ПРОГРАММИРОВАНИЕ, 2024, № 4, с. 27–40

ПРОГРАММНАЯ ИНЖЕНЕРИЯ, ТЕСТИРОВАНИЕ И ВЕРИФИКАЦИЯ ПРОГРАММ

УДК 004.89

АНАЛИТИЧЕСКИЙ ОБЗОР КОНФИДЕНЦИАЛЬНОГО ИСКУССТВЕННОГО ИНТЕЛЛЕКТА: МЕТОДЫ И АЛГОРИТМЫ РЕАЛИЗАЦИИ В ОБЛАЧНЫХ ВЫЧИСЛЕНИЯХ

© 2024 г. Е. М. Ширяев[a,*], А. С. Назаров[a,], Н. Н. Кучеров[a,***], М. Г. Бабенко[a,b,****]**

[a] *Северо-Кавказский федеральный университет*
355017 Ставрополь, ул. Пушкина, д. 1, Россия
[b] *Институт системного программирования РАН им. В. П. Иванникова*
109004 Москва, ул. А. Солженицына, д. 25, Россия
**E-mail: eshiriaev@ncfu.ru*
***E-mail: anazarov@ncfu.ru*
****E-mail: nkucherov@ncfu.ru*
*****E-mail: mgbabenko@ncfu.ru*

Поступила в редакцию 21.02.2024
После доработки 18.03.2024
Принята к публикации 18.03.2024

Технологии искусственного интеллекта и облачных систем в последнее время активно развиваются и внедряются. В связи с этим обострился вопрос их совместного использования, актуальный уже несколько лет. Проблема сохранения конфиденциальности данных в облачных вычислениях приобрела статус критической задолго до возникновения необходимости их совместного использования с искусственным интеллектом, который сделал ее еще более сложной. В данной статье представлен обзор как самих методов искусственного интеллекта и облачных вычислений, так и методов обеспечения конфиденциальности данных. В обзоре рассмотрены методы, использующие дифференциальную конфиденциальность; схемы разделения секрета; гомоморфное шифрование; гибридные методы. Проведенное исследование показало, что каждый рассмотренный метод имеет свои плюсы и минусы, обозначенные в работе, однако универсальное решение отсутствует. Было установлено, что теоретические модели гибридных методов, основанных на схемах разделения секрета и полностью гомоморфном шифровании, позволяют существенно повысить конфиденциальность обработки данных с использованием искусственного интеллекта.

Ключевые слова: облачные вычисления, искусственный интеллект, нейронная сеть, схемы разделения секрета, гомоморфное шифрование, система остаточных классов

DOI: 10.31857/S0132347424040036, **EDN:** PTIGVO

1. ВВЕДЕНИЕ

Технологии искусственного интеллекта (ИИ) получают все большее распространение в производственной и повседневной жизни общества. Рост популярности и быстрое развитие технологий приводят к усложнению задач, решаемых с помощью ИИ, и, как следствие, к тому, что обработка информации на стандартном пользовательском устройстве выполняется критически неэффективно, что неприемлемо для конечного потребителя. В данном случае выходом является применение облачных технологий (ОТ), которые в свою очередь уже обрели широкую популярность и получили развитие своей методологии. Стоит также отметить, что ИИ и ОТ вызывают большой интерес и в научном сообществе. Например, такие проекты как [1, 2], обрели популярность в повседневной сфере деятельности при этом являются научными проектами в области ИИ. При совместном использовании методов ИИ и ОТ появляется ряд как стандартных, так и специфических проблем, связанных с надежностью, безопасностью и конфиденциальностью. Учитывая, что из всего функционала, реализуемого ОТ, ИИ использует в основном облачные вычисления (ОВ), для ИИ с использованием ОТ имеют место все проблемы, характерные для ОВ. Угрозы безопасности ОВ в целом можно разделить на две категории: внешние угрозы и внутренние. К внешним угрозам относятся различные атаки

злоумышленников в целях кражи информации (например взлом) или повреждения информации (например DDOS-атаки) [3]. Внутренние угрозы в целом представляют собой совокупность всех возможных способов компрометации системы безопасности изнутри. Для борьбы с последними используют, например, схемы разделения секрета [4]. Однако наиболее высокий уровень безопасности достигается при использовании методов, которые позволяют обрабатывать данные в зашифрованном виде. В таком случае вероятность компрометации системы сокращается до минимума. Возможным решением проблемы конфиденциальной обработки данных является гомоморфное шифрование. В данной работе представлено исследование современных используемых на практике методов и алгоритмов обеспечения конфиденциальности ИИ и предиктивное исследование методов, которые возможно будут применяться в будущем.

Работа состоит из 4 разделов. В разделе 2 рассмотрены основные аспекты, связанные с ИИ и облачными технологиями. В разделе 3 рассмотрены подходы к построению сохраняющего конфиденциальность ИИ. В разделе 4 представлены результаты проведенного аналитического обзора.

2. ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ И ОБЛАЧНЫЕ ТЕХНОЛОГИИ

2.1. Искусственный интеллект

Искусственный интеллект (ИИ) изначально развивался в рамках интеллектуальных систем и до сих пор считается одной из их составляющих, а именно способностью выполнять творческие функции. Изначально необходимость в интеллектуальных системах была обусловлена автоматизацией принятия решений. То есть от системы ожидалась определенная реакция на определенные события. С развитием вычислительной техники, методов и алгоритмов, а также способов разработки систем и приложений данные реакции усложнялись и становились более гибкими.

Исторически первыми методами ИИ были методы машинного обучения (МО). МО представляет собой класс методов ИИ, задачей которых является не прямое решение задачи, а нахождение его за счет обучения на основе анализа множества решений сходных задач [5]. Постановку задачи МО можно определить следующим образом. Существует некая неизвестная зависимость между двумя множествами. Известны только прецеденты, т. е. пары из этих двух множеств, которые называются обучающей выборкой. На основе этих данных ставится задача восстановления зависимости, т. е. построения алгоритма, который с заданной точностью создаст новую пару. По сути, перед МО ставится задача аппроксимации функции, но не обязательно другой функцией, а неким алгоритмом.

Методы ИИ также можно разделить по способу обучения. Так, например, существуют модели обучения с подкреплением [6], среди таких моделей можно выделить генетические алгоритмы. Генетические алгоритмы представляют собой эвристические алгоритмы поиска, использующиеся для решения задач оптимизации и моделирования, путем случайного подбора, комбинирования и вариации исходных параметров, которые основаны на методах, подобных естественному отбору в природе [7]. Также существует обучение без учителя, подобными методами решается, например, задача кластеризации [8]. Однако наибольшую группу методов составляют методы, использующие обучение с учителем, где для множества прецедентов (известных пар входных и выходных данных) необходимо построить алгоритм, возвращающий требуемое решение [9].

Искусственные нейронные сети, или просто нейронные сети (НС), представляют собой математическую модель, которая построена по принципу функционирования сетей нервных клеток живого организма [10]. Развитие вычислительной техники позволило создавать модели НС большой сложности. В этом контексте можно выделить несколько событий, которые позволили расширить применение ИИ, это появление различных аппаратных и графических ускорителей [11–13], а также вентильных матриц [14–17] для ИИ, которые позволяют решать более широкий спектр задач, недоступный ранее. Работа таких моделей НС осуществляется за счет применения, так называемого глубокого обучения (ГО). По сути, ГО это процесс обучения многослойных НС. Теоретически 2–3-слойных НС достаточно для решения широкого круга задач, однако, для решения сложных задач зачастую используется ГО, которое показывает хорошие результаты [18]. К ГО относят такие методы как ограниченная машина Больцмана. Также на основе ГО строятся сверточные нейронные сети, которые используют на различных слоях различные формы сверток [19]. Применяются они зачастую для распознавания образов [20]. Одной из наиболее популярных современных технологий ИИ является

технология GPT. Даже в начале развития GPT модели требовали больших обучающих выборок, которые занимают десятки гигабайт на стадии предварительного обучения, а учитывая тот факт, что количество пользователей моделей с GTP превышает 100 млн, объемы потребляемых ресурсов превышают ресурсы доступные одному устройству, если рассматривать среднестатистический офисный компьютер. Для эффективной работы таких технологий требуется применение распределенных вычислительных систем.

2.2. Облачные вычисления

Рассмотрим более подробно облачные технологии (ОТ). По сути, ОТ и ОВ являются синонимами, так как любая обработка данных в облаке предполагает какие-либо вычисления. Аналогично ситуация обстоит с облачными хранилищами, так как при обработке информации (например, поиске) также производятся определенные вычисления.

ОВ являются развитием модели распределенных вычислений [21] за некоторыми исключениями. Распределенные вычисления предполагают наличие параллелизма в вычислениях, а именно объединение вычислительных ресурсов в параллельную вычислительную систему. Реализация такой системы возможна и на одном физическом устройстве, например серверной стойке или суперкомпьютере [22]. Переходной точкой между распределенными вычислениями и облачными вычислениями можно считать грид вычисления [23]. Главное же отличие ОВ заключается в концептуализации [24]. В отличие от распределенных и грид вычислений, которые являются прежде всего совокупностью средств и способов решения вычислительно сложных задач, ОВ прежде всего являются сервисом (услугой) по предоставлению возможностей для реализации эластичных вычислений. Принято классифицировать ОВ по видам моделей, в рамках которых предоставляется данная услуга:

• Software as a Service (SaaS) – эта модель подразумевает использование облачной инфраструктуры для реализации ОВ посредством предоставления пользователю пакета прикладного программного обеспечения. Контроль и управление инфраструктурой производится исключительно провайдером услуг [25];

• Platform as a Service (PaaS) – в таком случае пользователю представляется инфраструктура для размещения различного базового программного обеспечения. Обычно это инструментальные средства для создания программного обеспечения и, например, системы управления базами данных. Также провайдером могут предоставляться различные среды для работы с языками программирования [26];

• Infrastructure as a Service (IaaS) – в данном варианте поставщик предоставляет пользователю наиболее полные права по пользованию облаком. Здесь пользователю предлагается базовая инфраструктура, в рамках которой он самостоятельно организует процессы управления вычислительными ресурсами, а также построения сети и хранения данных. Также пользователь самостоятельно контролирует операционные системы, которые разворачиваются в облаке. Еще одно отличие от предыдущих категорий – пользователю предоставляется ограниченный контроль над сетевыми сервисами выделенного ему облачного пространства [27, 28].

На основе данных трех моделей выделяют и различные гибридные, такие как Data Base as a Service [29], Monitoring as a Service [30] и т. п., однако гибридные виды, по сути, разделяются с точки зрения потребностей клиента и скорее направлены на узкую специализацию применения того или иного вида ОВ. На основе вышеизложенной информации можно выделить преимущества, которые дают ОВ для ИИ. Как уже было сказано, методы ИИ являются достаточно вычислительно сложными. При проведении исследований и разработке приложений и сервисов, сопряженных с ИИ, исследователь/разработчик сталкивается с трудностями, связанными с ограниченностью вычислительных ресурсов. В случае исследователя, в теории, можно воспользоваться распределенной вычислительной системой своего учреждения, однако не в каждом учреждении имеется собственная мощная вычислительная система, либо к ней затруднен доступ. ОВ позволяют исследователю/разработчику арендовать вычислительные ресурсы у поставщика.

2.3. Облачные вычисления и искусственный интеллект

Рассмотрим применимость различных моделей ОВ для ИИ. В целом ИИ может быть развернут в рамках любой из трех моделей предоставления услуги. Однако есть несколько нюансов, в случае SaaS поставщик помимо вычислительных ресурсов поставляет и ИИ, в данном случае ИИ также является услугой. Примером такой услуги является MLaaS [31]. В случае PaaS поставщик предоставляет пользователю, помимо

вычислительных мощностей, инструментарий по разработке ИИ. В данную категорию можно отнести различные облачные сервисы для разработки ПО. Ярким примером является Google Colab [32]. Несмотря на то что данное решение не является специализированным, оно все же предоставляет пользователю возможности по разработке и исследованию ИИ. При этом существуют решения, направленные именно на работу с ИИ [33]. IaaS и ИИ достаточно трудно выделить в специализированную категорию, так как в данном случае поставщик услуг предоставляет исключительно вычислительные ресурсы и инфраструктуру связывающую их. В данном случае специализация на ИИ достигается за счет двух факторов: за счет того, что аппаратное обеспечение технической составляющей ОВ подобрано таким образом, чтобы ИИ работал максимально эффективно; а также за счет самой позиции поставщика услуг [34]. В рассмотренных случаях ИИ выступает как объект, либо являющийся частью услуги, либо являющийся фактическим потребителем вычислительных мощностей, однако существуют работы, в которых рассматриваются и другие возможности по применению ИИ в ОВ [35–38]. Например, ИИ может применяться для построения инфраструктуры и балансировки нагрузки.

2.4 Конфиденциальность данных в облачных вычислениях

Если проекты, как в случаях [1, 2], являются открытыми, т. е. изначально не содержат конфиденциальной информации, то требования к их безопасности невелики. Однако любой ИИ может применяться в задачах, связанных с обработкой конфиденциальных данных. Это может быть медицинская информация [39–44], банковские [45–48], государственные [49, 50] и всевозможные личные данные. В данных случаях возникает ситуация, в связи с которой ОВ часто подвергаются критике как в обществе в целом, так и в научном сообществе. При обработке какой-либо информации в облаке к ней имеют доступ обе стороны – пользователь и поставщик услуги. В этом случае степень конфиденциальности данных устанавливается соглашением между пользователем и поставщиком, а то, как соблюдается данное соглашение находится в зоне ответственности поставщика услуг. Не редки случаи, когда конфиденциальные данные, хранимые или обрабатываемые различными поставщиками услуг, были скомпрометированы. В таком случае пользователь вынужденно отдает предпочтение поставщикам, которые либо зарекомендовали себя, как надежные с точки зрения обеспечения конфиденциальности обрабатываемых данных, либо утверждают о низкой вероятности компрометации за счет применения эффективных методов защиты. При этом возникает ряд вопросов. Какой способ обеспечения конфиденциальности является эффективным для хранения данных? Какие методы способны обеспечить необходимый уровень конфиденциальности в условиях обработки данных с использованием ИИ в целом и Больших данных в частности?

Целью данной работы является ответить на заданные вопросы и предоставить читателю возможность ознакомится с актуальными методами обеспечения конфиденциальности ОВ, использующих ИИ.

Для более наглядной демонстрации ретроспективного обзора распределенных вычислительных технологий и ИИ была составлена таблица (табл. 1). Рассмотренные выше работы представлены не в хронологическом порядке, табл. 1 призвана его восстановить. Отметим, что некоторые технологии впервые представлены намного раньше, чем они получили развитие и были изучены в полном объеме.

Анализируя данные табл. 1 и представленный выше обзор, можно заметить, что теоретические модели ИИ активно развивались с середины XX в., однако практические модели получили свое развитие в конце XX – начале XXI в. Так же из таблицы видно, что развитие практических моделей ИИ совпадает по времени с поздними этапами развития распределенных вычислений (до конца XX в.) и зарождением ОВ в начале XXI в. Данное совпадение не случайно и связано с тем, что развитие распределенных и облачных вычислений все это время шло по пути агрегирования все больших вычислительных мощностей, которых так не хватало большим моделям ИИ. Таким образом, справедливо утверждение, что распределенные вычисления внесли немалый вклад в развитие моделей ИИ, а ОВ позволяют развивать и создавать современные вычислительно затратные модели ИИ, сложность которых растет с каждым днем.

Учитывая вышесказанное, можно утверждать, что проблема конфиденциальности ОВ имеет такое же значение для ИИ, как для любого “потребителя” ОВ, коим ИИ по сути и является. Далее будут рассмотрены методы обеспечения конфиденциальности ИИ в ОВ.

Таблица 1. Историческая справка по рассмотренным технологиям

Год	Распределенные вычисления	Облачные технологии	Искусственный интеллект
1943	—	—	Концепция НС [51, 52]
1954	—	—	Зарождение генетических алгоритмов [53]
1959	—	—	Машинное обучение [54]
1962	Модель коллективных вычислений [55]	—	—
1966	—	—	Появление языковых моделей [56]
1978	Принципы распределения работы между процессорами [57]	—	—
1980		—	Теоретическое описание ГО [58]
1992	Зарождение GRID [59]	—	—
1996	Проект GIMPS по поиску целых чисел [60]	—	—
1999	Проект SETI на базе BOINC [61]	—	—
2000	—	—	Начало практического применения ГО [62]
			Компьютерное зрение [63]
2006	—	Зарождение концепции облачных вычислений [64]	—
2008	—	Определение концепции облачных вычислений как услуги [65]	—
2009	—	Запуск Google Apps [66]	—
2011	—	Стандартизация SaaS, PaaS и IaaS как моделей обслуживания в ОВ [25–27]	—
2015	—	Развитие туманных вычислений как основы для Интернета Вещей [67]	—
		Запуск OpenFog [68]	
2018	—	—	GPT [69]

3. СОВРЕМЕННЫЕ МЕТОДЫ И АЛГОРИТМЫ ОБЕСПЕЧЕНИЯ КОНФИДЕНЦИАЛЬНОСТИ ИИ

3.1. Дифференциальная конфиденциальность

Понятие конфиденциальности в некоторой степени является не строгим. В зависимости от ситуации она может рассматриваться с разных сторон. Например, с точки зрения медицины достаточно, чтобы данные были анонимны, т. е. история болезни была обезличена. В работе [70] проводится исследование методов ГО для обеспечения мягкой конфиденциальности. Авторы достигают своей цели путем внедрения дифференциальной конфиденциальности, т. е. путем перемешивания частных обезличенных данных вместе с синтезированными данными. Авторы продемонстрировали высокую точность данного метода. Однако данный метод имеет ряд недостатков: основой конфиденциальности является внесение дополнительного шума и расстояний между точками во время выполнения градиентного спуска. Это не только повышает вычислительную сложность расчетов, но и не обеспечивает надлежащей безопасности данных, так как если злоумышленник получит доступ к данным во время обработки – он сможет убрать лишний шум.

С другой стороны, в работе [71] рассмотрен принципиально отличающийся от вышеприведенного подход к обеспечению конфиденциальности данных. Авторы рассматривают возможность обучения и использования нейронной сети группой пользователей без необходимости раскрытия конфиденциальных данных друг другу. Подобная возможность обеспечивается особенностью стохастического градиентного спуска, которая позволяет выполнять его параллельно и асинхронно. Авторы также утверждают, что их решение позволяет участникам обучать нейронную сеть независимо на своих собственных (конфиденциальных) наборах, делясь подмножествами ключевых параметров. В целом, данный метод позволяет обеспечить конфиденциальность среди участников группы, однако, оказывается неэффективным в случае предварительного сговора участников либо внешней атаки.

3.2. Схемы разделения секрета

В контексте построения конфиденциального ИИ в ОВ необходимо также рассмотреть схемы разделения секрета (СРС) [4]. Приведем кратко основные теоретические аспекты, связанные с СРС. Секрет S разделяется дилером D между N участниками таким образом, что для дешифровки какой-либо информации потребуется объединение долей S_i всех участников схемы обратно в секрет, $i \in 1,2,\ldots,N$, где N – количество участников. Если секретом является, например, ключ для схемы шифрования, конфиденциальность системы возрастает.

Существуют два типа схем разделения секрета: полные [4] и пороговые [72]. Полные предполагают, что для восстановления необходимы все доли секрета, пороговые же, что необходимо определенное количество долей, но не все. Пороговые схемы разделения секрета на практике используются гораздо чаще ввиду своей гибкости. Пороговые СРС позволяют снизить вычислительную нагрузку, при этом порог задается такой, чтобы злоумышленник не смог овладеть необходимым количеством долей секрета, либо чтобы не смог состояться предварительный сговор необходимого количества участников.

В работе [73] авторы предлагают распределенное ГО, когда участники обучают нейронную сеть с помощью своих конфиденциальных наборов. Авторы приводят результаты исследования, в котором удалось сохранить конфиденциальность при распределенном ГО в облаке с недоверенными участниками. Работа демонстрирует функционал, предоставляемый СРС при работе с ИИ, и сам факт возможности конфиденциального обучения. Авторами применяется схема Шамира [72]. За время своего существования схема доказала свою пригодность с точки зрения безопасности. Однако если мы рассматриваем применение методов СРС с точки зрения ОВ, то на проблемы безопасности также накладываются проблемы надежности, дополнительные корректирующие коды накладывают на систему дополнительные нагрузки. Для нивелирования необходимости применения дополнительных корректирующих кодов можно применять СРС, основанные на системе остаточных классов (СОК) [74], например, такие как СРС Асмута–Блума [75] и СРС Миньотта [76].

В работе [77] рассмотрена организация федеративного обучения нейронных сетей в облачных системах с применением СРС Асмута–Блума для обеспечения конфиденциальности [78]. Также авторами рассматривается СРС Миньотта, однако существуют работы доказывающие ее непригодность для обеспечения безопасности.

В работе [79] авторы предлагают протокол конфиденциальной передачи данных в условиях работы с нейронными сетями. Протокол передачи основан на СРС. Основной упор в статье сделан на скорость обработки данных при сохранении их конфиденциальности. Авторами показана эффективность их решения по сравнению с некоторыми методами гомоморфного шифрования, которое также применимо при построении конфиденциального ИИ в ОВ.

3.3. Гомоморфное шифрование

Первые работы по гомоморфному шифрованию (ГШ) появились несколько десятилетий назад. ГШ допускает обработку зашифрованных данных без необходимости проведения операции дешифрования. Гомоморфные схемы шифрования бывают гомоморфными по сложению, умножению или одновременно по обеим перечисленным арифметическим операциям. Такая особенность характерна для различных ассиметричных шифров. Так, например, гомоморфное сложение поддерживается схемами, представленными в работах [80, 81], а гомоморфное умножение схемами из работ [82, 83]. Стоит отметить, что применение таким схемам шифрования находится и сейчас, например, в работе [84] авторы применяют модифицированную схему Эль-Гамаля для обучения нейронной сети, основываясь на том факте, что схема Эль-Гамаля гомоморфна по умножению. Авторы используют линейную аппроксимацию функции активации. Данный

метод является довольно узконаправленным и трудномасштабируем.

В 2009 г. Джентри разработал полностью гомоморфную схему шифрования (ПГШ) [85]. Данная схема была не достаточно вычислительно эффективна, однако дальнейшие исследования его последователей позволили повысить эффективность до уровня, достаточного для реального практического применения ПГШ [86–91]. ПГШ позволяет выполнять как гомоморфное сложение, так и умножение (однако имеет ограничение на количество умножений с одним ключом), кроме того, существуют схемы, которые позволяют выполнять полиномиальные операции над шифротекстами [92].

Такой набор операций позволяет реализовать большое количество алгоритмов ИИ, что было достаточно быстро обнаружено многими исследователями ИИ. Рассмотрим разработанные ими методы.

CryptoNets [93]. В данном решении авторы предлагают применение ГШ в работе конфиденциального облачного хранилища с возможностью обработки нейронными сетями. Основным достигнутым результатом является точность работы нейронной сети по распознаванию образов, равная 99%.

В случае SEALion [94] предлагается решение, основанное на TensorFlow [95] и SEAL [96]. TensorFlow реализует вычисления в тензорах, SEAL само ГШ. С помощью разработанного решения авторы реализуют сверточные нейронные сети (СНС), в качестве метода обучения применяется метод опорных векторов. В работе представлены несколько моделей НС, точность определяется на основе качества распознавания изображений цифр. Авторы показывают, что их решения более эффективны чем, например, в работе [93]. Похожим решением является TenSEAL [97], эта библиотека предоставляет возможности по работе с МО, НС и СНС совместно с ГШ. В качестве недостатка можно выделить сложности при обучении на зашифрованных данных, при этом обученная НС демонстрирует быстродействие и дает точные результаты.

В работе [98] рассматриваются вариации конфиденциальных СНС с ГШ. Авторами предлагается СНС, в которой применяется целочисленная схема ПГШ – BGV [99]. Основной акцент сделан на возможности вычисления значения функции активации ReLU от шифротекстов, полученных согласно схеме BGV. Авторами было определенно три требования для НС – точность, конфиденциальность и эффективность, результатом работы является полиномиально приближенная функция ReLU, которая позволила уменьшить мультипликативную глубину, тем самым повысив эффективность работы сети. Данную работу можно считать достаточно важной по нескольким причинам: во-первых, показан сам факт возможности полностью конфиденциальных расчетов; во-вторых, показана возможность построения и применения нейронных сетей с ПГШ.

В работе [100] рассматривается сохраняющее конфиденциальность ГО. Авторы противопоставляют свои исследования многим работам, которые были рассмотрены ранее, сосредотачиваясь на устранении недостатков, которые были озвучены выше. Основной результат основан на применении методов аппроксимации для различных функций активации, например, таких как ReLU и Softmax. Для обеспечения большого количества гомоморфных умножений в схеме CKKS авторы предлагают свою модификацию процедуры bootstrapping, которая основана на том факте, что схема CKKS [92] применяет систему остаточных классов для ускорения арифметических вычислений. Этот результат интересен тем, что повышает быстродействие методов ГШ для конкретных задач, тогда как повышение быстродействия и применимости ГШ в целом является основным направлением исследований в этой области в настоящее время.

ГШ позволяет решить проблемы безопасности ОВ и ОТ в целом, создав прозрачную среду для свободных и конфиденциальных вычислений в облаках. Проведенный анализ методов ГШ в ИИ, позволяет заявить о том, что такого мнения придерживаются многие исследователи. Выше рассмотрены схемы ПГШ, которые в большей степени основаны на криптографических решетках, однако существуют схемы ПГШ, которые построены на СРС. Такой подход позволяет расширить возможности ПГШ в ОТ. Далее рассмотрены гибридные модели, которые совмещают методы ПГШ и СРС.

3.4. Гибридные системы гомоморфного шифрования и схем разделения секрета

Учитывая результаты в направлении повышения конфиденциальности ИИ, полученные за счет применения СРС и ПГШ, перспективной выглядит идея создания гибридной системы ПГШ–СРС. Такая гибридная система позволит усилить безопасность СРС за счет обработки информации в зашифрованном виде с помощью ПГШ в пространстве одного локального узла. Для передачи данных между узлами применяются алгоритмы безопасности на основе СРС. Таким

образом появляется возможность регулировать баланс, смещая акцент в сторону безопасности либо производительности.

В работе [101] предложено решение, в котором используются СРС и ПГШ для реализации сохраняющей конфиденциальность НС. Авторы описывают различные алгоритмы, применяемые в данной системе, например, алгоритмы безопасности, генерации ключей, работы СРС и т. п., а также подробно разбирают, как происходит работа НС при таком подходе к обеспечению безопасности. Основной упор делается на демонстрации эффективности, безопасности, точности полученной НС и ее сравнении с другими решениями, однако полученные данные являются теоретическими. Рассмотренная работа замечательна тем, что несмотря на отсутствие конкретики в плане применяемых ПГШ и СРС, она показывает саму возможность реализации такой системы и описывает ее характеристики, побуждая других исследователей заниматься разработками в данной сфере.

4. АНАЛИЗ ПОЛУЧЕННЫХ РЕЗУЛЬТАТОВ

В ходе исследования были проанализированы три группы методов обеспечения конфиденциальности искусственного интеллекта (табл. 2). Представленные ниже данные основаны на информации, представленной в обозреваемых работах.

Анализируя полученные результаты (табл. 2), можно подвести следующий итог.

Методы дифференциальной конфиденциальности преимущественно основаны на изменениях обучающей выборки. Предлагается добавление лишних шумов в данные, за счет которого возрастает степень конфиденциальности, одновременно с которой серьезно снижается эффективность расчетов. В строках 1 и 2 жирным шрифтов выделены значения вычислительной сложности. В первом случае ε – означает добавочный шум. Во втором случае δ – количество асинхронных узлов. Учитывая достигаемый при этом уровень конфиденциальности, при передаче данных требуется применение дополнительных мер защиты. При обработке степень конфиденциальности средняя – для кражи данных злоумышленнику необходимо будет либо отфильтровать информацию от шума, либо взять под контроль несколько узлов, чтобы отследить поток данных. Кроме того, дополнительное влияние на эффективность системы окажет необходимость применения корректирующих кодов во избежание возникновения коллизий либо потери данных.

В случае СРС схемы, основанные на СОК, позволяют повысить надежность системы за счет

Таблица 2. Результаты аналитического обзора методов

№	Метод	Вычислительная сложность	Конфиденциальность при передаче данных	Конфиденциальность при обработке данных	Обеспечение надежности	Рассмотренный метод ИИ
1	Модифицированный градиентный спуск	$\boldsymbol{O\left(D^2N+(D+\varepsilon)^3\right)}$	Низкая	Средняя	Отсутствует	AC-GAN [70]
2	Асинхронный градиентный спуск	$\boldsymbol{O\left(D^2(N\delta)+D^3\right)}$	Низкая	Средняя	Отсутствует	СНС [71]
3	СРС Шамира	$O(N^2)$	Высокая	Низкая	Отсутствует	ГО [73]
4	СРС Асмута–Блума	$O\left(\log_2^2(N)+\log_2(N^2)\right)$	Высокая	Низкая	Корректирующие коды СОК	Федеративное обучение [77, 79]
5	СРС Миньота	$O(N^2)$	Низкая	Низкая	Корректирующие коды СОК	Федеративное обучение [77]
6	BFV	$O(N\log N)$	Высокая	Высокая	Корректирующие коды СОК (в теории)	НС, СНС [93, 94]
7	BGV	$O(N\log N)$	Высокая	Высокая	Отсутствует	НС, СНС [98]
8	СККС	$O(N\log N)$	Высокая	Высокая	Корректирующие коды СОК (в теории)	НС, СНС, ГО [93, 94, 97, 100]
9	Гибрид СРС–ПГШ	$O(N\log N)$	*Высокая*	*Высокая*	*Корректирующие коды СОК*	*НС*, *СНС*, *ГО* [101]

использования самокорректирующих свойств СОК. Однако конфиденциальность во время обработки данных необходимо обеспечивать дополнительными методами шифрования. Также, в данном случае, для перехвата данных злоумышленнику потребуется взять под контроль пороговое количество узлов. При построении СРС порог для восстановления секрета рассчитывается так, чтобы время, затраченное на компрометацию долей секрета, было большим, чем время актуальности данных.

Обеспечить полную конфиденциальность позволяет ПГШ. Однако оно имеет существенный недостаток. Несмотря на то, что вычислительная сложность шифрования является невысокой, по сравнению с другими алгоритмами, вычислительная сложность обработки данных – намного выше, чего не скрывают авторы схем.

Последней из рассмотренных категорий были гибридные СРС–ПГШ-методы. В настоящий момент они представлены лишь теоретическими моделями, однако, анализируя их, можно оценить возможные характеристики системы.

Проведенное исследование показало, что разные исследовательские группы заинтересованы в разработке ИИ обладающего высоким уровнем конфиденциальности. Однако в настоящий момент релевантных методов достижения подобного уровня не существует. В дальнейшем планируется исследование наиболее перспективных с точки зрения обеспечения конфиденциальности методов, основанных на СРС–ПГШ. В частности планируется разработка НС, использующей схему ПГШ CKKS, а также СРС Асмута–Блума. Обосновывая выбор именно этих алгоритмов, можно заметить, что среди множества схем ПГШ именно CKKS представляет для исследователей наибольший интерес, а среди множества СРС, основанных на СОК, именно СРС Асмута–Блума имеет наилучшие характеристики с точки зрения безопасности. Использование СОК в CKKS позволит повысить эффективность решения, а в СРС Асмута–Блума – надежность обработки данных. Важной частью исследования будет являться определение точности результатов вычислений, ввиду целочисленной природы СОК, а также ошибки приближения в CKKS.

5. ЗАКЛЮЧЕНИЕ

В данной работе были исследованы методы построения, сохраняющего конфиденциальность ИИ в ОВ. На первом этапе исследования проведен аналитический обзор как методов ИИ, так и методов ОВ. По результатам обзора были определены критерии безопасности ИИ в ОВ. Далее был выполнен второй этап аналитического обзора, а именно обзор методов обеспечения конфиденциальности ИИ, на основе которого было выделено четыре группы методов:

- дифференциальная конфиденциальность;
- схемы разделения секрета;
- гомоморфное шифрование;
- гибридные методы, основанные на СРС–ПГШ.

На основании результатов проведенного исследования были определенны положительные и отрицательные стороны рассмотренных методов, сформировано представление о современном состоянии проблемы обеспечения конфиденциальности ИИ в ОВ, а также обозначены подходы к ее решению. Аналитический обзор показал, что в настоящий момент релевантного решения не существует. Решения, обеспечивающие наиболее высокий уровень конфиденциальности, имеют низкую эффективность ввиду необходимости выполнения сложных вычислений. ПГШ поддерживает операции сложения и умножения над зашифрованными значениями, такие операции как определение знака числа, деление, матричные операции – не реализованы в полном объеме. СРС не позволяют обрабатывать информацию в зашифрованном виде. Конфиденциальность обеспечивается за счет неразглашения источника части набора данных одного участника перед другими. Дифференциальная конфиденциальность обеспечивает анонимность во время работы нейронной сети, при этом не имеет защиты от перехвата данных. В качестве возможного решения теоретически установлен гибридный метод, основанный на СРС–ПГШ, однако он требует детального исследования.

В будущих работах будет проведено исследование СРС–ПГШ, а именно разработка прототипа и исследование его характеристик.

ИСТОЧНИК ФИНАНСИРОВАНИЯ

Работа выполнена при поддержке Российского научного фонда 19-71-10033, https://rscf.ru/project/19-71-10033/.

СПИСОК ЛИТЕРАТУРЫ

1. *Brown T. et al.* Language models are few-shot learners // Advances in neural information processing systems. 2020. V. 33. P. 1877–1901.
2. OpenAI, GPT-4 Technical Report. arXiv, 27 март 2023 г. https://doi.org/10.48550/arXiv.2303.08774

3. *Douligeris C., Mitrokotsa A.* DDoS attacks and defense mechanisms: classification and state-of-the-art // Computer networks. 2004. V. 44. № 5. P. 643–666.
4. *Beimel A.* Secret-Sharing Schemes: A Survey // Coding and Cryptology, Y.M. Chee, Z. Guo, S. Ling, F. Shao, Y. Tang, H. Wang, and C. Xing, Eds., in Lecture Notes in Computer Science. Berlin, Heidelberg: Springer. 2011. P. 11–46. https://doi.org/10.1007/978-3-642-20901-7_2
5. *Mahesh B.* Machine learning algorithms-a review // International Journal of Science and Research (IJSR). [Internet]. 2020. V. 9. № 1. P. 381–386.
6. *Kaelbling L.P., Littman M.L., Moore A.W.* Reinforcement learning: A survey // Journal of artificial intelligence research. 1996. V. 4. P. 237–285.
7. *Srinivas M., Patnaik L.M.* Genetic algorithms: A survey // Computer. 1994. V. 27. № 6. P. 17–26.
8. *Spragins J.* Learning without a teacher // IEEE Transactions on Information Theory. 1996. V. 12. № 2. P. 223–230.
9. *Liu B.* Supervised Learning // Web Data Mining. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011. P. 63–132. https://doi.org/10.1007/978-3-642-19460-3_3
10. *Wang S.-C.* Artificial Neural Network // Interdisciplinary Computing in Java Programming. Boston, MA: Springer US, 2003. P. 81–100. https://doi.org/10.1007/978-1-4615-0377-4_5
11. *Park H., Kim S.* Chapter Three – Hardware accelerator systems for artificial intelligence and machine learning // Advances in Computers. V. 122, S. Kim and G.C. Deka, Eds., in Hardware Accelerator Systems for Artificial Intelligence and Machine Learning. V. 122. Elsevier, 2021. P. 51–95. https://doi.org/10.1016/bs.adcom.2020.11.005
12. *Hwang D. H., Han C.Y., Oh H.W., Lee S.E.* ASimOV: A Framework for Simulation and Optimization of an Embedded AI Accelerator // Micromachines. 2021. V. 12. № 7. https://doi.org/10.3390/mi12070838
13. *Mishra A., Yadav P., Kim S.* Artificial Intelligence Accelerators // Artificial Intelligence and Hardware Accelerators, A. Mishra, J. Cha, H. Park, and S. Kim, Eds. Cham: Springer International Publishing, 2023. P. 1–52. https://doi.org/10.1007/978-3-031-22170-5_1
14. *Carminati M., Scandurra G.* Impact and trends in embedding field programmable gate arrays and microcontrollers in scientific instrumentation // Review of Scientific Instruments. 2021. V. 92.№ 9. https://pubs.aip.org/aip/rsi/article-abstract/92/9/091501/1030652
15. *Shawash J., Selviah D.R.* Real-time nonlinear parameter estimation using the Levenberg–Marquardt algorithm on field programmable gate arrays // IEEE Transactions on industrial electronics. 2012. V. 60. № 1. P. 170–176.
16. *Ruiz-Rosero J., Ramirez-Gonzalez G., Khanna R.* Field programmable gate array applications – A scientometric review // Computation. 2019. V. 7. № 4. P. 63.
17. *Mellit A., Kalogirou S.A.* MPPT-based artificial intelligence techniques for photovoltaic systems and its implementation into field programmable gate array chips: Review of current status and future perspectives // Energy. 2014. V. 70. P. 1–21.
18. *Goodfellow I., Bengio Y., Courville A.* Deep learning. MIT press, 2016. https://books.google.com/books?hl=ru&lr=&id=omivDQAAQBAJ&oi=fnd&pg=PR5&dq=Deep+Learning&ots=MNV5aolzSS&sig=waXAS6C-_v-48H2qbW9rMFkEhFY
19. *Bouvrie J.* Notes on convolutional neural networks. 2006. http://web.mit.edu/jvb/www/papers/cnn_tutorial.pdf
20. *Rawat W., Wang Z.* Deep convolutional neural networks for image classification: A comprehensive review // Neural computation. 2017. V. 29; № 9. P. 2352–2449.
21. *Needham R.M., Herbert A.J.* The Cambridge distributed computing system, 1983.
22. *Adiga N.R. et al.* An overview of the BlueGene/L supercomputer // SC'02: Proceedings of the 2002 ACM/IEEE Conference on Supercomputing, IEEE, 2002. P. 60–60. https://ieeexplore.ieee.org/abstract/document/1592896/
23. *Jacob B., Brown M., Fukui K., Trivedi N.* Introduction to grid computing // IBM redbooks, 2005. P. 3–6.
24. *Foster I., Zhao Y., Raicu I., Lu S.* Cloud computing and grid computing 360-degree compared // 2008 grid computing environments workshop, IEEE, 2008. P. 1–10. https://ieeexplore.ieee.org/abstract/document/4738445/?casa_token=TbNOHOEaljQAAAAA:j6MuEJKmrGL8iCvH-HzRnmI2k5UKn5y1w7hC4MNJanJXZPfiBC_XKLoTFsCImP1RYzyKfRKiCE0
25. *Cusumano M.* Cloud computing and SaaS as new computing platforms // Commun. ACM, April, 2010. V. 53. № 4. P. 27–29. https://doi.org/10.1145/1721654.1721667
26. *Rodero-Merino L., Vaquero L.M., Caron E., Muresan A., Desprez F.* Building safe PaaS clouds: A survey on security in multitenant software platforms // Computers & security. 2012. V. 31. № 1. P. 96–108.
27. *Bhardwaj S., Jain L., Jain S.* Cloud computing: A study of infrastructure as a service (IAAS) // International Journal of engineering and information Technology. 2010. V. 2. № 1. P. 60–63.
28. *Manvi S.S., Shyam G.K.* Resource management for Infrastructure as a Service (IaaS) in cloud computing: A survey // Journal of network and computer applications. 2014. V. 41. P. 424–440.
29. *Lehner W., Sattler K.-U.* Database as a service (DBaaS) // 2010 IEEE 26th International Conference on Data Engineering (ICDE2010), IEEE, 2010. P. 1216–1217.

https://ieeexplore.ieee.org/abstract/document/5447723/?casa_token=uaXogPZV0C0AAAAA:4Dg_40-GvhUsuHXFKUOgxZ_ZyGlCOqjcZtpRo-K6UosB-k-_Wh5wAmJIBtHYRE9OLXZ1xwVKuLAE

30. *Meng S., Liu L.* Enhanced monitoring-as-a-service for effective cloud management // IEEE Transactions on Computers. 2012. V. 62. № 9. P. 1705–1720.
31. *Weng Q. et. al.* {MLaaS} in the wild: Workload analysis and scheduling in {Large-Scale} heterogeneous {GPU} clusters // 19th USENIX Symposium on Networked Systems Design and Implementation (NSDI 22), 2022. P. 945–960. https://www.usenix.org/conference/nsdi22/presentation/weng
32. *Bisong E.* Google Colaboratory // Building Machine Learning and Deep Learning Models on Google Cloud Platform. Berkeley, CA: Apress, 2019. P. 59–64. https://doi.org/10.1007/978-1-4842-4470-8_7
33. H2O AI Cloud. https://h2o.ai/platform/ai-cloud/
34. NVIDIA NGC | NVIDIA. https://www.nvidia.com/en-us/gpu-cloud/
35. *Tang J.* Artificial intelligence-based e-commerce platform based on SaaS and neural networks // 2020 Fourth International Conference on Inventive Systems and Control (ICISC). IEEE, 2020. P. 421–424. https://ieeexplore.ieee.org/abstract/document/9171193/?casa_token=TmYwFdLDXq0AAAAA:8P5V-VcZS_KWCXEnEm8xk2RPMV5kfWF27K9S9O9Z-5fYh273EkseT7j0Jf7jZYAMOnPUX0l-5sCbs
36. *Yathiraju N.* Investigating the use of an Artificial Intelligence Model in an ERP Cloud-Based System // International Journal of Electrical, Electronics and Computers. 2022. V. 7. № 2. P. 1–26.
37. *Mishra S., Tripathi A.R.* AI business model: an integrative business approach // J. Innov. Entrep. Dec. 2021. V. 10. № 1. P. 18. https://doi.org/10.1186/s13731-021-00157-5
38. *Mishra D., Shekhar S.* Artificial Intelligence Candidate Recruitment System using Software as a Service (SaaS) Architecture // International Research Journal of Engineering and Technology. 2018. V. 05. № 05. P. 3804–3808.
39. *Cadario R., Longoni C., Morewedge C.K.* Understanding, explaining, and utilizing medical artificial intelligence // Nature human behaviour. 2021. V. 5. № 12. P. 1636–1642.
40. *Kim M., Song Y., Wang S., Xia Y., Xiang X.* Secure logistic regression based on homomorphic encryption: Design and evaluation // JMIR medical informatics. 2018. V. 6. № 2. P. e8805.
41. *Klonoff D.C.* Fog computing and edge computing architectures for processing data from diabetes devices connected to the medical internet of things // Journal of diabetes science and technology. 2017. V. 11. № 4. P. 647–652.
42. *Kocabas O., Soyata T.* Utilizing homomorphic encryption to implement secure and private medical cloud computing // 2015 IEEE8th International Conference on Cloud Computing. IEEE, 2015. P. 540–547.
43. *Liu R., Rong Y., Peng Z.* A review of medical artificial intelligence // Global Health Journal. 2020. V. 4. № 2. P. 42–45.
44. *Sun X., Zhang P., Sookhak M., Yu J., Xie W.* Utilizing fully homomorphic encryption to implement secure medical computation in smart cities // Personal and Ubiquitous Computing. 2017. V. 21. № 5. P. 831–839.
45. *Kaya O., Schildbach J., AG D.B., Schneider S.* Artificial intelligence in banking // Artificial intelligence. 2019. https://www.dbresearch.com/PROD/RPS_EN-PROD/PROD0000000000495172/Artificial_intelligence_in_banking%3A_A_lever_for_pr.pdf
46. *Rahman M., Ming T.H., Baigh T.A., Sarker M.* Adoption of artificial intelligence in banking services: an empirical analysis // International Journal of Emerging Markets. 2021. https://www.emerald.com/insight/content/doi/10.1108/IJOEM-06-2020-0724/full/html
47. *Sadok H., Sakka F., El Maknouzi M.E.H.* Artificial intelligence and bank credit analysis: A review // Cogent Economics & Finance. Dec. 2022. V. 10. № 1. P. 2023262. https://doi.org/10.1080/23322039.2021.2023262
48. *Smith A., Nobanee H.* Artificial intelligence: in banking A mini-review // Available at SSRN3539171, 2020. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3539171
49. *Reis J., Santo P.E., Melão N.* Artificial Intelligence in Government Services: A Systematic Literature Review // New Knowledge in Information Systems and Technologies. V. 930. Á. Rocha, H. Adeli, L.P. Reis, and S. Costanzo, Eds., in Advances in Intelligent Systems and Computing. V. 930. Cham: Springer International Publishing. 2019. P. 241–252. https://doi.org/10.1007/978-3-030-16181-1_23
50. *Valle-Cruz D., Alejandro Ruvalcaba-Gomez E., Sandoval-Almazan R., Ignacio Criado J.* A Review of Artificial Intelligence in Government and its Potential from a Public Policy Perspective // Proceedings of the 20th Annual International Conference on Digital Government Research. Dubai United Arab Emirates: ACM, June 2019. P. 91–99. https://doi.org/10.1145/3325112.3325242
51. *Pitts W.* The linear theory of neuron networks: The dynamic problem // The bulletin of mathematical biophysics. 1943. V. 5. P. 23–31.
52. *Khare S.S., Gajbhiye A.R.* Literature Review on Application of Artificial Neural Network (Ann) In Operation of Reservoirs // International Journal of computational Engineering research (IJCER). June 2013. V. 3. № 6. P. 63.
53. *Seesing A.* Evotest: Test case generation using genetic programming and software analysis // Operations Research. 1954. V. 2. P. 393–410.

54. *Samuel A.L.* Machine learning // The Technology Review. 1959. V. 62. № 1. P. 42–45.
55. *Evreinov È.V., Kosarev I.* Однородные универсальные вычислительные системы высокой производительности (No Title), 1966. https://cir.nii.ac.jp/crid/1130282272859765760
56. *Gold E.M.* Language identification in the limit // Information and control. 1967. V. 10. № 5. P. 447–474.
57. *Глушков В.М.* Вычислительная система, 1996. https://elibrary.ru/item.asp?id=41074434
58. *Huang X.* Deep-learning based climate downscaling using the super-resolution method, 1981. https://pdfs.semanticscholar.org/cf5c/3b29559ababba5a889444632e1c91d6b78fc.pdf
59. *Smarr L., Catlett C.E.* Metacomputing // Grid Computing, 1st ed., F. Berman, G. Fox, and T. Hey, Eds., Wiley, 2003. P. 825–835. https://doi.org/10.1002/0470867167.ch37
60. *Buske D., Keith S.* GIMPS Finds Another Prime! // Math Horizons. April 2000. V. 7. № 4. P. 19–21. https://doi.org/10.1080/10724117.2000.11975124
61. *Anderson D.P.* Boinc: A system for public-resource computing and storage // Fifth IEEE/ACM international workshop on grid computing. IEEE, 2004. P. 4–10. https://ieeexplore.ieee.org/abstract/document/1382809/?casa_token=cjAKtADFAKwAAAAA:-WGH_xmovZAUi-kr_PA-h3nXtuizBL829DPFlC0B6pbcCoApRKDCZLwFWxzfYdT0WauFC5c6EQw1
62. *Du T., Shanker V.* Deep learning for natural language processing // Eecis. Udel. Edu, 2009. P. 1–7.
63. *Davies E.R.* Machine vision: theory, algorithms, practicalities. Elsevier, 2004. https://books.google.com/books?hl=ru&lr=&id=uY-Z3vORugwC&oi=fnd&pg=PP1&dq=Machine+Vision+:+Theory,+Algorithms,+Practicalities&ots=QOl9U9_MBf&sig=w0poN6d3IGeXs4oacagO4MlnxYs
64. *Mell P., Grance T.* The NIST Definition of Cloud Computing // National Institute of Standards and Technology Special Publication. 2011. V. 53. P. 1–7.
65. *Finkelstein R.* Analyzing Trend of Cloud Computing and it's Enablers using Gartner Strategic Technology, 2004. https://www.researchgate.net/profile/Amol-Adamuthe/publication/308747055_Analyzing_Trend_of_Cloud_Computing_and_it's_Enablers_using_Gartner_Strategic_Technology/links/59a929d3a6fdcc2398414d6f/Analyzing-Trend-of-Cloud-Computing-and-its-Enablers-using-Gartner-Strategic-Technology.pdf
66. A history of cloud computing // Computer Weekly. https://www.computerweekly.com/feature/A-history-of-cloud-computing
67. *Dolui K., Datta S.K.* Comparison of edge computing implementations: Fog computing, cloudlet and mobile edge computing // 2017 Global Internet of Things Summit (GIoTS), IEEE. 2017. P. 1–6.
68. OpenFog, OPC Foundation. https://opcfoundation.org/markets-collaboration/openfog/
69. *Radford A., Narasimhan K., Salimans T., Sutskever I.* Improving language understanding by generative pre-training" 2018. https://www.mikecaptain.com/resources/pdf/GPT-1.pdf
70. *Beaulieu-Jones B.K. et al.* Privacy-Preserving Generative Deep Neural Networks Support Clinical Data Sharing // Circ: Cardiovascular Quality and Outcomes. Jul. 2019. V. 12. № 7. P. e005122. https://doi.org/10.1161/CIRCOUTCOMES.118.005122
71. *Shokri R., Shmatikov V.* Privacy-Preserving Deep Learning // Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security. Denver Colorado USA: ACM, Oct. 2015. P. 1310–1321. https://doi.org/10.1145/2810103.2813687.
72. *Shamir A.* How to share a secret // Communications of the ACM. 1979. V. 22. № 11. P. 612–613.
73. *Duan J., Zhou J., Li Y.* Privacy-preserving distributed deep learning based on secret sharing // Information Sciences. 2020. V. 527. P. 108–127.
74. *Akushsky I.A., Yuditsky D.I.* Modular arithmetic in residue classes // Soviet Radio, 1968.
75. *Asmuth C., Bloom J.* A modular approach to key safeguarding // IEEE transactions on information theory. 1983. V. 29. № 2. P. 208–210.
76. *Mignotte M.* How to share a secret // Workshop on cryptography. Springer, 1982. P. 371–375.
77. *Tian T., Wang S., Xiong J., Bi R., Zhou Z., Bhuiyan M.Z.A.* Robust and privacy-preserving decentralized deep federated learning training: Focusing on digital healthcare applications // IEEE/ACM Transactions on Computational Biology and Bioinformatics, 2023. https://ieeexplore.ieee.org/abstract/document/10058838/
78. *Barzu M., Ţiplea F.L., Drăgan C.C.* Compact sequences of co-primes and their applications to the security of CRT-based threshold schemes // Information Sciences. 2013. V. 240. P. 161–172.
79. *Ge Z., Zhou Z., Guo D., Li Q.* Practical Two-party Privacy-preserving Neural Network Based on Secret Sharing. http://arxiv.org/abs/2104.04709
80. *Paillier P.* Public-Key Cryptosystems Based on Composite Degree Residuosity Classes // Advances in Cryptology – EUROCRYPT '99. V. 1592, J. Stern, Ed., in Lecture Notes in Computer Science. V. 1592. Berlin, Heidelberg: Springer Berlin Heidelberg, 1999. P. 223–238. https://doi.org/10.1007/3–540–48910-X_16.
81. *Benaloh J.* Dense probabilistic encryption // Proceedings of the workshop on selected areas of cryptography, 1994. P. 120–128. https://sacworkshop.org/proc/SAC_94_006.pdf

82. *Rivest R. L., Shamir A., Adleman L.* A method for obtaining digital signatures and public-key cryptosystems // Commun. ACM. Feb. 1978. V. 21. № 2. P. 120–126. https://doi.org/10.1145/359340.359342
83. *ElGamal T.* A public key cryptosystem and a signature scheme based on discrete logarithms // IEEE transactions on information theory. 1985. V. 31. № 4. P. 469–472.
84. *Chen T., Zhong S.* Privacy-preserving backpropagation neural network learning // IEEE Transactions on Neural Networks. 2009. V. 20. № 10. P. 1554–1564.
85. *Gentry C.* A fully homomorphic encryption scheme // Stanford university, 2009.
86. *Gentry C.* Computing arbitrary functions of encrypted data // Communications of the ACM. 2010. V. 53. № 3. P. 97–105.
87. *Gentry C., Halevi S.* Implementing gentry's fully-homomorphic encryption scheme // Advances in Cryptology–EUROCRYPT 2011: 30th Annual International Conference on the Theory and Applications of Cryptographic Techniques. Tallinn, Estonia, May 15–19, 2011. Proceedings 30, Springer, 2011. P. 129–148.
88. *Gentry C., Halevi S., Peikert C., Smart N.P.* Ring Switching in BGV-Style Homomorphic Encryption // Security and Cryptography for Networks. V. 7485. I. Visconti and R. De Prisco, Eds. Lecture Notes in Computer Science. V. 7485. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012. P. 19–37. https://doi.org/10.1007/978-3-642-32928-9_2
89. *Gentry C., Sahai A., Waters B.* Homomorphic encryption from learning with errors: Conceptually-simpler, asymptotically-faster, attribute-based // Annual Cryptology Conference. Springer, 2013. P. 75–92.
90. *van Dijk M., Gentry C., Halevi S., Vaikuntanathan V.V.* Fully homomorphic encryption over the integers // Annual international conference on the theory and applications of cryptographic techniques. Springer, 2010. P. 24–43.
91. *van Dijk M., Gentry C., Halevi S., Vaikuntanathan V.* Fully Homomorphic Encryption over the Integers // Advances in Cryptology – EUROCRYPT 2010. V. 6110. H. Gilbert, Ed., Lecture Notes in Computer Science. V. 6110. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010. P. 24–43. https://doi.org/10.1007/978-3-642-13190-5_2
92. *Cheon J. H., Kim A., Kim M., Song Y.* Homomorphic encryption for arithmetic of approximate numbers // International conference on the theory and application of cryptology and information security. Springer, 2017. P. 409–437.
93. *Gilad-Bachrach R., Dowlin N., Laine K., Lauter K., Naehrig M., Wernsing J.* Cryptonets: Applying neural networks to encrypted data with high throughput and accuracy // International conference on machine learning, PMLR, 2016. P. 201–210. https://proceedings.mlr.press/v48/gilad-bachrach16.html
94. *van Elsloo T., Patrini G., Ivey-Law H.* SEALion: a Framework for Neural Network Inference on Encrypted Data. http://arxiv.org/abs/1904.12840
95. TensorFlow. https://www.tensorflow.org/?hl=ru
96. Microsoft SEAL. Microsoft. https://github.com/microsoft/SEAL
97. *Benaissa A., Retiat B., Cebere B., Belfedhal A.E.* TenSEAL: A Library for Encrypted Tensor Operations Using Homomorphic Encryption. http://arxiv.org/abs/2104.03152
98. *Chabanne H., De Wargny A., Milgram J., Morel C., Prouff E.* Privacy-preserving classification on deep neural network // Cryptology ePrint Archive, 2017. https://eprint.iacr.org/2017/035
99. *Brakerski Z., Gentry C., Vaikuntanathan V.* (Leveled) fully homomorphic encryption without bootstrapping // ACM Transactions on Computation Theory (TOCT). 2014. V. 6. № 3. P. 1–36.
100. *Lee J.-W. et al.* Privacy-preserving machine learning with fully homomorphic encryption for deep neural network // IEEE Access. 2022. V. 10. P. 30039–30054.
101. *Ryffel T., Tholoniat P., Pointcheval D., Bach F.* ARIANN: Low-Interaction Privacy-Preserving Deep Learning via Function Secret Sharing. arXiv, October 28, 2021. http://arxiv.org/abs/2006.04593

ANALYTICAL REVIEW OF CONFIDENTIAL ARTIFICIAL INTELLIGENCE: METHODS AND ALGORITHMS FOR DEPLOYMENT IN CLOUD COMPUTING

© 2024 E. M. Shiriaev[a], A. S. Nazarov[a], N. N. Kucherov[a], M. G. Babenko[a, b]

[a] *North Caucasus Federal University, Pushkin st. 1, Stavropol, 355017, Russia*

[b] *Ivannikov Institute for System Programming of the Russian Academy of Sciences, Alexander Solzhenitsyn st. 25, Moscow, 109004, Russia*

The technologies of artificial intelligence and cloud systems have recently been actively developed and implemented. In this regard, the issue of their joint use, which has been topical for several years, has become more acute. The problem of data privacy preservation in cloud computing acquired the status of critical long before the necessity of their joint use with artificial intelligence, which made it even more complicated. This paper presents an overview of both the artificial intelligence and cloud computing techniques themselves, as well as methods to ensure data privacy. The review considers methods that utilize differentiated privacy; secret sharing schemes; homomorphic encryption; and hybrid methods. The conducted research has shown that each considered method has its pros and cons outlined in the paper, but there is no universal solution. It was found that theoretical models of hybrid methods based on secret sharing schemes and fully homomorphic encryption can significantly improve the confidentiality of data processing using artificial intelligence.

Keywords: cloud computing, artificial intelligence, neural network, secret sharing scheme, homomorphic encryption, residue number system