ОБОЗРЕНИЕ
ПРИКЛАДНОЙ И ПРОМЫШЛЕННОЙ
Том 23 МАТЕМАТИКИ Выпуск 5
2016

А. В. Зязин, **С. Ю. Катышев** (Москва, Московский технологический ун-т)**. Способ безопасного распределения баз данных между ограниченно доверенными облаками.**

Использование облачной инфраструктуры влечет за собой ряд рисков безопасности информации. В том числе, связанных с несанкционированным использованием данных пользователя провайдером облака. Для простого хранения информации, не предполагающего ее обработку на стороне провайдера, проблема может быть решена с помощью шифрования всего блока данных. Однако, при размещении баз данных такой подход вступает в противоречие с необходимостью осуществления операций поиска и обработки непосредственно в облаке. Сложившиеся подходы к защите облачных баз данных основаны на использовании нескольких типов криптографических примитивов, прежде всего это гомоморфное шифрование (HE), шифрование, сохраняющее порядок (OPE) и блочное шифрование (CBC). Исходя из структуры необходимых запросов к базе данных, для каждого поля выбирают подходящий тип шифрования. Например, если необходим поиска по полю числового формата, для его шифрования целесообразно использовать OPE.

В качестве типичных примеров практической реализации указанных подходов приведем СУБД CryptDB и StealthDB ([1], [2]). CryptDB строится поверх MySQL/PostgreSQL и использует принцип «луковичного» (послойного) шифрования значения поля указанными примитивами. При этом вместе с запросом к СУБД передается минимально необходимый для его выполнения набор ключей расшифрования. Очевидный недостаток такого способа — потенциальная возможность быстрого накопления у провайдера облака ключей для расшифрования полей до слоев OPE или HE, за исключением быть может неиспользуемых при поиске полей. При этом OPE и HE являются криптографически уязвимыми, прежде всего относительно атак по парам открытый-шифрованный текст.

В StealthDB, помимо собственно шифрования содержимого полей, используется идея предшествующего разделения поля между несколькими местами хранения (ОЗУ, файловая система на стороне клиента, одно или несколько облаков). При этом должна быть обеспечена возможность соответствующей переформулировки поисковых запросов. Пример такого подхода — раздельное хранение старшего и младшего байтов 16-битного представления целого беззнакового числа.

В докладе предлагается новый способ повышения защищенности базы данных, основанный на ее разделении между облаками, поддерживаемыми разными провайдерами (предполагается, что они не вступают в сговор против владельца базы данных). В отличие от StealthDB, это разделение касается служебной информации, необходимой для выполнения запросов к б. д. Такой подход близок к использовавшемуся в [3].

Реляционная база данных состоит из набора прямоугольных таблиц, для простоты будем считать что таблица только одна. Пусть Ω, Z — конечные непустые множества. Далее под базой данных будем подразумевать произвольное непустое подмножество B множества $\Omega \times Z^m$, $m \in \mathbf{N}$, такое, что любые два несовпадающих элемента из B отличаются в первой координате. Элементы B называют записями.

С базой данных свяжем совокупность индексных множеств

$$I_{i,z} = \{\omega \in \Omega \,|\, \exists\, b_1, \ldots, b_{i-1}, b_{i+1}, \ldots, b_m : (\omega, b_1, \ldots, b_{i-1}, z, b_{i+1}, \ldots, b_m) \in B\}.$$

Под поиском в б. д. подразумеваем вычисление результата любой конечной последовательности теоретико-множественных операций над индексными множествами.

Выберем вспомогательное множество R достаточно большой мощности, ключ для блочного шифрования на множестве $Z \times \{1, \ldots, m\} \times R$ и m ключей для OPE на множестве $Z \times \{1, \ldots, m\}$. Соответствующие функции обозначим CBC_1 и $\mathrm{OPE}_1, \ldots, \mathrm{OPE}_m$. В первом облаке храним таблицу из двух столбцов, строки которой имеют вид

$$(\mathrm{OPE}_i(z, i), \mathrm{CBC}_1(\omega, i, r)),$$

где $z \in Z$, $i \in \{1, \ldots, m\}$, $\omega \in \Omega$, $r \in R$. Строки таблицы образуются «расщеплением» записей вида $(\omega, z_1, \ldots, z_m)$ из B — на $(\omega, z_1), \ldots, (\omega, z_m)$, с последующим случайным выбором значений r для каждой строки и применением соответствующих функций шифрования. Шифрование производится на стороне владельца базы данных, в первое облако ключи не передаются.

Для работы со вторым облаком владелец базы данных предоставляет ключ для расшифрования CBC_1. Кроме того, выбирается функция блочного шифрования CBC_2 множестве на Z^m. Ключ в облако не передается.

Затем во втором облаке формируются две таблицы. Первая состоит из двух столбцов, ее строки получаются преобразованием каждой строки $(\omega, z_1, \ldots, z_m)$ из в строку вида

$$(\omega; \mathrm{CBC}_2(z_1, \ldots, z_m)).$$

Вторая таблица состоит из трех столбцов. Первый — для хранения номера сеанса при поиске, второй — типа Ω, последний — типа $\{1, \ldots, m\}$.

Выполнение поискового запроса в B выполняется в две стадии. Сначала средствами первого облака выделяются и передаются в зашифрованном виде во второе облако необходимые индексные множества. При этом в первом облаке необходимо выполнять запросы на поиск строк, удовлетворяющих двустороннему неравенству по первому столбцу, формируемые на стороне владельца. Их результаты — по второму столбцу — передаются во второе облако.

В нем полученный результат расшифровывается (CBC_1) и заносится во вторую таблицу. Переданная владельцем информация о необходимых операциях над индексными множествами позволяет сформировать итоговый результат с помощью стандартных поисковых запросов. Соответствующие ему шифрованные записи из второго столбца первой таблицы передаются владельцу.

По сравнению с рассмотренными примерами предложенный способ организации базы данных затрудняет накопление информации о ее содержимом путем анализа трафика независимыми провайдерами облаков, в том при возможности инициирования ими запросов через владельца.

REFERENCES

1. *Popa R. A.*, *Redfield C. M. S.*, *Zeldovich N.*, *Balakrishnan H.* CryptDB: Protecting Confidentiality with Encrypted Query Processing. — Proceedings of the 23rd ACM Symposium on Operating Systems Principles (SOSP). ACM, October 2011, p. 85–110.
2. *Josef Spillner*, *Martin Beck*, *Alexander Schill*, *Thomas Michael Bohnert.* Stealth Databases: Ensuring Secure Queries in Untrusted Cloud Environments. 8th IEEE/ACM UCC, 2015.

3. *Зязин А. В.* Использование электронных ключей для защиты локальных баз данных от несанкционированного копирования. — Обозрение прикл. и промышл. матем., 2002, т. 9, в. 2, с. 380–381.