ОБОЗРЕНИЕ
ПРИКЛАДНОЙ И ПРОМЫШЛЕННОЙ
Том 25 МАТЕМАТИКИ Выпуск 2
2018

Н. П. Варновский, **А. В. Шокуров** (Москва, Ин-т проблем информационной безопасности МГУ, ИСП РАН)**. Гомоморфное шифрование. Материалы для дискуссии.**

Доклад представляет собой обзор состояния дел в области гомоморфного шифрования, но обзор не типичный. Обсуждаются не достижения в этой области, а проблемы, а также «мифы» и «легенды», связанные с такими понятиями как гомоморфное шифрование и вычисления над зашифрованными данными. Содержание доклада может оказаться полезным для криптографов в качестве краткого введения в проблематику. Но в первую очередь обзор адресован математикам, которые решили заняться исследованиями гомоморфного шифрования, не имея специального криптографического образования.

Все рассказы о гомоморфном шифровании начинаются со ссылки на работу [1]. Зачастую утверждается, что именно в ней впервые была предложена концепция вычисления над зашифрованными данными. Но это утверждение больше похоже на «легенду». Идея вычислений над зашифрованными данными настолько естественна и очевидна, что ее появление лишь в 1978 г. выглядит весьма сомнительным. Вероятно, работа [1] — просто первая публикация, в которой эта идея обсуждалась.

За этим последовали десятилетия попыток найти методы шифрования, позволяющие выполнять произвольные вычисления над зашифрованными данными. Все эти попытки не приводили к успеху, и в результате сложилось мнение о несуществовании гомоморфного шифрования.

В 2009 г. вышла в свет известная работа Джентри [2]. Обычно утверждается, что в этой работе опровергнута гипотеза о несуществовании гомоморфного шифрования и решена, по крайней мере теоретически, проблема вычислений над зашифрованными данными. Здесь сразу два утверждения, которые можно причислить к «мифам»: об опровержении гипотезы и о решении проблемы. Но прежде чем обсуждать эти и другие «мифы», поясним кратко, каким образом строятся криптосистемы гомоморфного шифрования.

Всюду далее будем рассматривать открытые тексты из множества $\Sigma = \{0, 1\}$ и криптограммы из алгебраической системы A. Не претендуя на общность изложения, будем считать, что для вычислений и над однобитовыми открытыми текстами, и в алгебраической системе A используются операции сложения и умножения. Более того, допуская некоторую вольность, в обоих случаях будем обозначать эти операции через $+$ и $\cdot$, хотя, вообще говоря, это разные операции на разных множествах.

Наиболее простое объяснение идей, лежащих в основе конструкций криптосистем гомоморфного шифрования, можно найти в работе [4]. Это конструкция криптосистемы с секретным ключом.

Секретный ключ — достаточно большое натуральное число p.

Криптограмма c открытого текста $m \in \Sigma$ вычисляется по формуле $c = p \cdot q + 2r + m$, где q и r — случайные целые числа, выбираемые из предписанных интервалов. В частности, необходимо, чтобы выполнялось неравенство $|2r| < p/2$.

Дешифрование криптограммы c выполняется по формуле $m = (c \bmod p) \bmod 2$.

Очевидно, что если c_1 — криптограмма бита m_1, а c_2 — криптограмма бита m_2, то $c_1 + c_2$ и $c_1 \cdot c_2$ — криптограммы битов $m_1 + m_2$ и $m_1 \cdot m_2$ соответственно.

Но у этой несложной конструкции имеется серьезный дефект. Шум, изначально присутствующий в криптограммах ($2r$), при выполнении каждой операции, в особенности при умножении, возрастает. Поэтому можно выполнить лишь ограниченное количество операций над криптограммами, далее из-за слишком большого шума корректное дешифрование будет невозможно.

Такие криптосистемы называются невполне (somewhat) гомоморфными. Термин «невполне гомоморфная криптосистема» не слишком удачен, но единственную предлагаемую альтернативу «частично гомоморфная» следует зарезервировать за криптосистемами, гомоморфными относительно одной операции.

Проблему шума Джентри решил с помощью так называемой процедуры перешифрования (bootstrapping). Пусть c — криптограмма некоторого бита m с большим шумом. Проблему уменьшения шума можно решить, если расшифровать c, получить бит m и заново зашифровать его на ключе p. Пусть B — схема в базисе $(+, \cdot)$, выполняющая эту процедуру. Процедура перешифрования выполняет вычисления согласно схеме B гомоморфно: на вход поступают криптограмма c и ключ p, зашифрованный на ключе p.

Еще раз подчеркнем, что здесь мы лишь неформально описываем идею. За конструкцией криптосистемы отсылаем читателя к оригинальной статье [4]. Большинство определений, относящихся к гомоморфному шифрованию, можно найти в статье [8], которая представляет собой не столько обзор, сколько попытку навести порядок в понятиях и терминологии. Полные доказательства стойкости вполне гомоморфных криптосистем — большая редкость. Такое доказательство можно найти в диссертации Джентри [3].

Если невполне гомоморфная криптосистема позволяет выполнять процедуру перешифрования и еще операцию умножения, то вычисления над криптограммами могут продолжаться сколь угодно долго. Таким образом невполне гомоморфная криптосистема преобразуется во вполне (fully) гомоморфную.

Но главная заслуга Джентри — не процедура перешифрования, а доказательство, при некотором предположении, стойкости предложенной криптосистемы.

Теперь вернемся к упомянутым выше «мифам». Утверждение о том, что результат Джентри решает проблему вычислений над зашифрованными данными может быть истинным или ложным в зависимости от того, что понимать под такими вычислениями. Если зашифрованная информация хранится в базе данных и требуется обеспечить стандартный для баз данных набор запросов, то защита такой информации может оказаться невозможной в принципе [5].

Что касается первого «мифа», то здесь следует уточнить, что понимается под гомоморфным шифрованием. Пусть C — произвольная схема в базисе $\{+, \cdot\}$, которая получает на вход биты $m_1, \ldots, m_n$, и пусть $m = C(m_1, \ldots, m_n)$. С математической точки зрения гомоморфным шифрованием следует называть такое отображение битов $m_1, \ldots, m_n$ в криптограммы $c_1, \ldots, c_n$, что применив к этим криптограммам ту же схему C, но с другими операциями сложения и умножения, мы получим криптограмму c, из которой при дешифровании будет получен бит m.

При такой трактовке термина «гомоморфное шифрование» гипотеза о его несуществовании остается неопровергнутой. Конструкции, предложенные Джентри и его последователями, решают (теоретически) проблему вычислений специального вида над зашифрованными данными. А именно, проблему вычисления произвольных функций от зашифрованных данных.

Следует заметить, что как и во всякой недавно возникшей теории, в гомоморфном шифровании на данный момент имеются проблемы с терминологией. Ситуацию усугубляют большое количество работающих в этой области математиков, практически не знакомых с криптографией.

Так, во всех работах много внимания уделяется так называемому требованию компактности гомоморфных криптосистем. Оно означает, что какие бы вычисления мы ни проводили над криптограммами, длины последних остаются ограниченными. Это требование позволяет избежать тривиальности, и в нем нет ничего некорректного ни с математической, ни с криптографической точки зрения.

Но это требование попросту лишнее. Мы имеем дело с криптосистемами вероятностного шифрования, в которых при фиксированных открытом тексте и ключе криптограмма — случайная величина. Но уже в само определение криптосистемы должно входить требование конечности множества криптограмм для любой пары (открытый текст, ключ).

Определение множества криптограмм требует некоторой аккуратности. В это множество должна входить всякая криптограмма, из которой при дешифровании на данном ключе будет получен данный открытый текст.

Если требование конечности множества криптограмм опустить, то проблема гомоморфности давно решена. Примером может служить Poly Cracker [6]. Не вдаваясь в подробности, опишем неформально идею конструкции. Пусть $\{q_i\}$ — набор полиномов и y — такое значение, что $q_i(y) = 0$ для любого i. Тогда криптограмма бита m вычисляется по формуле $\sum g_i q_i + m$, где g_i — случайные полиномы.

Но эта конструкция, и ей подобные,— другой криптографический примитив, который некорректно называть криптосистемой. Возможно, здесь уместен исходный термин «privacy homomorphism».

Самая популярная и устойчивая «легенда», связанная с гомоморфным шифрованием, гласит, что последнее решает проблему облачных вычислений над конфиденциальными данными. На самом деле не решает и не может решить.

В достаточно общей постановке проблема защиты информации в облачных вычислениях такова: имеются n пользователей, которые хранят на облачном сервере свои конфиденциальные данные $m_1, \ldots, m_n$. Требуется вычислять произвольные функции $f(m_1, \ldots, m_n)$ таким образом, чтобы значения $m_1, \ldots, m_n$ оставались конфиденциальными. В работе [7] доказано, что уже в случае двух пользователей такие конфиденциальные вычисления невозможны.

Таким образом, для защиты информации в облачных вычислениях требуется либо расширить модель, дополнив ее новыми сущностями и (или) предположениями, либо ограничить множество вычисляемых функций, либо сделать и то и другое. Пример можно найти все в той же статье [7]: если ограничиться функциями f, существенно зависящими только от одной переменной, то, используя вполне гомоморфное шифрование, можно построить систему облачных вычислений над конфиденциальными данными.

После выхода в свет статьи Джентри появилось немало работ, авторы которых уверены, что поскольку доказано существование вполне гомоморфных криптосистем, главная нерешенная проблема теперь — эффективность таких криптосистем. Разумеется, такое мнение ошибочно. Основная проблема — стойкость предлагаемых конструкций. Если требование стойкости опустить, то проблема эффективности решается тривиальным образом.

Под обоснованием стойкости понимается либо ее доказательство в математической модели (при некоторых предположениях), либо достаточно солидная история попыток криптоанализа, как для используемых на практике криптографических схем. Авторы многих статей хотят жить в лучшем из миров: полагают, что для предлагаемых ими конструкций будет обоснование стойкости второго типа, хотя сами эти конструкции для практики не интересны.

Такова особенность криптографии: трудоемкость разработки новой схемы не сопоставима с трудоемкостью ее криптоанализа. Большинство опубликованных криптографических схем остаются невзломанными просто потому, что их никто всерьез не анализировал.

За годы, прошедшие с момента публикации работы Джентри, несмотря на все усилия, эффективная (с практической точки зрения) конструкция гомоморфной криптосистемы так и не найдена. В работе [8] приводятся данные о производительности экспериментальных реализаций некоторых вполне гомоморфных криптосистем. Эти данные не внушают никакого оптимизма.

Уверенность в том, что через некоторое время, после дополнительных усилий, эффективная конструкция наконец будет найдена, ничем не обоснована.

Вполне вероятно, что все исследования в данной области завершатся доказательством какого-либо отрицательного результата.

В заключение отметим основные направления исследований в области гомоморфного шифрования.

Основная нерешенная проблема — существование (несуществование) эффективных и доказуемо стойких вполне гомоморфных криптосистем.

Остается также открытым и вопрос о существовании вполне гомоморфных криптосистем, не имеющих теоретического обоснования стойкости, но по эффективности сравнимых, скажем, с RSA, т. е. пригодных к практическому применению. Стойкость такой криптосистемы может исследоваться традиционными для практической криптографии методами.

Еще одно направление исследований — вполне гомоморфные криптосистемы с открытыми текстами произвольной длины. Некоторые шаги в этом направлении уже сделаны (см. работу [9]).

Вызывает также интерес поиск новых криптографических предположений для доказательства стойкости вполне гомоморфных криптосистем. На данный момент все такие доказательства основаны на предположении о вычислительной трудности задачи выведования с ошибками (Learning With Erros — LWE).

Есть и такая тема для исследований: какие системы облачных вычислений над конфиденциальными данными можно построить на основе гомоморфных криптосистем?

И, наконец, остается открытой основная проблема: существует ли гомоморфное шифрование в указанном выше строгом математическом смысле.

СПИСОК ЛИТЕРАТУРЫ

1. *Rivest R. L.*, *Adleman L.*, *Dertouzos M. L.* On data banks and privacy homomorphisms. In: Foundations of Secure Computation–Workshop. (Atlanta, GA, October 3–5, 1978.) Proceedings./ Ed. by R. A. DeMillo, D. P. Dobkin, A. K. Jones, R. J. Lipton. N. Y.–Orlando, FL: Academic Press, 1978, p. 160–179.
2. *C. Gentry.* Fully homomorphic encryption using ideal lattices. In: STOC'09 Proceedings of the 2009 ACM International Symposium on Theory of Computing. (Bethesda, MD, May 31–June 02, 2009.) Ed. by M. Mitzenmacher. N. Y.: ACM, 2009, p. 169–178.
3. *Gentry C.* A fully homomorphic encryption scheme. PhD thesis. Stanford, CA: Stanford Univ., 2009, 209 p.
4. *van Dijk M.*,, *Gentry C.*, *Halevi S.*, *Vinod V.* Fully homomorphic encryption over the integers. In: Advances in Cryptology–EUROCRYPT 2010. 29th Annual International Conference on the Theory and Applications of Cryptographic Techniques. (French Riviera, May 30–June 3, 2010.) Proceedings./ Ed. by H. Gilbert. Heidelberg etc.: Springer, 2010, p. 24–43. (Ser. Lect. Notes Comput. Sci. V. 6110.)
5. *Варновский Н. П.*, *Шокуров А. В.* Гомоморфное шифрование. — Тр. ин-та системного программирования РАН, 2007, т. 12, с. 27–36.

6. *Fellows M. R.*, *Koblitz N.* Combinatorial cryptosystems galore! In: Finite Fields: Theory, Applications, and Algorithms. Second International Conference on Finite Fields: Theory, Applications, and Algorithms. (Las Vegas, NV, August 17–21, 1993.) Proceedings. / Ed. by G. L. Mullen, P. J.-Sh. Shiue Providence, RI: AMS, 1994, p. 51–61. (Ser. Contemporary Mathematics. V. 168.)

7. *van Dijk M.*, *Juels A.* On the impossibility of cryptography alone for privacy-preserving cloud computing. — Cryptology ePrint Archive, https://eprint.iacr.org/2010/305.

8. *Armknecht F.*, *Boyd C.*, *C.*, *Gjøsteen K.*, *Jäschke A.*, *Reuter C. A.*, *Strand M.*, A Guide to Fully Homomorphic Encryption. Cryptology ePrint Archive, https://eprint.iacr.org/2015/1192.

9. *Nuida K.*, *Kurosawa K.* (Batch) fully homomorphic encryption over integers for non-binary message spaces, In: Advances in Cryptology–EUROCRYPT 2015. 34th Annual International Conference on the Theory and Applications of Cryptographic Techniques. (Sofia, April 26–30, 2015.) Proceedings. Part I. / Ed. by E. Oswald, M. Fischlin. Heidelberg etc.: Springer, 2015, p. 537–555. (Ser. Lect. Notes Comput. Sci. V. 9056.)