УДК 004.042
DOI:10.30533/scidata-2023-14-01

Анализ подходов к обеспечению конфиденциальности пространственных вычислений в недоверенной вычислительной среде

Ю.Б. Брагина[1]✉

АФФИЛИАЦИИ

[1] Московский государственный университет геодезии и картографии, Москва, Россия
✉ bragina_y@mail.ru

ЦИТИРОВАНИЕ

Брагина Ю.Б. Анализ подходов к обеспечению конфиденциальности пространственных вычислений в недоверенной вычислительной среде // Пространственные данные: наука и технологии. 2023. Т. 14. № 1. С. 8–17. DOI:10.30533/scidata-2023-14-01.

КЛЮЧЕВЫЕ СЛОВА

обеспечение конфиденциальности данных, облачная модель распределенных вычислений, пространственные вычисления

АННОТАЦИЯ

В статье обсуждается общая проблематика обеспечения конфиденциальности данных в облачном окружении и выделяется специфическая проблема обеспечения конфиденциальности пространственных вычислений в такого рода среде, показаны причины этой специфичности и высказана идея возможного решения. Целью проводимого исследования является анализ подходов к обеспечению конфиденциальности пространственных вычислений в недоверенной вычислительной среде. Актуальность исследования обуславливается все более широким распространением различных носимых устройств, таких как умные очки или гарнитуры смешанной реальности, которые могут совмещать в поле зрения пользователя цифровую информацию из компьютерной системы

и из внешней среды, позволяя пользователю взаимодействовать с цифровым миром, сохраняя при этом осведомленность о физической среде. Для восприятия, анализа и интерпретации пространственной информации, поступающей из физической среды, необходимо выполнение пространственных вычислений, причем не только на самом носимом устройстве, но и в облачном окружении. Идея решения для обеспечения конфиденциальности пространственных вычислений в недоверенной вычислительной среде, предложенная автором, основана на шифровании пространственного индекса, построенного в виде R*-дерева.

1 Введение

Технологические достижения последнего времени революционизируют способы нашего взаимодействия с цифровой информацией. К такого рода технологическим достижениям следует отнести и технологию пространственных вычислений, которая предполагает использование компьютерных алгоритмов для восприятия, анализа и интерпретации пространственной информации, поступающей из физической среды. Источником такой пространственной информации могут быть различные носимые устройства, например умные очки или гарнитуры смешанной реальности. Эти устройства могут совмещать в поле зрения пользователя цифровую информацию из компьютерной системы и из внешней среды, позволяя пользователю взаимодействовать с цифровым миром, сохраняя при этом осведомленность о физической среде. Примеры успешного и эффективного использования такого рода взаимодействия приведены в работах [1–3]. Эти работы описывают использование пространственных вычислений и иммерсивных виртуальных технологий для целей управления современным городским хозяйством.

2 Материалы и методы

При использовании пространственных вычислений возникают важные вопросы, касающиеся конфиденциальности используемых данных и информационной безопасности самих пространственных вычислений. Для выполнения пространственных вычислений может быть использована (и практически всегда используется на практике) такая вычислительная среда, как облачные вычисления.

Согласно[1] облачные вычисления — это «модель, обеспечивающая удобный сетевой доступ по требованию к общим конфигурируемым вычислительным

1 Special Publication 800-145, The NIST Definition of Cloud Computing [Электронный ресурс]. Режим доступа: http://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-145.pdf (дата обращения: 06.11.2023).

ресурсам (сетям, серверам, хранилищам данных, приложениям и сервисам), который оперативно предоставляется с минимальными усилиями по управлению и взаимодействию с сервис-провайдером». Эта модель распределенных вычислений включает в себя пять основных базовых характеристик, три сервисные модели и четыре модели развертывания. Как отмечено в [4], специфика модели облачных вычислений требует исследования применимости тех методов защиты данных, которые использовались ранее, и, возможно, разработки новых методов в этой области. Особенно это касается методов обеспечения конфиденциальности данных: хранение данных в облаке не может считаться безопасным с точки зрения обеспечения конфиденциальности данных, поскольку нельзя быть полностью уверенным, что хранимые данные не могут быть раскрыты злоумышленником из числа сотрудников облачного сервис-провайдера третьей стороне. Для описания этой ситуации различными исследователями были предложены разные модели недоверенного облачного провайдера. В настоящей работе вычислительная среда облачного провайдера считается недоверенной по сути работы [5]. Это значит, что угроза нарушения конфиденциальности (раскрытия данных третьей стороне) считается актуальной, а угрозы нарушения целостности и доступности — нет.

Компании, заинтересованные в развитии модели облачных вычислений, сформировали Альянс облачной безопасности (англ. — «Cloud Security Alliance», CSA), который организует и проводит исследования в области защиты данных в облаке. В результате этих исследований Альянс выпустил важный и постоянно обновляемый документ, своего рода руководство по безопасности облачных вычислений [6]. Согласно этому документу при использовании моделей облачных вычислений IaaS (англ. — «Infrastructure as a Service»), PaaS (англ. — «Platform as a Service») или SaaS (англ. — «Software as a Service»), инсайдер (например, сотрудник облачного провайдера) может нарушить конфиденциальность хранимых данных. Даже шифрование данных не решает проблему, если ключ (или ключи) шифрования не хранится и не используется только на стороне клиента.

3 Результаты и обсуждение

Такой характер угроз требует внедрения соответствующих мер безопасности, основной из которых является использование различных технологий шифрования данных. Шифрование данных на стороне клиента, таким образом, становится необходимой частью мер защиты данных в облаке.

Однако при реализации этого подхода возникает проблема обработки зашифрованных данных. В частности, поиск в таких зашифрованных данных, хранящихся в облаке, не представляется тривиальной задачей. Для решения этой проблемы в научной литературе описаны различные подходы.

Одним из самых сильных, во всяком случае, с точки зрения безопасности, считается подход к поиску в зашифрованных данных, который основан на использовании забывчивой памяти (в иностранной научной литературе используется термин «Oblivious RAM», или ORAM). Непосредственно концепция забывчивой памяти была предложена в работе [6]. Эта концепция предлагала такое устройство памяти, которое бы позволило считывать и записывать данные в память без сохранения информации о том, в каких областях памяти были произведены эти операции чтения и записи. Реализация концепции должна была основываться на специальных протоколах для чтения и записи. Протокол чтения в качестве входных параметров использует N номер элемента памяти и секретный ключ, а в качестве выходного значения выдает значение этого N элемента памяти. Серверная часть в результате выполнения этого протокола не получает никакого значения. Протокол записи в качестве входного значения получает от пользователя номер элемента памяти N и данные, которые нужно записать в этот элемент. Серверная часть, выполняя протокол, передает ему область забывчивой памяти и получает ее обратно уже измененной. Наличие подобных аппаратно-реализованных протоколов дало бы возможность разработать решение для поискового процесса в зашифрованных данных. Для этого необходимо наличие двух модулей забывчивой памяти на стороне облачного сервис-провайдера. Первый модуль — для хранения зашифрованных данных, а второй — для хранения структуры, которая поддерживала бы операцию поиска, то есть поисковый индекс. Эта структура тоже должна храниться в зашифрованном виде. Процедура поиска будет представлять собой набор обращений клиента к сервис-провайдеру с использованием описанного протокола чтения, а процедура загрузки данных — набор обращений клиента к сервис-провайдеру с использованием описанного протокола записи. Стойкость такого подхода будет определяться стойкостью свойства «забывчивости» используемых модулей. Привлекательность этого решения состоит в том, что оно, будучи реализованным, скрывает от сервис-провайдера практически все, даже доступ к результатам поиска. Однако **решение имеет и существенные недостатки:**

1) оно требует соответствующих аппаратных решений на стороне сервис-провайдера;
2) недостатком использования забывчивой памяти будет относительно невысокая скорость поиска, что, в свою очередь, приведет к ограничениям на размер хранимых таким образом данных.

Если не принимать во внимание аппаратные решения проблемы поиска в зашифрованных данных на стороне недоверенного провайдера (таких как описанный выше вариант с использованием забывчивой памяти), то **наиболее очевидным решением подобной проблемы, на первый взгляд, может быть следующее:**

1) создать локальную копию зашифрованных данных, в которых должен быть произведен поиск;
2) расшифровать эту локальную копию данных;
3) произвести поиск в расшифрованной локальной копии данных;
4) получить результаты поиска.

Однако такое решение вопроса напрямую возможно, только если для проведения поиска достаточно части зашифрованных данных, причем эту часть возможно отделить от остальных зашифрованных данных, не расшифровывая их. Если же речь идет о полной выгрузке всех зашифрованных данных из облака для проведения поиска, то такое решение не представляется логичным и обессмысливает перенос данных в облако.

Решением проблемы поиска в зашифрованных данных могло бы быть использование дополнительных структур данных — поисковых индексов, хранимых на стороне клиента. При использовании такого подхода перед отправкой данных в облако выполняется создание поискового индекса одновременно с шифрованием данных. Затем этот индекс на клиентской стороне используется при отсылке запросов на получение зашифрованных данных из облака. Сам индекс хранится локально на клиентской стороне.

Подобный подход имеет следующие недостатки:

- необходимость хранить локальную копию поискового индекса на всех компьютерах, откуда предполагается получать доступ к данным, хранимым в облаке (следует заметить, что получение доступа к данным в облачном окружении с помощью носимых устройств — смартфонов и планшетов — уже стало ожидаемой функцией), что увеличивает вероятность реализации угрозы утечки этого поискового индекса;
- необходимость дополнительных вычислительных ресурсов на клиентской стороне для выполнения операции поиска в данных.

Таким образом, мы приходим к идее разработки систем, способных хранить поисковый индекс непосредственно в облаке. В этом случае, разумеется, этот **индекс должен храниться в зашифрованной форме, причем таким образом, что:**

- индекс не должен делать возможным раскрытие облачному сервис-провайдеру информации о хранимых данных;
- индекс должен обеспечивать эффективные в вычислительном смысле методы поиска в зашифрованных данных;
- доступ к использованию поискового индекса должен осуществляться с помощью секретного ключа клиента, недоступного облачному сервис-провайдеру.

Насколько известно автору настоящей статьи, впервые решение для защиты данных, отданных на аутсорсинговое хранение внешнему недоверенному провайдеру, было предложено в работе [7]. Решение состояло в конструировании индекса, основанного на зашифрованных данных, и введении добавочной информации для поддержки запросов к зашифрованным данным. Позднее в работе [8]

была предложена криптографическая схема для сохраняющего линейный порядок шифрования (в работе эта схема названа «order-preserving encryption scheme», OPES). Схема могла быть использована для любых одномерных данных, в которых задан линейный порядок. В работе была доказана вычислительная эффективность предложенной схемы, которая позволила бы реализовать эффективные запросы к данным у сервис-провайдера. Интересная идея в этой области, отличная от прежде изложенных здесь подходов, содержится в работе [9]. В этой работе был представлен подход к сохранению конфиденциальности данных в модели аутсорсингового хранения и обработки данных у облачного провайдера, который позволяет выполнять нечеткие запросы к зашифрованным строкам.

Однако все упомянутые подходы применимы либо к строкам, либо к одномерным данным с заданным порядком, и не могут быть непосредственно использованы для защиты пространственных вычислений.

Выполнение запроса на проверку пространственных отношений между объектами требует большого количества вычислений и при выполнении напрямую может сделать реализацию такого запроса чрезвычайно затратным как с точки зрения времени, так и с точки зрения требующихся вычислительных ресурсов. Поэтому для работы с пространственными данными в соответствующих расширениях современных реляционных баз данных используются специальные модели данных (R, R+, R*-деревья, BSP-деревья, k-мерные деревья, квадро- и окто-деревья и т.д.). Следует отметить, что в современных условиях использование реляционных баз данных для выполнения пространственных запросов не всегда является возможным. В работе [10] были исследованы специальные случаи, когда реляционные базы данных неприменимы для обработки пространственно-временных данных, и предложены соответствующие решения. Таким образом, в зависимости от типа пространственных запросов и условий их применения для их эффективного выполнения используются различные модели пространственных данных, и любой подход к поиску в зашифрованных пространственных или пространственно-временных данных должен учитывать эту специфику.

Автор считает, что для сохранения пространственно-временных данных в процессе выполнения пространственных вычислений возможно использовать шифрование пространственного индекса, построенного в виде R*-дерева. Общая схема передачи данных облачному провайдеру согласно предлагаемому подходу показана на **Рисунке 1.**

Рис. 1 Общая схема передачи данных облачному провайдеру для выполнения пространственных вычислений.

Fig. 1 Common scheme for transferring data to a cloud provider to perform spatial computing.

4 Выводы

Проведенный анализ показал, что описанные в научной литературе подходы к обеспечению конфиденциальности в недоверенной среде применимы либо к строкам, либо к данным с заданным порядком, и не могут быть непосредственно использованы для защиты пространственных вычислений. Это связано со спецификой моделей пространственных данных, и любой подход к организации выполнения пространственных запросов к зашифрованным пространственным или пространственно-временным данным должен учитывать эту специфику. Результатом анализа также стало предложение обеспечивать конфиденциальность пространственных вычислений в недоверенной вычислительной среде, шифруя не только сами данные, но и пространственный индекс, построенный в виде R*-дерева. Определение вычислительной эффективности такого подхода требует проведения дополнительных исследований.

БЛАГОДАРНОСТИ

Вычислительные эксперименты, связанные с проверкой применимости и возможной эффективности этого подхода, проводятся в настоящее время на базе Научного центра «Лаборатория пространственных вычислений и искусственного интеллекта» ФГБОУ ВО «Московский государственный университет геодезии и картографии» (МИИГАиК).

БИБЛИОГРАФИЯ

1. Hämäläinen M. Urban development with dynamic digital twins in Helsinki city // IET Smart Cities. 2021. No. 3(4). P. 201–210. DOI:10.1049/smc2.12015.
2. Hao H., Wang Y. Smart Curb Digital Twin: Inventorying Curb Environments Using Computer Vision and Street Imagery // IEEE Journal of Radio Frequency Identification. 2022. Vol. 7. P. 168–172. DOI:10.1109/JRFID.2022.3225733.
3. He X., Ai Q., Wang J., et al. Situation Awareness of Energy Internet of Thing in Smart City Based on Digital Twin: From Digitization to Informatization // IEEE Internet Things Journal. 2023. Vol. 10. No. 9. P. 7439–7458. DOI:10.1109/JIOT.2022.3203823.
4. Информационная безопасность цифрового пространства / под ред. Е.В. Стельмашонок, И.Н. Васильевой. СПб.: Изд-во СПбГЭУ, 2019. 155 с.
5. Damiani E., De Capitani, Vimercati S., et al. Balancing confidentiality and efficiency in untrusted relational DBMSs // Proceedings of the 10th ACM conference on Computer and communications security (CCS '03). 2003. P. 93–102. DOI:10.1145/948109.948124.
6. Goldreich O., Ostrovsky R. Software protection and simulation on oblivious RAMs // Journal of the ACM. 1996. No. 3. P. 431–473.
7. Hacigümüş H., Iyer B., Li C., Mehrotra S. Executing SQL over encrypted data in the database-service-provider model // Proceedings of the ACM SIGMOD International Conference on Management of Data (SIGMOD '02). 2002. P. 216–227.
8. Agrawal R., Kiernan J., Srikant R., Xu Y. Order preserving encryption for numeric data // Proceedings of the ACM SIGMOD International Conference on Management of Data (SIGMOD '04). 2004. P. 563–574.
9. Huang R.W., Gui X.L., Yu S., Zhuang W. Study of privacy-preserving framework for cloud storage // Computer Science and Information Systems. 2011. Vol. 8(3). P. 801–819.
10. Матерухин А.В. Теоретические основы и методология обработки потоков пространственно-временных данных: дис. ... д-ра техн. наук: 25.00.35. 2018. 173 с.

АВТОР

Брагина Юлия Борисовна
ФГБОУ ВО «Московский государственный университет геодезии и картографии» (МИИГАиК), Москва, Россия
0009-0006-3509-2517

Поступила 22.05.2023. Принята к публикации 23.06.2023. Опубликована 30.06.2023.

UDC 004.042
DOI:10.30533/scidata-2023-14-01

Analysis of approaches to ensuring confidentiality of spatial computing in the untrusted computing environment

Yulia B. Bragina[1]

AFFILIATIONS

[1] Moscow State University of Geodesy and Cartography, Moscow, Russia
bragina_y@mail.ru

CITATION

Bragina YuB. Analysis of approaches to ensuring confidentiality of spatial computing in the untrusted computing environment. *Spatial Data: science, research and technology.* 2023;14(1): 8–17. DOI:10.30533/scidata-2023-14-01.

KEYWORDS

data confidentiality, cloud model of distributed computing, spatial computing

ABSTRACT

The article discusses a general problem of ensuring data confidentiality in a cloud environment and highlights a specific problem of ensuring the confidentiality of spatial computing in this kind of environment, shows the reasons for this specialty and suggests a possible solution. The purpose of this research is to analyze the approaches for ensuring the confidentiality of spatial computing in the untrusted computing environment. The relevance of the study is determined by the increasingly widespread use of various wearable devices, such as smart glasses or mixed reality headsets, which can combine digital information from the computer system and from the external environment in the user's field of view, allowing him to interact with the digital world while maintaining awareness of the physical world. To perceive, analyze and interpret spatial information coming from the physical environment, it is necessary

to perform spatial computing, not only by a wearable device itself, but also in a cloud environment. The idea of the solution for ensuring the confidentiality of spatial computing in the untrusted computing environment is proposed by the author and is based on encryption of a spatial index constructed in the form of the R*-tree.

REFERENCES

1. Hämäläinen M. Urban development with dynamic digital twins in Helsinki city. *IET Smart Cities*. 2021;3(4): 201–210. DOI:10.1049/smc2.12015.
2. Hao H, Wang Y. Smart Curb Digital Twin: Inventorying Curb Environments Using Computer Vision and Street Imagery. *IEEE Journal of Radio Frequency Identification.* 2022;7: 168–172. DOI:10.1109/JRFID.2022.3225733.
3. He X, Ai Q, Wang J, et al. Situation Awareness of Energy Internet of Thing in Smart City Based on Digital Twin: From Digitization to Informatization. *IEEE Internet Things Journal.* 2023;10(9): 7439–7458. DOI:10.1109/JIOT.2022.3203823.
4. *Informacionnaja bezopasnost' cifrovogo prostranstva* [Information security of the digital space]. Ed. Stelmashonok EV, Vasilevoj IN. Saint Petersburg: SPbGEU; 2019. 155 p. (In Russian).
5. Damiani E, De Capitani, Vimercati S, et al. Balancing confidentiality and efficiency in untrusted relational DBMSs. *Proceedings of the 10th ACM conference on Computer and communications security (CCS '03)*. 2003; 93–102. DOI:10.1145/948109.948124.
6. Goldreich O, Ostrovsky R. Software protection and simulation on oblivious RAMs. *Journal of the ACM*. 1996;3: 431–473.
7. Hacigümüş H, Iyer B, Li C, Mehrotra S. Executing SQL over encrypted data in the database-service-provider model. *Proceedings of the ACM SIGMOD International Conference on Management of Data (SIGMOD '02)*. 2002; 216–227.
8. Agrawal R, Kiernan J, Srikant R, Xu Y. Order preserving encryption for numeric data. *Proceedings of the ACM SIGMOD International Conference on Management of Data (SIGMOD '04)*. 2004; 563–574.
9. Huang RW, Gui XL, Yu S, Zhuang W. Study of privacy-preserving framework for cloud storage. *Computer Science and Information Systems*. 2011;8(3): 801–819.
10. Materuhin AV. *Teoreticheskie osnovy i metodologija obrabotki potokov prostranstvenno-vremennyh dannyh* [Theoretical foundations and methodology for processing spatiotemporal data streams: dissertation]. Dr. of Sci. thesis. Moscow: 2018; 173 p. (In Russian).

AUTHOR

Yulia B. Bragina
Moscow State University of Geodesy and Cartography, Moscow, Russia
0009-0006-3509-2517

Submitted: May 22, 2023. Accepted: June 23, 2023. Published: June 30, 2023.