных клиентов для участия в раунде обучения, и они выполняют ука-

занную FL-задачу на основе определенного алгоритма выбора клиента (например, FedCS [9]);

2. Конфигурация: для выбранных устройств конфигурация FL-сервера варьируется в зависимости от выбранного метода агрегации (например, простого метода, описанного в [2], или безопасного, представленного в [2]). FL-план и FL-контрольная точка отправляются на каждое выбранное устройство вместе с глобальной моделью;

3. Передача сообщений: сервер FL-параметров ожидает обновлений от клиентов, принимающих участие в обучении. По мере поступления обновлений сервер агрегирует их с использованием предопределенных алгоритмов (например, FedAvg [2]), и уведомляет подключенные устройства о времени повторного подключения. С течением времени к серверу подключается достаточное количество клиентов и федеративное обучение выполняется под его контролем – глобальная модель сервера обновляется; в противном случае раунд будет отменен. Параметры обновления модели часто отправляются на сервер по зашифрованной связи. В целях устранения объективной несогласованности локальных обновлений, выполняемых каждым клиентом в каждом коммуникационном раунде, может быть применен метод FedNova [10], используемый для нормализованного усреднения при сохранении сходимости быстрой ошибки. Существуют и другие методы компенсации неоднородности локальных наборов данных клиентов и скорости их обработки, которые влияют на архитектуру соединения FL-устройств и на дизайн FL-сервера.

Заключение

В данной работе рассмотрена концепция федеративного обучения, которая отлично подходит для приложений граничных вычислений, поскольку она может использовать преимущества периферийных серверов обработки и высоко распределенных периферийных устройств, генерирующих данные. Федеративное обучение позволяет использовать совместную структуру модели глубокого обучения для оптимизации сети граничных вычислений, что может существенно повысить эффективность этой технологии в периферийных вычислительных сетях. В этой статье представлены основы граничных вычислений и федеративного обучения. А также выполнен обзор и анализ реализации федеративного обучения в граничных вычислениях с помощью одноименного протокола FL.

Примечания

1. Довгаль В.А., Довгаль Д.В. Модель взаимодействия анализирующих туманно-облачных вычислений для обработки информации о положении беспилотных летательных аппаратов // Осенние математические чтения в Адыгее: материалы III Международной научной конференции. 2019. С. 149–154.

2. Jakub Konečný, H. Brendan Mc Mahan, Felix X. Yu, Ananda TheerthaSuresh&DaveBacon. Federated Learning: Strategies for Improving Communication Efficiency [Электронный ресурс] URL: https://arxiv.org/pdf/1610.05492.pdf

3. Warnat-Herresthal, Stefanie & Schultze, Hartmut & Shastry, Krishna & Manamohan, Sathyanarayanan & Mukherjee, Saikat & Garg, Vishesh & Sarveswara, Ravi & Händler, Kristian & Pickkers, Peter & Aziz, N. Ahmad & Ktena, Sofia & Siever, Christian & Kraut, Michael & Desai, Milind & Monet, Bruno & Saridaki, Maria & Siegel, Charles & Drews, Anna & Nuesch Germano, Melanie & Schultze, Joachim. (2020). Swarm Learning as a privacy-preserving machine learning approach for disease classification. 10.1101/2020.06.25.171009.

4. Federated Learning: Challenges, Methods and Future Directions / T. Li, A.K. Sahu, A. Talwalkar, V. Smith // IEEE Signal Process. Mag. 2020. No. 37. P. 50–60.

5. Nishio T., Yonetani R. Client Selection for Federated Learning with Heterogeneous Resources in Mobile Edge // 2019 IEEE International Conference on Communications (ICC), Shanghai, China, 20–24 May 2019. P. 1–7.

6. Communication-efficient learning of deep networks from decentralized data / B. McMahan, E. Moore, D. Ramage, S. Hampson, B. Aguera y Arcas // 20th International Conference on Artificial Intelligence and Statistics, PMLR. 2017. No. 54. P. 1273–1282.

7. Dovgal V.A. Swarm learning based on the artificially intelligent edge // CEUR Workshop Proceedings. Ser.: DLT 2021 – Selected Papers of the 6th International Scientific and Practical Conference "Distance Learning Technologies", 2021. P. 260–265.

8. Towards Federated Learning at scale: System design. K. Bonawitz, H. Eichner, W. Grieskamp, D. Huba, A. Ingerman, V. Ivanov // J Roselander. arXiv 2019. arXiv:1902.01046.

9. Nishio T., Yonetani R. Client Selection for Federated Learning with Heterogeneous Resources in Mobile Edge. In Proceedings of the ICC 2019 // IEEE International Conference on Communications (ICC) – 2019, Shanghai, China, 20–24 May 2019. P. 1–7.

10. A Novel Framework for the Analysis and Design of Heterogeneous Federated Learning / J. Wang, Q. Liu, H. Liang, J. Gauri, H.V. Poor // IEEE Transport Signal Processing. 2021. No. 69. P. 5234–5249.

References

1. Dovgal V.A., Dovgal D.V. Model of interaction of analyzing fog-cloud computing for processing information about the position of unmanned aerial vehicles // Autumn Mathematical Readings in Adyghea: Materials of the 3rd International Scientific Conference. 2019. P. 149–154.

2. Jakub Konečný, H. Brendan Mc Mahan, Felix X. Yu, Ananda TheerthaSuresh&DaveBacon. Federated Learning: Strategies for Improving Communication Efficiency [Electronic resource]. URL: https://arxiv.org/pdf/1610.05492.pdf

3. Warnat-Herresthal, Stefanie & Schultze, Hartmut & Shastry, Krishna & Manamohan, Sathyanarayanan & Mukherjee, Saikat & Garg, Vishesh & Sarveswara, Ravi & Händler, Kristian & Pickkers, Peter & Aziz, N. Ahmad & Ktena, Sofia & Siever, Christian & Kraut, Michael & Desai, Milind & Monet, Bruno & Saridaki, Maria & Siegel, Charles & Drews, Anna & Nuesch Germano, Melanie & Schultze, Joachim. (2020). Swarm Learning as a privacy-preserving machine learning approach for disease classification. 10.1101/2020.06.25.171009.

4. Federated Learning: Challenges, Methods and Future Directions / T. Li, A.K. Sahu, A. Talwalkar, V. Smith // IEEE Signal Process. Mag. 2020. No. 37. P. 50–60.

5. Nishio T., Yonetani R. Client Selection for Federated Learning with Heterogeneous Resources in Mobile Edge // 2019 IEEE International Conference on Communications (ICC), Shanghai, China, 20–24 May 2019. P. 1–7.

6. Communication-efficient learning of deep networks from decentralized data / B. McMahan, E. Moore, D. Ramage, S. Hampson, B. Aguera y Arcas // 20th International Conference on Artificial Intelligence and Statistics, PMLR. 2017. No. 54. P. 1273–1282.

7. Dovgal V.A. Swarm learning based on the artificially intelligent edge // CEUR Workshop Proceedings. Ser.: DLT 2021 – Selected Papers of the 6th International Scientific and Practical Conference "Distance Learning Technologies", 2021. P. 260–265.

8. Towards Federated Learning at scale: System design. K. Bonawitz, H. Eichner, W. Grieskamp, D. Huba, A. Ingerman, V. Ivanov // J Roselander. arXiv 2019. arXiv:1902.01046.

9. Nishio T., Yonetani R. Client Selection for Federated Learning with Heterogeneous Resources in Mobile Edge. In Proceedings of the ICC 2019 // IEEE International Conference on Communications (ICC) – 2019, Shanghai, China, 20–24 May 2019. P. 1–7.

10. A Novel Framework for the Analysis and Design of Heterogeneous Federated Learning / J. Wang, Q. Liu, H. Liang, J. Gauri, H.V. Poor // IEEE Transport Signal Processing. 2021. No. 69. P. 5234–5249.

Статья поступила в редакцию 27.11.2022; *одобрена после рецензирования 20.12.2022*; *принята к публикации 21.12.2022*.

The article was submitted 27.11.2022; *approved after reviewing 20.12.2022*; *accepted for publication 21.12.2022*.