УДК 519.2: 003.26

АКТУАЛЬНЫЕ ПРОБЛЕМЫ КРИПТОЛОГИИ

Ю.С. ХАРИН, И.Б. БЕРЕЖНОЙ

Белорусский государственный университет,
НИИ прикладных проблем математики и информатики,
г. Минск, Республика Беларусь

Введение

История развития криптологии, составными частями которой являются криптография и криптоанализ, сложна и включает 3 этапа [1]:

Этап 1 (с древних времен до 1949 г.) – этап докомпьютерной криптологии.

Этап 2 (1949–1976 гг.) принято отсчитывать с момента публикации 60-страничной статьи «Теория в секретных системах» американским математиком-прикладником К. Шенноном. Следует отметить, что исторически первым строгое математическое обоснование того, что симметричные алгоритмы шифрования с одноразовыми ключами являются абсолютно стойкими, дал еще в июне 1941 г. академик АН СССР В.А. Котельников. Однако данное обоснование было опубликовано в секретном отчете и на протяжении всего XX века так и не было представлено в открытой печати [2].

Этап 3 (1976 г. – настоящее время) принято отсчитывать с момента публикации статьи У. Диффи и М. Хеллмана об асимметричных (двухключевых) алгоритмах шифрования.

Статья посвящена современным актуальным проблемам криптологии.

1. Источники случайности (энтропии)

В криптографии генерация случайности сводится к задаче генерации равномерно распределенной случайной последовательности (РРСП) $x_t \in \{0,1\}$, $t = 1, 2, \ldots$, которая определяется двумя базовыми свойствами:

C_1) для любого n и случайных индексов $1 \le t_1 < \cdots < t_n$ случайные величины $x_{t_1}, \ldots, x_{t_n}$ независимы в совокупности;

C_2) для любого номера t случайная величина x_t имеет равномерное распределение: $\mathrm{P}\{x_t = a\} = 1/2$, $a \in \{0,1\}$.

Генерация РРСП представляет собой важную задачу не только для генерации «гаммы» в поточных криптосистемах, но и в других системах криптографической защиты информации для выработки ключей и векторов инициализации, одноразовых чисел (называемых в англоязычной литературе Nonce) в криптографических протоколах и схемах ЭЦП.

Генераторы РРСП можно разделить на следующие подмножества: 1) аппаратные генераторы; 2) программно-аппаратные генераторы; 3) программные генераторы.

Аппаратные (физические) генераторы РРСП позволяют вырабатывать случайные последовательности на основе физических процессов в объектах, параметры которых меняются хаотически (например, использование полупроводникового лазера позволяет вырабатывать РРСП на скоростях от 160 Гбит/с [3], а шумовой диод ND103L производства Республики Беларусь – на скорости около 0,5 Мбит/с [4]). Теоретически такие процессы непредсказуемы, однако на практике на хаотические процессы постоянно влияют окружающая среда и измерительная аппаратура. Отсюда проистекают **недостатки аппаратных генераторов**: необходимость постоянного контроля свойств получаемой последовательности и невозможность репликации последовательности на другом устройстве классической компьютерной архитектуры.

В целях устранения необходимости в постоянном контроле аппаратных генераторов применяются программно-аппаратные генераторы, в которых происходит дополнительная программная обработка данных с аппаратного источника, что позволяет получить лучшие статистические характеристики и увеличить скорость генерации РРСП. Однако **проблема**

репликации последовательности ограничивает сферу применения и для программно-аппаратных генераторов.

Широко распространенной альтернативой аппаратным генераторам РРСП стало использование программных генераторов, к которым предъявляются следующие два требования «практической криптостойкости» (т.е. криптостойкости при условии «разумных» ограничений у противника на время и вычислительные ресурсы для анализа перехваченного сообщения): 1) выходную последовательность генератора невозможно отличить от РРСП на основании определенного набора статистических тестов; 2) знание некоторого отрезка выходной последовательности не позволяет предсказать последующие элементы последовательности.

Последовательность («гамма»), полученная из такого программного генератора, называется псевдослучайной последовательностью (ПСП). Следует отметить, что так как ПСП программных генераторов полностью определяются начальным значением (называемым в англоязычной литературе seed), то **недостатком таких генераторов является проблема источника энтропии для формирования случайного равномерно распределенного значения seed.** Другими словами, обеспечение «практической криптостойкости» программного генератора ПСП происходит за счет не только математических, но и физических методов, хотя и в сравнительно малом объеме.

Существующие криптографические генераторы ПСП принято классифицировать в два подмножества [1, 5]: 1) простые (элементарные) генераторы, использующие единственную псевдослучайную последовательность для генерации выходной последовательности: а) конгруэнтные генераторы; б) генераторы с регистром сдвига; в) генераторы Фибоначчи; г) генераторы на основе односторонних функций; д) генераторы, основанные на проблемах теории чисел; 2) составные генераторы, использующие несколько «простейших» псевдослучайных последовательностей для генерации выходной последовательности: а) фильтрующие и комбинирующие генераторы; б) композиционные генераторы (каскады); в) генераторы с динамическим изменением законов рекурсии; г) **генераторы с дополнительной памятью [6].**

2. Статистическое тестирование криптографических генераторов случайных и псевдослучайных последовательностей

Для определения статистических свойств выходных последовательностей криптографических генераторов ПСП используются наборы (так называемые батареи) статистических тестов, каждый из которых вычисляет вероятность (P-значение) события, состоящего в том, что «генератор истинно случайных чисел» выработает последовательность с определенной статистикой «хуже», чем у проверяемой. Другими словами, батарея тестов выдает вектор вероятностей события, состоящего в том, что исследуемая последовательность близка к РРСП.

Каждый статистический тест представляет собой решающее правило $d = d(X) \in \{0,1\}$ для проверки гипотезы H_0 о том, что исследуемая последовательность $X = \{x_1, \ldots, x_T\}$ является равномерно распределенной случайной последовательностью против альтернативной гипотезы $H_1 = \bar{H}_0$. Для исследуемой последовательности вычисляется некоторая статистика, распределение вероятностей которой при истинной гипотезе H_0 известно. На основании значения статистики или (P-значения) гипотеза H_0 принимается либо отклоняется.

На текущий момент существует ряд известных батарей тестов: DieHarder [7], NIST STS [8], gjrand [9], testU01 [10].

Проведенный в [1] обзор существующих статистических тестов выявляет следующие **недостатки**: 1) многие из существующих тестов не ориентированы на проверку главного свойства C_1, а лишь частных случаев свойств C_1, C_2, т.е. частных случаев альтернативы $H_1 = \bar{H}_0$; 2) многие из известных тестов построены «эвристически» и вообще не фиксируют

семейство альтернатив H_1; 3) многие тесты не имеют оценок мощности $w = \mathrm{P}\{d = 1 \mid H_1\}$; 4) при включении нескольких тестов в батарею не удается учесть их вероятностную зависимость и оптимизировать «составной» тест.

Следует еще отметить, что в марте 2022 г. вышла статья [11], авторы которой обнаружили прямые ошибки в реализации одного из тестов батареи NIST STS, а также проблемы с вычислением P-значений у данной батареи на последовательностях большой длины (≈ 1 Гбит). В связи с этим в апреле 2022 г. **NIST объявил [12] о начале процесса пересмотра собственной батареи и специально обратил внимание на непригодность текущей версии для оценивания криптографических генераторов ПСП.**

Для преодоления указанных недостатков **предлагается строить статистические тесты *s*-мерной равномерности для заданных семейств альтернатив H_1 с использованием энтропийных профилей [13, 14] и строить асимптотические оценки мощностей известных NIST-тестов [15]**.

Существующие статистические тесты и построенные на их основе батареи выходных последовательностей криптографических генераторов обладают **еще одним недостатком, приводящим к следующему парадоксу**: при увеличении длины $T \to +\infty$ наблюдаемых последовательностей реально существующих генераторов вероятность принятия альтернативы H_1 стремится к единице: $w_T \to 1$. Иначе говоря, «увеличением длины проверяемой последовательности T можно забраковать любой реальный генератор».

Объяснение этого парадокса состоит в следующем. Реальные генераторы не идеальны, они отклоняются от гипотезы H_0 по своему вероятностному распределению на некоторую величину $\varepsilon \neq 0$. Все тесты, включаемые в батареи тестов, как известно, обладают оптимальным свойством проверки **простой** гипотезы H_0 против сложной альтернативы ($H_1 = \bar{H}_0$) – свойством состоятельности: $w_T ::= \mathrm{P}\{d = 1 \mid H_1\} \to 1$. Таким образом, **этот парадокс порожден некорректностью математической модели нулевой гипотезы H_0**. Для того чтобы избежать парадокса, **гипотеза H_0 должна быть сложной: она должна задавать величину допуска ε отклонений от идеальной модели РРСП.** Приведем пример сложной гипотезы об *s*-мерном равномерном распределении двоичной случайной последовательности. Обозначим: $V = \{0,1\}$ – двоичный алфавит, $p_{J_1^s} = \mathrm{P}\{x_1 = j_1, \ \ldots, \ x_s = j_s\}$, $J_1^s = (j_1, \ \ldots, j_s) \in V^s$, – *s*-мерное распределение вероятностей; $\boldsymbol{P} = \left\{ p = \left(p_{J_1^s}\right) : p_{J_1^s} \geq 0, \ \sum_{J_1^s \in V^s} p_{J_1^s} = 1 \right\}$ – множество всевозможных *s*-мерных распределений вероятностей;

$$\boldsymbol{P}_0^{(\varepsilon)} = \left\{ p \in \boldsymbol{P} : \ \sum_{J_1^s \in V^s} \left(p_{J_1^s} - \frac{1}{2^s} \right)^2 \leq \varepsilon^2 \right\}.$$

Сложная гипотеза об *s*-мерной равномерности:

$$H_0^{(\varepsilon)} = \left\{ p = \left(p_{J_1^s}\right) \in \boldsymbol{P}_0^{(\varepsilon)} \right\}.$$

Отметим, что при $\varepsilon \to 0$ получаем ранее рассмотренную простую гипотезу H_0:

$$H_0^{(\varepsilon)} \to H_0 .$$

Заметим еще, что при $s = 1$ гипотеза $H_0^{(\varepsilon)}$ принимает вид:

$$H_0^{(\varepsilon)} = \left\{ \left| \mathrm{P}\{x_t = 1\} - \frac{1}{2} \right| = \left| \mathrm{P}\{x_t = 0\} - \frac{1}{2} \right| \le \frac{\varepsilon}{\sqrt{2}} \right\}.$$

Удается построить критерий обобщенного отношения правдоподобия для проверки сложных гипотез $H_0^{(\varepsilon)}$, $H_1 = \bar{H}_0^{(\varepsilon)}$.

3. Искусственные нейронные сети и машинное обучение в криптологии

Нейронные сети и машинное обучение составляют основу искусственного интеллекта. В сжатом виде в истории развития искусственных нейронных сетей (ИНС) можно выделить 3 этапа. **Этап I** (1940–1960 гг.) развивался под лозунгом «кибернетика» и был направлен на построение математической модели биологического нейрона в работах McCulloch, Pitts (1943 г.), Hebb (1949 г.), Rosenblatt (1958 г.). **Этап II** (1961–1990 гг.) развивался под лозунгом 1-слойных и 2-слойных ИНС и был направлен на разработку алгоритмов их обучения (Rummelhart, 1986 г.). **Этап III** (1991 г. – настоящее время) развивается под лозунгом «deep learning – глубокое обучение многослойных нейронных сетей».

Отметим, что теория ИНС базируется на решенной в 1957 г. академиком А.Н. Колмогоровым 13-ой проблемы Гилберта о композициях функций.

Анализ литературы позволяет выявить следующие направления использования ИНС и машинного обучения в решении актуальных задач криптологии: конструирование криптосистем на основе ИНС; классификация зашифрованного трафика; стеганоанализ; криптоанализ алгоритмов шифрования; криптоатаки на основе ИНС и машинного обучения; обнаружение вторжений и аномалий в сетях; обнаружение и анализ вредоносов.

В 1991 г. основатель RSA P. Райвест на конференции ASIACRYPT'1991 представил приглашенный доклад «Криптография и машинное обучение», в котором впервые была указана необходимость использования ИНС в криптографии. В качестве примера использования ИНС в криптологии приведем информацию о результатах, полученных M.M. Alani [16]: успешно выполнен криптоанализ алгоритма DES на основе выборки 2^{11} пар (и 3-DES на основе выборки 2^{12} пар) «открытый текст – шифртекст» с одним и тем же ключом при затратах машинного времени 51 мин (72 мин) с использованием 4-слойной нейронной сети (256 нейронов в первом слое). Атака имеет цель обучить ИНС для восстановления блока открытого текста по доступному блоку шифртекста без восстановления ключа.

Еще одна новая актуальная проблема, связанная с ИНС и с машинным обучением, – **доверенный искусственный интеллект**.

4. Постквантовая криптография

Стойкость основных на сегодня асимметричных (с открытым ключом) криптографических платформ IFC (Integer Factorization Crypto, криптография на основе задачи факторизации), FFC (Finite Field Crypto, криптография в конечных полях), ECC (криптография на эллиптических кривых) основана на вычислительной сложности лежащих в их основе математических задач. На классическом компьютере эти задачи могут быть решены в лучшем случае за субэкспоненциальное время. Ситуация кардинально изменится, если будет построен промышленный квантовый компьютер. С его помощью целевые задачи можно будет **эффективно решить за полиномиальное время. Симметричная (с секретным ключом) криптография также окажется под ударом: фактическое сокращение примерно в два раза битовых длин ключей.**

Криптографическое сообщество активно прорабатывает **запасные криптографические платформы, которые сохранят стойкость даже после «квантового апокалипсиса»**. Такие платформы принято называть постквантовыми. Основные из них: HBC (Hash-Based Crypto, **криптография на основе функций хеширования**), CBC (Code-Based Crypto, **криптография на основе задач теории кодирования**), LBC (Lattice-Based Crypto, **криптография на основе теории решеток**), MBC (Multivariate-polynomial-Based Crypto, **криптография на основе многочленов от нескольких переменных**).

В 2017 г. в США под эгидой Национального института стандартов и технологий (NIST) начался конкурс по разработке постквантовых алгоритмов и протоколов. В настоящий момент конкурс близится к завершению. В четвертый, заключительный, раунд вышли 5 систем платформы LBC (3 ШОК и 2 ЭЦП) и по одному представителю платформ CBC и MBC, т.е. LBC является безоговорочным лидером [17].

5. Проблемы совершенствования подготовки кадров в области информационной безопасности (ИБ)

• расширение перечня специальностей и их направлений для удовлетворения растущего спроса в кадрах по ИБ;

• привлечение специалистов-практиков для организации учебных, производственных, преддипломных практик и проведения учебных занятий в университетах с целью укрепления связи образования в области ИБ с практической деятельностью;

• формирование государственного кадрового заказа на специалистов по ИБ, особенно в интересах силовых министерств и комитетов, органов госуправления;

• привлечение организаций, являющихся потребителями выпускников-специалистов по ИБ, к укреплению материально-технической базы и финансированию научных исследований по ИБ в университетах;

• мониторинг использования специалистов по ИБ;

• согласование Перечней специальностей подготовки кадров высшей квалификации в РБ и РФ в области ИБ.

Литература

1. Математические и компьютерные основы криптологии. / Ю.С. Харин, В.И. Берник, Г.В. Матвеев, С.В. Агиевич. – Мн.: Новое знание, 2003. – 413 с.
2. Молотков С. Квантовая криптография и теоремы В.А. Котельникова об одноразовых ключах и об отсчетах. / С. Молотков // Успехи физических наук. – 2006. – Т. 176, № 7. – С. 777–788.
3. Li Nianqiang Two approaches for ultrafast random bit generation based on the chaotic dynamics of a semiconductor laser. / N. Li, B. Kim, V.N. Chizhevsky, A. Locquet, M. Bloch, D.S. Citrin. W. Pan // Opt. Express. – 2014. – Vol. 22(6). – P. 6634–6646.
4. Пикуза М.О., Михневич С.Ю. Тестирование аппаратного генератора случайных чисел при помощи набора статистических тестов NIST. / М.О. Пикуза, С.Ю. Михневич // Доклады БГУИР. – 2021. – Т. 19 (6). – С. 37–42.
5. Основы криптографии / А.П. Алферов, А.Ю. Зубов, А.С. Кузьмин, А.В. Черемушкин. – М.: Гелиос АРВ, 2005. – 480 с.
6. Бережной И.Б., Харин Ю.С. Вероятностная модель динамики памяти криптографических генераторов Макларена – Марсальи. / И.Б. Бережной, Ю.С. Харин // Информатика. – 2014, № 1. – С. 105–115.
7. Brown, R. Dieharder: A Random Number Test Suite. – [Electronic resource]. http://www.phy.duke. edu/~rgb/General/rand_rate.php.
8. NIST SP 800-22: Download Documentation and Software. – [Electronic resource]. https: //csrcnist. gov/projects/random-bit-generation/document at ion-and-software.
9. Jones, G. Gjrand random numbers official site. – [Electronic resource]. http://gj rand. sourcef orge. net/.
10. L'Ecuyer P. TestUOl – Empirical Testing of Random Number Generators. – [Electronic resource]. http://simul.iro.umontreal.ca/ testuOl/tuOl.html.
11. Kowalska K.A. On the revision of NIST 800-22 Test Suites. – [Electronic resource]. https: //eprint.iacr.org/2022/540.pdf.
12. Decision to Revise NIST SP 800-22 Rev. 1a. – [Electronic resource]. https://csrc.nist.gov/news/2022/ decision-to-revise-nist-sp-800-22-rev-1a.
13. Программный комплекс для энтропийного анализа дискретных последовательностей. / В. Палуха, Ю. Харин, М. Мальцев [и др.] // Информационные системы

и технологии = Information Systems and Technologies: материалы междунар. науч. конгресса по информатике. – Т. 1. – 2022. – С. 102–107.

14. Палуха В.Ю., Харин Ю.С. Тестирование криптографических генераторов случайных и псевдослучайных последовательностей на основе энтропийных профилей. // Сб. конф. Комплексная защита информации. – Москва: Авангард. – 2022. – С. 117–122.

15. Волошко В.А., Трубей А.И. О мощности тестов многомерной дискретной равномерности, используемых для статистического анализа генераторов случайных последовательностей. / В.А. Волошко, А.И. Трубей // Журнал Белорусского государственного университета. Математика. Информатика. – 2022, № 1. – С. 26–37.

16. Alani M.M. Neuro-cryptanalysis of DES and triple-DES / M.M. Alani // Int. Conf. on Neural information processing. – N.Y.: Springer. – 2012. – P. 637–646.

17. Матвеев Г.В. Аналитический обзор методов и алгоритмов постквантовой криптографии. – Научно-технический отчет по НИР. – Минск: НИИ ППМИ БГУ, 2022. – 56 с.