УДК 004.056.5

М. П. Мищенко

ГИБРИДНАЯ АТАКА НА ЗАДАЧУ ОБУЧЕНИЯ С ОШИБКАМИ (LWE) С РАЗРЯЖЕННЫМ СЕКРЕТОМ

В 2007 г. Н. Хаугрейв-Грэм предложил атаку на криптосистему NTRU, которая заключается в совмещении техники редукции решеток и комбинаторного метода «встречи посередине» (атаки meet-in-the-middle, MiTM). В данной работе этот подход применяется к задаче Learning with Errors (LWE) с секретом малой длины. Задача LWE является одной из самых важных в теории криптографии на решетках. Безопасность большого количества криптографических схем, начиная от простых подписей и схем и заканчивая продвинутыми схемами, такими как групповые подписи и полностью гомоморфное шифрование, основывается на предположениях о сложности LWE. В статье представлен обзор гибридной атаки, а также используемых в ней алгоритмов. Это необходимо для дальнейшей практически значимой реализации атаки, главной целью которой является верификация корректности применения метода MiTM к гибридной атаке на LWE.

In 2007, Howgrave-Graham proposed attack against NTRU cryptosystem, which consists of two parts, combining lattice reduction technique and a combinatorial method called meet-in-the-middle (MiTM). In this article, we apply hybrid attack to the Learning with Errors Problem (LWE) with sparse secret. The LWE problem is considered to be one of the most important in lattice-based cryptography. Large number of cryprographic schemes ranging from basic signature and encryption schemes to advanced schemes like group signatures and fully homomorphic encryption, base their security on the hardness assumption of LWE. In this paper, we review the hybrid attack and the algorithms it is based on. It is required for further practical implementation of the attack, whose main objective is to verify correctness of MiTM to the hybrid attack against the LWE problem.

Ключевые слова: гибридная атака, криптография на решетках, обучение с ошибками, метод «встречи посередине».

Keywords: hybrid attack, lattice based cryptography, LWE, MiTM.

Решетки

В 1996 г. венгерский математик Миклош Айтаи опубликовал работу [2], в которой доказал, что можно построить случайную решетку с коротким вектором такую, что любой алгоритм нахождения этого вектора в данной решетке будет сводим к задаче нахождения приблизительно кратчайшего вектора в любой решетке. Этот факт положил начало направлению в современной криптографии, которое называют lattice based cryptography, или криптография на решетках.

Схемы, реализованные с помощью криптографии на решетках, имеют огромный потенциал и вызывают большой интерес к их изучению.

Вестник Балтийского федерального университета им. И. Канта.
Сер.: Физико-математические и технические науки. 2020. № 4. С. 10 – 15.

В настоящее время происходит процесс стандартизации постквантовых криптографических протоколов (в частности, на решетках) в России и США. В 2018 г. Национальный институт стандартов и технологий опубликовал список из открытых вопросов, касающихся безопасности криптографических схем, основанных на решетках. Гибридная атака, изучаемая в данной работе, также релевантна для криптосистемы NTRU — кандидата к стандартизации. Таким образом, выбор параметров криптосистем, основанных на решетках, является открытым вопросом, требующим детального анализа.

Пусть $\mathbf{B} = (\mathbf{b}_0, \ldots, \mathbf{b}_{n-1}) \subset \mathbb{R}$ — множество линейно независимых векторов. Решеткой $\mathcal{L}$, порожденной векторами $\mathbf{b}_0, \ldots, \mathbf{b}_{n-1}$, называют множество всех линейных комбинаций этих векторов с коэффициентами в $\mathbb{R}$:

$$\mathcal{L}(\mathbf{B}) = \{\alpha_0\mathbf{b}_0 + \cdots + \alpha_{n-1}\mathbf{b}_{n-1} | \alpha_0, \ldots, \alpha_{n-1} \in \mathbb{Z}\} \in \mathbb{R}^m.$$

Целочисленной называют решетку, все векторы которой имеют целочисленные координаты. Таким образом, целочисленная решетка является аддитивной подгруппой $\mathbb{Z}^m$ для некоторого $m \geq 1$. Базисом $\mathcal{L}$ называют любой набор независимых векторов, которые порождают $\mathcal{L}$, размерностью $\mathcal{L}$ — количество векторов в любом базисе $\mathcal{L}$.

Криптография на решетках основывается на вычислительной сложности решения определенных задач. Рассмотрим задачи, необходимые для описания атаки на LWE.

В задаче поиска ближайшего вектора (CVP) для $\mathbf{t} \notin \mathcal{L}$ требуется найти $\mathbf{v} \in \mathcal{L}$:

$$\|\mathbf{v} - \mathbf{t}\| \text{ минимально для всех } \mathbf{v} \in \mathcal{L}.$$

Пусть $dist(t, \mathcal{L}) \leq \frac{1}{\gamma}\lambda_1(\mathcal{L})$. Тогда получаем задачу о декодировании с ограниченным расстоянием (Bounded distance decoding, BDD): для базиса $\mathbf{B}$ решетки $\mathcal{L}$ и целевого вектора $\mathbf{t} \in \mathbb{R}^m$ найти вектор $\mathbf{e} \in \mathbb{R}^m$ такой, что выполняется $\|\mathbf{e}\| < \alpha\lambda_1(\mathcal{L}(\mathbf{B}))$ и $\mathbf{t} - \mathbf{e} \in \mathcal{L}(\mathbf{B})$.

Эвристики и алгоритмы

Теорема *(Теорема Эрмита). Для любой решетки $\mathcal{L}$ размерности n существует такой ненулевой вектор $\boldsymbol{b} \in LL$, что*

$$\|\mathbf{b}\| \leq \sqrt{n}\det(\mathcal{L})^{\frac{1}{n}}.$$

Гауссова эвристика предсказывает, что количество $|\mathcal{L} \cap \mathcal{B}|$ точек решетки, лежащих внутри измеримого объекта $\mathcal{B} \subset \mathbb{R}^n$, приблизительно равно $\mathcal{B}/Vol(\mathcal{L})$. Применяя эту эвристику для шара в n-мерном евклидовом пространстве, получим следующую оценку $\lambda_1(\mathcal{L})$ для заданной решетки $\mathcal{L}$.

Гауссова эвристика. Обозначим с помощью $gh(\mathcal{L})$ первый ожидаемый минимум решетки $\mathcal{L}$ согласно гауссовой эвристике. Для решетки $\mathcal{L}$ полного ранга в $\mathbb{R}^d$ имеем

$$gh(\mathcal{L}) = \sqrt{d/2e}Vol(\mathcal{L})^{\frac{1}{d}}.$$

Качество базиса после редукции можно измерить величиной, называемой корень-фактором Эрмита.

Корень-фактор Эрмита (эрмитова дельта) для базиса **B** решетки размерности d определяется как

$$\delta = (|\mathbf{b}_0/Vol(\mathbf{B})^{1/d})^{1/d}.$$

Замечание: рассмотрим версию теоремы, которая дает более точную (верхнюю) оценку для нормы кратчайшего вектора. Константой Эрмита δ_n называют такое наименьшее число, что для любой решетки $\mathcal{L}$ размерности n найдется вектор $\mathbf{w} \in \mathcal{L}$, который удовлетворяет

$$\| \mathbf{w} \|^2 \leq \delta_n \det(\mathcal{L})^{\frac{2}{n}}.$$

Алгоритм Бабая (Nearest Plane Algorithm). На вход принимается базис $\mathbf{B} \subset \mathbb{Z}^n$ решетки $\mathcal{L}$ и целевой вектор $\mathbf{t} \in \mathbb{R}^n$. Выходом алгоритма является вектор $\mathbf{e} \in \mathbb{R}^n$ такой, что $\mathbf{t} - \mathbf{e} \in \mathcal{L}(\mathbf{B})$.

Под редукцией базиса **B** решетки $\mathcal{L}$ понимают процесс нахождения последовательных базисов $\mathcal{L}$, векторы которых более короткие и ближе к тому, чтобы быть ортогональными. Важными понятиями редукции являются алгоритмы HKZ и БКЗ-β reduction. Алгоритм БКЗ принимает на вход базис решетки $\mathcal{L}$ и размер блока β, возвращая БКЗ-β редуцированный базис решетки $\mathcal{L}$.

Блочный алгоритм Эрмита — Коркина — Золотарёва, БКЗ (Block Korkine — Zolotarev Algorithm, BKZ). Базис $\mathbf{B} = (\mathbf{b}_0, \dots, \mathbf{b}_{d-1})$ решетки $\mathcal{L}$ называется БКЗ-редуцированным (BKZ) с блоком размера β (БКЗ-β), если для его векторов справедливо

$$\mathbf{b}_i^* = \lambda_1(\mathcal{L}_{[i:\min(i+\beta,d)]}, \forall i < d.$$

Другими словами, БКЗ-редукция требует, чтобы конкретный вектор в базисе был наикратчайшим только при рассмотрении локальной спроектированной подрешетки таким образом, что локальность зависит от β. Трудность выполнения алгоритма БКЗ растет вместе с ростом β. В случае БКЗ-β-редукции дельта Эрмита является довольно предсказуемой величиной. Для блоков маленького размера дельта Эрмита вычисляется экспериментально, для блоков большого размера справедлива асимптотическая формула

$$\delta(\beta)^{2\beta-1} = (\beta/(2\pi e))(\beta\pi)^{\frac{1}{\beta}}.$$

LLL-алгоритм (Lenstra-Lenstra-Lovász, LLL) описывает нахождение оптимального ортогонального базиса для решеток любой размерности. Пусть $\mathbf{B} = (\mathbf{b}_0, \dots, \mathbf{b}_{n-1})$ — базис решетки $\mathcal{L}$ и $\mathbf{B}^* = (\mathbf{b}_0^*, \dots, \mathbf{b}_{n-1}^*)$ — соответствующий ортогональный базис. Говорят, что базис $\mathbf{B}^*$ LLL-редуцирован, если выполняются следующие два условия:

1) условие «размера»:

$$\| \mu_{i,j} \| = \frac{|\mathbf{b}_i \cdot \mathbf{b}_j^*|}{\|\mathbf{b}_j^*\|^2} \leq \frac{1}{2}, \text{для } 0 \leq j \leq n-1;$$

2) условие Ловаса:

$$\| \mathbf{b}_i^* \| \geq \left(\frac{3}{4} - \mu_{i,i-1}^2\right) \| \mathbf{b}_{i-1}^* \|^2, \text{для } 0 \leq j \leq n-1.$$

Задача обучения с ошибками (Learning with Errors, LWE)

Задача LWE [3]: Пусть n, q — положительные целые, χ — некоторое вероятностное распределение на $\mathbb{Z}$, $\mathbf{s}$ — некоторый (секретный) вектор в $\mathbb{Z}_q^n$. Обозначим с помощью $\Psi_{\mathbf{s},\chi,q}$ вероятностное распределение на $\mathbb{Z}_q^n \times \mathbb{Z}_q$, полученное путем выбора векторов $\mathbf{a} \in \mathbb{Z}_q^n$ случайно и равномерно и $\mathbf{e} \in \mathbb{Z}_q^m$ согласно χ. Требуется вернуть пару $(\mathbf{a}, \mathbf{c}) = (\mathbf{a}, \mathbf{as} + \mathbf{e}) \in$ $\in \mathbb{Z}^n \times \mathbb{Z}_q$.

В качестве вероятностного распределения $\chi \in \mathbb{Z}_q$ обычно берутся распределения, близкие к нормальному (распределение Гаусса N). Входными данными является пара $(\mathbf{A}, \mathbf{v})$, где $\mathbf{A} \in \mathbb{Z}_q^{m \times n}$ выбирается случайно согласно некоторому равномерному распределению, и вектор $\mathbf{As} + \mathbf{e}$ для равномерно выбранного $\mathbf{s} \in \mathbb{Z}_q^n$ и вектора $\mathbf{e} \in \mathbb{Z}_q^m$, выбранного с помощью распределения χ^m. Проблему составляет определение, каким путем был получен вектор, заданный $\mathbf{v}$. Эта задача может быть сведена к задаче BDD в q-арных решетках: для данной матрицы $\mathbf{A} \in \mathbb{Z}_q^{m \times n}$ и вектора $\mathbf{v} \in \mathbb{Z}_q^m$ определить, выбран вектор $\mathbf{v}$ равномерно из $\mathbb{Z}_q^m$ или же получен путем перебора всех координат случайной точки в решетке $\mathcal{L}_q(\mathbf{A}^T)$ согласно распределению χ.

Криптосистемы с разряженным секретом, то есть с секретом, коэффициенты которого принадлежат ограниченному интервалу, эффективны в прикладных задачах, так как позволяют быстрее производить вычисления. Рассмотрим общий вид атаки для случая, когда секрет является тринарным ($\mathbf{s} \in \{0, \pm 1\}^n$).

Гибридная атака на LWE

Для целых чисел $m, n, q \in \mathbb{Z}$, где $m > n$, положим

$$(\mathbf{A}, \mathbf{b}) = \mathbf{As} + \mathbf{e} \bmod q. \tag{1}$$

Это экземпляр LWE, где $\mathbf{A} \in \mathbb{Z}_q^{m \times n}, \mathbf{b} \in \mathbb{Z}_q^n$, секрет $\mathbf{s} \in \{0, \pm 1\}^n$ и вектор ошибки $\mathbf{e} \in \mathbb{Z}_q^m$.

Рассмотрим порожденную базисом решетку $\mathcal{L}(\mathbf{B})$ вида

$$\mathcal{L}(\mathbf{B}) = \begin{pmatrix} q\mathbf{I}_n & \mathbf{A} \\ \mathbf{0} & \mathbf{I}_n \end{pmatrix} \in \mathbb{Z}_q^{2n \times 2n}. \tag{2}$$

Вектор $(\mathbf{b}, \mathbf{0}) \in \mathbb{Z}_q^{2n}$ близок к решетке $\mathcal{L}(\mathbf{B})$ на $\|(\mathbf{e}, \mathbf{s})\|$. Разобьем матрицу $\mathbf{A}$ на две части: $\mathbf{A} = (\mathbf{A}_l | \mathbf{A}_g)$ так, что $\mathbf{A}_l \in \mathbb{Z}^{n \times l}$, $\mathbf{A}_g \in \mathbb{Z}^{n \times g}$ и $l + g = n$. Аналогичным образом разбиваем секрет $\mathbf{s}$ на две части: $\mathbf{s} = (\mathbf{s}_l | \mathbf{s}_g)$. Таким образом, получим новый экземпляр LWE:

$$\mathbf{b}' = \left(\mathbf{A}_l \middle| \mathbf{A}_g\right) \begin{pmatrix} \mathbf{s}_l \\ \mathbf{s}_g \end{pmatrix} + \mathbf{e} = \mathbf{A}_l \mathbf{s}_l + \mathbf{A}_g \mathbf{s}_g + \mathbf{e}' \bmod q.$$

Матрица $\mathbf{A}_l$ формирует подрешетку решетки $\mathcal{L}(\mathbf{B}')$, порожденную базисом:

$$\mathcal{L}(\mathbf{B}') = \begin{pmatrix} q\mathbf{I}_n & \mathbf{A}_l \\ \mathbf{0} & \mathbf{I}_l \end{pmatrix} \in \mathbb{Z}_q^{n+l \times n+l}. \quad (3)$$

Полученный таким образом новый экземпляр LWE решить гораздо проще, так как определитель $\det(L_q(A_g)) = q^m - n + g$ новой решетки существенно больше определителя $\det(L_q(A)) = q^m - n$ решетки $\mathbf{A}$. Для решения этой задачи можно свести проблему BDD к нахождению CVP.

Перед началом атаки вызывается процедура BDDPreprocess, которая редуцирует решетку $L_q(B')$ для поиска параметров атаки, в том числе размера блока β BKZ-редукции. Зная β, задачу BDD для решетки $\mathcal{L}(\mathbf{B}')$ можно свести к решению CVP в решетке $\mathcal{L}(\mathbf{B}')$.

Первую часть атаки (угадывание) можно ускорить с помощью метода «встречи посередине». Используя этот подход, возможно сократить количество попыток угадать вектор до квадратного корня от количества попыток при наивном подходе методом «грубой силы».

Алгоритм. Гибридная атака на LWE (MITM).

Input: $n, q \in \mathbb{Z}$ и $g \in \mathbf{N}$,

Output: $\mathbf{A} = (\mathbf{A}_l | \mathbf{A}_g)$, где $\mathbf{A}_l \in \mathbb{Z}^{n \times l}, \mathbf{A}_g \in \mathbb{Z}^{n \times g}, \mathbf{b} \in \mathbb{Z}_q^n$ вектор $\mathbf{s} \in \{0, \pm 1\}^n$ такой, что $\mathbf{b} = \mathbf{A}\mathbf{s} - \mathbf{e} \bmod q$.

1. Запустить алгоритм предобработки $BDDPreprocess(\mathcal{L}(\mathbf{B}'))$ для базиса $\mathbf{B}'$, заданного уравнением (3). Получим решетку $\mathbf{B}'_{red}$.

2. Сформируем два списка векторов L_1 и L_2:

$$L_1 = \{\mathbf{A}_l \mathbf{s}_1, \mathbf{s}_1)\} \in \mathbb{Z}_q^{n+l},$$

$$L_2 = \{\mathbf{b} - \mathbf{A}_l \mathbf{s}_2, \mathbf{s}_2)\} \in \mathbb{Z}_q^{n+l},$$

такие, что $\exists(\mathbf{s}_1, \mathbf{s}_2) \in L_1[2] \times L_2[2]$, в том числе $\mathbf{s}_g = \mathbf{s}_1 + \mathbf{s}_2$.

3. $V_1 = BDDQuery(\mathbf{B}'_{red}, L_1)$.

4. $V_2 = BDDQuery(\mathbf{B}'_{red}, L_2)$.

5. Вызвать $ClosestPairs(V_1, V_2, d)$ для целевого расстояния $d = \|(\mathbf{e}, \mathbf{s}_l)\|$.

Для того чтобы применить метод «встречи посередине», процедура BDDQuery должна быть аддитивно-гомоморфной. В качестве ее предлагается взять алгоритм Бабая, но его эвристика нуждается в проверке.

Также предполагается, что существует эффективная хеш-функция, которая решает проблему близких векторов ClosestPairs (V_1, V_2): для данных векторов V_1, V_2 и целевого расстояния (в евклидовой метрике) d найти все пары $(\mathbf{v}_1, \mathbf{v}_2) \in V_1 \times V_2$ такие, что $k\mathbf{v}_1 - \mathbf{v}_2$ для некоторого целого $k \leq d$. В качестве такой хеш-функции в данной реализации предлагается использовать алгоритм SimHash [4].

Работа над алгоритмом ведется на языке Python с использованием библиотек fpylll и G6k. На данный момент реализованы алгоритм SimHash, процедуры BDDPreprocess и ClosestPairs. Также предстоит реа-

лизовать процедуру BDDQuery, тем самым завершив предварительную реализацию атаки в общем виде. В дальнейшем предстоит проверить корректность применения MiTM к гибридной атаке, провести анализ времени выполнения алгоритма, определить вероятность успешного выполнения атаки, минимизировать время выполнения путем нахождения оптимальных параметров.

Список литературы

1. *Howgrave-Graham N.* A hybrid lattice-reduction and meet-in-the-middle attack against NTRU // Advances in Cryptology — CRYPTO 2007: Proceedings of 27th Annual International Cryptology Conference. Springer, 2007. Vol. 4622. P. 150—169. (Lecture Notes in Computer Science).

2. *Ajtai M.* Generating hard instances of lattice problems // Proceedings of the 28th annual ACM symposium on Theory of computing. Springer, 1996. P. 99—108.

3. *Regev O.* On lattices, learning with errors, random linear codes, and cryptography // Journal of the ACM. 2009. Vol. 56, iss. 6. P. 1—40.

4. *Fitzpatrick R., Bishof Ch. H., Buchmann J. et al.* Tuning GaussSieve for speed // International Conference on Cryptology and Information Security in Latin America. Springer, 2014. P. 288—305.

5. *Babai L.* On Lovász' lattice reduction and the nearest lattice point problem // Combinatorica. 1986. Vol. 6, iss. 1. P. 1—13.

6. *Buchmann J., Göpfert F., Player R., Wunderer Th.* On the hardness of LWE with binary error: Revisiting the hybrid lattice-reduction and meet-in-the-middle attack // International Conference on Cryptology in Africa. Springer, 2016. P. 24—43.

8. *Wunderer T.* Revisiting the Hybrid Attack: Improved Analysis and Refined Security Estimates // Cryptology ePrint Archive, Report 2016/733. URL: https://eprint.iacr.org/2016/733 (дата обращения: 15.10.2020).

9. *Micciancio D., Goldwasser S.* Complexity of lattice problems: a cryptographic perspective. Springer Science & Business Media, 2012. Vol. 671.

10. *Micciancio D.* Lattice-based cryptography // Encyclopedia of Cryptography and Security. Springer, 2011. P. 713—715.

11. *Conway J.H., Sloane N.J.A.* Sphere packings, lattices and groups. Springer Science & Business Media, 2013.

Об авторе

Максим Павлович Мищенко — студент, Балтийский федеральный университет им. И. Канта, Россия.

E-mail: mishchenkom_rel@bk.ru

The author

Maksim P. Mishchenko, Student, Immanuel Kant Baltic Federal University, Russia.

E-mail: mishchenkom_rel@bk.ru