УДК 004.85

Иванов Илья Александрович, магистрант, Московский государственный технический университет им. Н.Э. Баумана, г. Москва

СРАВНИТЕЛЬНЫЙ АНАЛИЗ МЕТОДОВ ФЕДЕРАТИВНОГО ОБУЧЕНИЯ

Аннотация. В статье проводится сравнительный анализ методов федеративного обучения – распределённой парадигмы машинного обучения без передачи исходных данных клиентов на сервер. Выделены четыре группы: базовые алгоритмы агрегации, методы для гетерогенных данных, персонализированное федеративное обучение и приватные подходы. Сравнение проведено по семи критериям; сформулированы рекомендации по выбору метода в зависимости от характера задачи.

Abstract. This article presents a comparative analysis of federated learning methods – a distributed machine learning paradigm without sending raw client data to the server. Four groups are identified: baseline aggregation algorithms, methods for heterogeneous data, personalized federated learning, and privacy-preserving approaches. The comparison was performed using seven criteria; recommendations are formulated for selecting a method depending on the task.

Ключевые слова: федеративное обучение, распределённое обучение, FedAvg, гетерогенность данных, non-IID, персонализированные модели, дифференциальная приватность, безопасная агрегация, машинное обучение

Keywords: federated learning, distributed learning, FedAvg, data heterogeneity, non-IID, personalized models, differential privacy, secure aggregation, machine learning

Введение

Актуальность исследования связана с двумя встречными тенденциями: законодательная база (GDPR, ФЗ № 152) всё сильнее ограничивает централизованный сбор персональных данных, а пользовательские устройства приобретают вычислительные возможности, достаточные для обучения нейронных сетей непосредственно на устройстве. Федеративное обучение (federated learning, FL) отвечает на эти изменения: сырые пользовательские данные никогда не покидают устройства, а обмен с сервером ограничен промежуточными весами модели [8, 12].

К настоящему моменту предложены десятки методов федеративного обучения, различающихся по способу агрегации, устойчивости к гетерогенности данных и гарантиям приватности. Выбор подходящего метода для конкретного приложения требует учёта статистики данных, вычислительных возможностей клиентов и требований к защите информации [3].

Обзоры по отдельным ветвям FL не снимают главной сложности: в литературе нет единого набора критериев, в которых можно было бы сопоставить методы, выросшие из разных идей, и выбрать подходящий подход под конкретный сценарий внедрения.

Научная новизна работы заключается в построении сводной классификации актуальных методов федеративного обучения с оценкой по единому набору семи критериев, охватывающих качество, эффективность коммуникации, инженерные ограничения и гарантии приватности.

Цель настоящей работы – провести сравнительный анализ основных групп методов федеративного обучения, выявить их преимущества и ограничения и сформулировать рекомендации по выбору метода в зависимости от характера решаемой задачи.

1. Постановка задачи и таксономия методов

В классической постановке федеративного обучения [8] N клиентов хранят локальные наборы D_k объёма n_k; передача D_k за пределы клиента не предусматривается. Глобальная цель – минимизация функции потерь, усреднённой по локальным выборкам клиентов:

$$F(w) = \Sigma_k (n_k/n) \cdot F_k(w), \quad (1)$$

где $F_k(w)$ – локальная функция потерь k-го клиента, n_k – объём его локальных данных, $n = \Sigma_k n_k$ – суммарный объём, w – параметры модели.

Алгоритм FedAvg [8], лежащий в основе большинства FL-систем, выражает обновление как взвешенное усреднение локальных весов w_{t+1}^k, полученных k-м клиентом после нескольких внутренних эпох SGD:

$$w_{t+1} = \Sigma_k (n_k/n) \cdot w_{t+1}^k. \quad (2)$$

Известные подходы разделены на четыре группы, различающиеся по характеру вмешательства в протокол FedAvg: базовые алгоритмы агрегации, методы для гетерогенных данных, персонализированное FL и приватные подходы.

2. Базовые алгоритмы агрегации

Первая группа работает в предположении, что данные клиентов порождаются одним распределением (IID). В FedSGD обновление модели выполняется по усреднённым градиентам, собираемым на каждом шаге; при равном глобальном батче это поведение совпадает с централизованным SGD, однако требует постоянного сетевого обмена.

Алгоритм FedAvg [8] переходит от обмена градиентами к обмену весами: между соседними раундами клиенты проводят несколько внутренних эпох обучения на локальных выборках, что уменьшает число сетевых взаимодействий в 10–100 раз. Допускается включение в раунд лишь подмножества клиентов, что критично при развёртываниях с миллионами устройств.

К сильным сторонам базовых алгоритмов относятся минимальные инженерные требования и совместимость с типовыми обучающими пайплайнами. Слабое место – резкое падение итогового качества, когда локальные распределения уходят друг от друга (эффект клиентского дрейфа).

3. Методы для гетерогенных данных

Следующая группа отвечает на главный эмпирический вызов FL: выборки на разных устройствах порождаются различными распределениями, в результате локальные параметры смещаются от истинного оптимума, ухудшая качество агрегированной модели [13].

FedProx [6] вводит в локальную целевую функцию проксимальный член $(\mu/2)\cdot\|w - w_t\|^2$, удерживающий локальные решения вблизи текущих глобальных весов и разрешающий клиентам выполнять разное число шагов.

SCAFFOLD [4] вводит оценки разницы между средним и локальными градиентами и вычитает их из локального шага, благодаря чему сходимость не зависит от степени разнородности выборок.

FedNova [11] перевзвешивает вклады пропорционально проделанному числу шагов, избавляя итоговые веса от перекоса, вызванного различиями в вычислительных возможностях клиентов, и применяется как надстройка над любым из этих алгоритмов.

Преимущество группы – сохранение схемы FedAvg при точечном введении поправок; расходы возникают из-за добавления вспомогательных величин в обмен и из необходимости подбора веса регуляризатора.

4. Персонализированное федеративное обучение

Третья группа отказывается от поиска «общей модели для всех» в пользу совместного обучения персональных w_k, разделяющих общие представления, но способных адаптироваться к индивидуальным распределениям.

Per-FedAvg [2] опирается на идеи мета-обучения MAML: глобальные веса подбираются как «хорошая стартовая точка», из которой один градиентный шаг на данных клиента даёт качественную персональную модель.

pFedMe [10] разводит глобальные и персональные веса через огибающую Моро: локальная модель оптимизируется под конкретного клиента, а глобальная выступает опорной точкой.

Ditto [7] разносит две задачи: обычное усреднение FedAvg вырабатывает общие веса, а каждый клиент отдельно дообучает собственную модель вблизи этих весов, что одновременно даёт индивидуальное качество и снижает влияние вредоносных клиентов.

Группа эффективна при сильно различающихся локальных данных, но расходует вдвое больше памяти на хранение двух наборов весов и требует пересмотра процедуры валидации.

5. Безопасные и приватные подходы

Четвёртая группа работает в предположении, что сам агрегатор или внешний наблюдатель может быть недобросовестным. Перехваченные обновления при определённых условиях позволяют восстановить входные примеры клиента – так называемые атаки инверсии градиентов [3].

Secure Aggregation [1] маскирует индивидуальные обновления парными псевдослучайными векторами; при суммировании всех вкладов маски сокращаются, и агрегатор видит только результат. Протокол остаётся работоспособным даже при потере до трети участников раунда.

DP-FedAvg [9] вводит (ε, δ)-дифференциальную приватность: локальные вклады урезаются по норме, после чего к сумме добавляется гауссов шум с дисперсией, калиброванной под целевой уровень (ε, δ). Строгость гарантий достигается ценой точности: чем больше шума, тем ниже финальное качество модели.

Схемы гомоморфного шифрования (CKKS, BFV) обеспечивают приватную агрегацию криптографическими средствами, выполняя сложение над зашифрованными векторами; их применение ограничено двумя–тремя порядками просадки производительности относительно FedAvg.

Ключевое достоинство этих методов – строгая доказуемость защиты, востребованная в сферах здравоохранения, финансов и других регулируемых отраслях. Обратная сторона – полный или частичный отказ от использования

сырых градиентов и, как следствие, повышенные вычислительные и сетевые издержки.

6. Сводное сравнение групп методов

Для сравнения выделенных групп методов использованы семь критериев, отражающих как качественные характеристики моделей, так и инженерные аспекты их применения: точность на IID-данных, робастность к non-IID, коммуникационная стоимость, вычислительная нагрузка на клиента, гарантии приватности, скорость сходимости и простота реализации. Результаты сравнения приведены в таблице 1.

Таблица 1 – Сводное сопоставление групп методов федеративного обучения

Критерий	Базовые алгоритмы	Гетерогенные данные	Персонализированные	Приватные подходы
Точность на IID	Высокая	Высокая	Высокая	Средняя
Робастность к non-IID	Низкая	Высокая	Высокая	Средняя
Коммуникационная стоимость	Низкая	Средняя	Средняя	Высокая
Нагрузка на клиента	Низкая	Средняя	Средняя	Высокая
Гарантии приватности	Отсутствуют	Отсутствуют	Слабые	Формальные
Скорость сходимости	Высокая (IID)	Средняя	Средняя	Низкая
Простота реализации	Высокая	Средняя	Средняя	Низкая

Данные таблицы 1 показывают, что каждая из четырёх групп оптимальна лишь в своём режиме использования, и выбор, по сути, сводится к балансу четырёх параметров: однородность данных, инженерная сложность, степень индивидуальности клиентов и строгость требований к приватности.

На близких по распределению выборках оптимален FedAvg – его легко интегрировать в существующие пайплайны. При выраженной гетерогенности

выборок предпочтительны FedProx или SCAFFOLD, восстанавливающие сходимость без отказа от единой модели. Там же, где поведение пользователей сильно различается (предсказание ввода текста, персональные видео- и музыкальные рекомендации), персонализированные подходы дают заметный выигрыш. Наконец, в сферах с юридическими требованиями к защите данных приватные техники комбинируются с любым из предыдущих подходов: безопасная агрегация не вносит потерь качества модели, а DP-FedAvg добавляется, когда нужны формальные гарантии.

Заключение

По результатам проведённого исследования сформулированы следующие выводы.

1. Предложена таксономия методов федеративного обучения, выделяющая четыре группы по характеру решаемой задачи: базовые алгоритмы агрегации, методы для гетерогенных данных, персонализированное федеративное обучение и приватные подходы.

2. Показано, что ни одна из выделенных групп не превосходит остальные по всем семи критериям одновременно: качество при гетерогенных данных достигается ценой вычислительных затрат, а формальная приватность – ценой снижения точности.

3. Перспективно комбинирование подходов: SCAFFOLD или FedProx для устойчивости к гетерогенности совместно с безопасной агрегацией и персонализацией на стороне клиента.

4. Открытыми остаются вопросы об устойчивости федеративного обучения к скоординированным атакам отравления и о метриках оценки персонализированных моделей.

Список источников

1. Bonawitz K. [и др.]. Practical secure aggregation for privacy-preserving machine learning [Electronic resource] // Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security. – 2017. – P. 1175–1191. – DOI: 10.1145/3133956.3133982. – Access mode: https://doi.org/10.1145/3133956.3133982 (accessed: 22.04.2026).

2. Fallah A., Mokhtari A., Ozdaglar A. Personalized federated learning with theoretical guarantees: a model-agnostic meta-learning approach [Electronic resource] // Advances in Neural Information Processing Systems. – 2020. – Vol. 33. – P. 3557–3568. – Access mode: https://arxiv.org/abs/2002.07948 (accessed: 19.04.2026).

3. Kairouz P. [и др.]. Advances and open problems in federated learning [Electronic resource] // Foundations and Trends in Machine Learning. – 2021. – Vol. 14, No. 1–2. – P. 1–210. – DOI: 10.1561/2200000083. – Access mode: https://arxiv.org/abs/1912.04977 (accessed: 17.04.2026).

4. Karimireddy S.P. [и др.]. SCAFFOLD: stochastic controlled averaging for federated learning [Electronic resource] // Proceedings of the 37th International Conference on Machine Learning. – 2020. – P. 5132–5143. – Access mode: https://arxiv.org/abs/1910.06378 (accessed: 18.04.2026).

5. Konečný J. [и др.]. Federated learning: strategies for improving communication efficiency [Electronic resource] // arXiv preprint arXiv:1610.05492. – 2016. – Access mode: https://arxiv.org/abs/1610.05492 (accessed: 21.04.2026).

6. Li T. [и др.]. Federated optimization in heterogeneous networks [Electronic resource] // Proceedings of Machine Learning and Systems. – 2020. – Vol. 2. – P. 429–450. – Access mode: https://arxiv.org/abs/1812.06127 (accessed: 17.04.2026).

7. Li T. [и др.]. Ditto: fair and robust federated learning through personalization [Electronic resource] // Proceedings of the 38th International Conference on Machine Learning. – 2021. – P. 6357–6368. – Access mode: https://arxiv.org/abs/2012.04221 (accessed: 22.04.2026).

8. McMahan B. [и др.]. Communication-efficient learning of deep networks from decentralized data [Electronic resource] // Proceedings of the 20th International Conference on Artificial Intelligence and Statistics. – 2017. – Vol. 54. – P. 1273–1282. – Access mode: https://arxiv.org/abs/1602.05629 (accessed: 14.04.2026).

9. McMahan B. [и др.]. Learning differentially private recurrent language models [Electronic resource] // International Conference on Learning Representations. – 2018. – Access mode: https://arxiv.org/abs/1710.06963 (accessed: 16.04.2026).

10. T Dinh C., Tran N., Nguyen J. Personalized federated learning with Moreau envelopes [Electronic resource] // Advances in Neural Information Processing Systems. – 2020. – Vol. 33. – P. 21394–21405. – Access mode: https://arxiv.org/abs/2006.08848 (accessed: 22.04.2026).

11. Wang J. [и др.]. Tackling the objective inconsistency problem in heterogeneous federated optimization [Electronic resource] // Advances in Neural Information Processing Systems. – 2020. – Vol. 33. – P. 7611–7623. – Access mode: https://arxiv.org/abs/2007.07481 (accessed: 15.04.2026).

12. Yang Q. [и др.]. Federated machine learning: concept and applications [Electronic resource] // ACM Transactions on Intelligent Systems and Technology. – 2019. – Vol. 10, No. 2, Article 12. – P. 1–19. – DOI: 10.1145/3298981. – Access mode: https://doi.org/10.1145/3298981 (accessed: 13.04.2026).

13. Zhao Y. [и др.]. Federated learning with non-IID data [Electronic resource] // arXiv preprint arXiv:1806.00582. – 2018. – Access mode: https://arxiv.org/abs/1806.00582 (accessed: 18.04.2026).

References

1. Bonawitz K. et al. Practical secure aggregation for privacy-preserving machine learning [Electronic resource]. Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security, 2017, pp. 1175–1191. DOI: 10.1145/3133956.3133982. – Access mode: https://doi.org/10.1145/3133956.3133982 (accessed: 22.04.2026).

2. Fallah A., Mokhtari A., Ozdaglar A. Personalized federated learning with theoretical guarantees: a model-agnostic meta-learning approach [Electronic resource]. Advances in Neural Information Processing Systems, 2020, vol. 33, pp. 3557–3568. – Access mode: https://arxiv.org/abs/2002.07948 (accessed: 19.04.2026).

3. Kairouz P. et al. Advances and open problems in federated learning [Electronic resource]. Foundations and Trends in Machine Learning, 2021, vol. 14, no. 1–2, pp. 1–210. DOI: 10.1561/2200000083. – Access mode: https://arxiv.org/abs/1912.04977 (accessed: 17.04.2026).

4. Karimireddy S.P. et al. SCAFFOLD: stochastic controlled averaging for federated learning [Electronic resource]. Proceedings of the 37th International Conference on Machine Learning, 2020, pp. 5132–5143. – Access mode: https://arxiv.org/abs/1910.06378 (accessed: 18.04.2026).

5. Konečný J. et al. Federated learning: strategies for improving communication efficiency [Electronic resource]. arXiv preprint arXiv:1610.05492, 2016. – Access mode: https://arxiv.org/abs/1610.05492 (accessed: 21.04.2026).

6. Li T. et al. Federated optimization in heterogeneous networks [Electronic resource]. Proceedings of Machine Learning and Systems, 2020, vol. 2, pp. 429–450. – Access mode: https://arxiv.org/abs/1812.06127 (accessed: 17.04.2026).

7. Li T. et al. Ditto: fair and robust federated learning through personalization [Electronic resource]. Proceedings of the 38th International Conference on Machine Learning, 2021, pp. 6357–6368. – Access mode: https://arxiv.org/abs/2012.04221 (accessed: 22.04.2026).

8. McMahan B. et al. Communication-efficient learning of deep networks from decentralized data [Electronic resource]. Proceedings of the 20th International Conference on Artificial Intelligence and Statistics, 2017, vol. 54, pp. 1273–1282. – Access mode: https://arxiv.org/abs/1602.05629 (accessed: 14.04.2026).

9. McMahan B. et al. Learning differentially private recurrent language models [Electronic resource]. International Conference on Learning Representations, 2018. – Access mode: https://arxiv.org/abs/1710.06963 (accessed: 16.04.2026).

10. T Dinh C., Tran N., Nguyen J. Personalized federated learning with Moreau envelopes [Electronic resource]. Advances in Neural Information Processing Systems, 2020, vol. 33, pp. 21394–21405. – Access mode: https://arxiv.org/abs/2006.08848 (accessed: 22.04.2026).

11. Wang J. et al. Tackling the objective inconsistency problem in heterogeneous federated optimization [Electronic resource]. Advances in Neural Information Processing Systems, 2020, vol. 33, pp. 7611–7623. – Access mode: https://arxiv.org/abs/2007.07481 (accessed: 15.04.2026).

12. Yang Q. et al. Federated machine learning: concept and applications [Electronic resource]. ACM Transactions on Intelligent Systems and Technology, 2019, vol. 10, no. 2, article 12, pp. 1–19. DOI: 10.1145/3298981. – Access mode: https://doi.org/10.1145/3298981 (accessed: 13.04.2026).

13. Zhao Y. et al. Federated learning with non-IID data [Electronic resource]. arXiv preprint arXiv:1806.00582, 2018. – Access mode: https://arxiv.org/abs/1806.00582 (accessed: 18.04.2026).