ИЗВЕСТИЯ РАН. СЕРИЯ ФИЗИЧЕСКАЯ, 2019, том 83, № 3, с. 420–423

УДК 530.145.1

СТАТИСТИЧЕСКАЯ ОЦЕНКА ФИЗИЧЕСКОГО ГЕНЕРАТОРА СЛУЧАЙНЫХ ЧИСЕЛ

© 2019 г. Н. С. Перминов[1, 2, *], О. И. Банник[1], Л. Р. Гилязов[1], К. С. Мельник[1], Д. Ю. Таранкова[1, 3]

[1]Казанский квантовый центр, Федеральное государственное бюджетное образовательное учреждение высшего образования "Казанский национальный исследовательский технический университет им. А.Н. Туполева-КАИ" (КНИТУ-КАИ), Казань, Россия

[2]Казанский физико-технический институт им. Е.К. Завойского – обособленное структурное подразделение Федерального государственного бюджетного учреждения науки "Федеральный исследовательский центр "Казанский научный центр Российской академии наук", Казань, Россия

[3]Институт радиоэлектроники и телекоммуникаций, Федеральное государственное бюджетное образовательное учреждение высшего образования "Казанский национальный исследовательский технический университет им. А.Н. Туполева-КАИ" (КНИТУ-КАИ), Казань, Россия

*E-mail: nperminov@kazanqc.org

Поступила в редакцию 03.09.2018 г.
После доработки 10.09.2018 г.
Принята к публикации 22.10.2018 г.

Изучены методы робастного статистического анализа физических генераторов случайных чисел. Предложена оценка меры случайности на основе метода ранжированных амплитуд, дополняющая энтропийный анализ. Сделан анализ качества случайности исходных данных с аналого-цифрового преобразователя для квантового генератора случайных чисел на основе гомодинного детектирования.

DOI: 10.1134/S036767651903030X

ВВЕДЕНИЕ

Криптография – одна из старейших прикладных наук о контроле секретных свойств информации, история которой насчитывает несколько тысяч лет. Именно в ней понятие "полезной случайности" постепенно выросло в отдельное научное направление, связанное с фундаментальным изучением понятий хаоса и предсказуемости. На данный момент в мире существует огромное количество разных подходов к количественному определению хаотических и случайных явлений, многие из которых в своем корне опираются на меру сравнения набора из данных друг относительно друга. Именно здесь кроются проблемы окончательного осмысления таких важных понятий как "измерение", "повторяемость" и "закон природы", заложенных математиками-механиками в физику около 300 лет назад. Несмотря на видимые значительные успехи в техническом развитии цивилизации за эти века, сами основы анализа и обработки данных в прикладных естественных науках не получили достаточного внимания со стороны исследователей-экспериментаторов. В этой работе мы хотим продемонстрировать читателю применение новых прикладных методов анализа физических данных, кардинально меняющих взгляд на проблему измерения, случайности и хаоса в естественных науках.

Наше исследование посвящено изучению случайности в рамках высокоскоростной квантовой криптографии на непрерывных переменных [1–4], где требуется большой расход случайных чисел. По сути, системы квантовой криптографии представляют собой распределенную систему согласования двух независимых случайных последовательностей на передающей и приемной сторонах при помощи квантовых состояний. Сами случайные последовательности реализуются с помощью генераторов случайных чисел (ГСЧ). При разработке ГСЧ для квантовой криптографии недостаточно того, что тесты на случайность по некоторому критерию были пройдены, что является лишь необходимым условием. Принципиально важен источник первичной случайности [1], который используется для получения равнораспределенной последовательности 0 и 1 и который был бы источником случайности по другим физическим соображениям [4], не зависящим от тестов. Многие псевдослучайные ГСЧ проходят тесты, но не являются, очевидно, истинно случайны-

ми. Все ГСЧ можно разделить на два класса. Первый класс – физические генераторы, когда случайная последовательность извлекается из некоторого физического процесса. Второй класс – математические ГСЧ, когда случайная последовательность получается как результат математического преобразования, как правило, рекуррентного, некоторого затравочного числа. Любой математический генератор выдает псевдослучайную последовательность, которая полностью предсказуема, если известны исходные затравочные числа.

Поэтому процесс генерации случайных чисел на основе результатов измерений физических величин (физический ГСЧ) на существующий момент является единственным источником случайности, который признан научным миром надежным для использования в квантовых коммуникациях. Этот процесс разделяется на 2 стадии: измерение (вместе с получением цифровых данных) и затем экстракция битовой последовательности. Экстракция конечной битовой последовательности, которая будет проходить известные тесты на случайность, достаточно хорошо изучена, и многочисленные эффективные алгоритмы экстракции присутствуют в огромном количестве в свободной печати [5, 6].

В этой работе мы делаем акцент на изучении первой стадии, то есть на изучении качества исходных сырых данных, полученных в ходе измерений физической системы. Нами выявлены новые дискретно-интегральные характеристики шумов и предложена робастная оценка меры случайности на основе метода ранжированных амплитуд, дополняющая традиционный энтропийный анализ Шеннона–Реньи в криптографии. Сделан анализ качества случайности исходных данных для квантового ГСЧ на основе гомодинного детектирования.

ФИЗИЧЕСКИЙ ГСЧ И ИДЕЯ ЕГО АНАЛИЗА

Основным вопросом, решаемым в рамках построения физического ГСЧ на стадии измерения, остается количественный вопрос о “качестве” исходных цифровых данных, полученных с аналого-цифрового преобразователя (АЦП) устройства. Статистические меры “качества” в криптографии, стеганографии и анализе временных рядов сильно различаются. Однако обеспечение физической случайности ГСЧ на первой стадии связано в первую очередь с мерой совпадения экспериментальных данных с теоретически предсказуемыми в пространстве состояний АЦП или его расширенном пространстве (случай более идеального АЦП). Такое сравнение невозможно на уровне стандартных энтропийных мер Шеннона–Реньи, к которым прибегают все известные авторам методы реализации физических ГСЧ.

Поэтому мы предлагаем использовать более естественную меру случайности, основанную на статистике дробных моментов, построенных на основе последовательности ранжированных амплитуд (ПРА) [7–9]. Метод ПРА, называемый в математической литературе порядковыми статистиками, является основным промежуточным звеном трансформации данных, которое включает в себя всю исходную статистическую информацию без потерь (неинвазивность). ПРА-выборки $\{s_k\}$ – это ее упорядоченная выборка $\mathrm{SRA}[\{s_k\}] = \{x_k\}$, где $x_k \geq x_{k+1}$. Для случайных последовательностей самой широкой природы, не имеющих явно выраженного тренда и заранее известной физически или математически детерминированной природы, ПРА-анализ является единственно возможным достоверным статистическим методом гладкого анализа дискретной случайности.

КАЧЕСТВО ГСЧ И ИНТЕГРАЛЬНАЯ СТАТИСТИКА

Дискретное интегрирование и центрирование (вычитание среднего) известны в дискретной статистике как одни из немногих методов локального неинвазивного анализа, позволяющих эффективно отобразить хаотический ряд чисел во множество с уникальным трендом. В математике многократное интегрирование и дифференцирование исходных функций системы позволяют построить ее расширенное фазовое пространство (теория струй), где “законы” (дифференциальные уравнения) приобретают максимально простой вид. На этом пути были достигнуты значительные успехи, связанные с построением формальных решений дифференциальных уравнений, что в корне изменило возможности анализа теоретической физики. Дискретным аналогом такой схемы является изучение рекуррентных соотношений. На уровне прикладной статистики интересным становится изучение трендов многократных дискретных интегралов от данных с целью их идентификации.

Для экспериментальных данных, полученных на основе гомодинного детектирования лазерного излучения, мы построили несколько выборок и задались целью их сравнить. Мы взяли 2 подряд идущих подвыборки и вычли их друг из друга. Для полученной разницы, которая номинально также должна быть случайной величиной, мы 3 раза провели обратимую операцию дискретного интегрирования и центрирования. В результате мы получили кривую y_n, показанную на рис. 1, имеющую явный синусоидальный

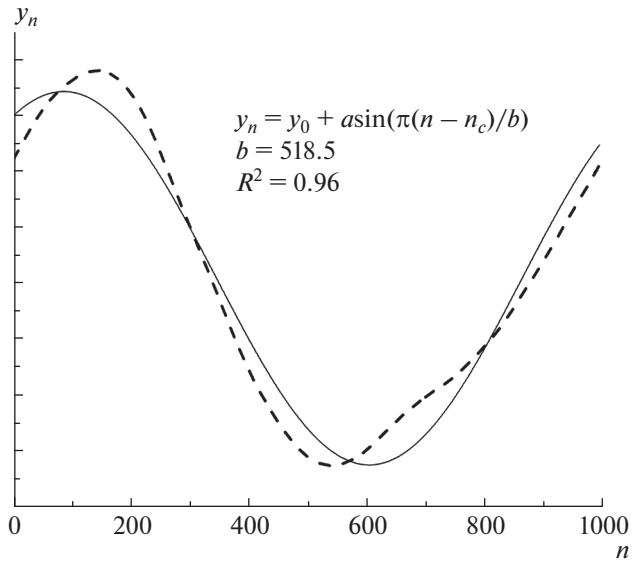

Рис. 1. Типичный результат тройного дискретного интегрирования y_n разницы двух коротких выборок (масштаб для наглядности опущен): штриховая линия – эксперимент, сплошная линия – явно выраженный синусоидальный тренд.

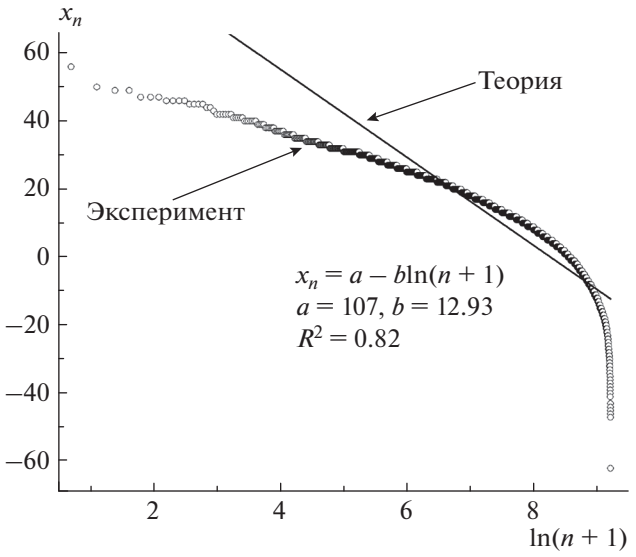

Рис. 2. ПРА данных с АЦП после гомодинного детектора: кружками отмечены экспериментальные данные, прямая линия – теория (отвечает распределению Пуассона). Качество данных, определяемое коэффициентом детерминации R^2, составляет 0.82.

тренд. Этот факт говорит об отсутствии полной случайности и наличии не выявленных пока сложных закономерностях. Эти тренды интегральных кривых допускают количественную параметризацию, являющуюся уникальным идентификатором физического ГСЧ [8]. Дискретно-интегральные характеристики шумов также могут быть использованы в качестве аналогов хэш-функций для первичных данных физических ГСЧ.

ПРА-МЕРА КАЧЕСТВА ДАННЫХ

По сравнению с общеизвестными информационными энтропийными мерами Шеннона–Реньи, предложенные техники работают на естественном фазовом пространстве физической измеряемой системы для набора дискретных данных. При увеличении разрядности АЦП и качества установки вид и форма ПРА-кривых останутся практически неизменными, в отличие от энтропии Шеннона–Реньи, которая больше подходит для определения качества затравочной случайности абстрактных чисел (не привязанных к фазовому пространству физической системы) в схемах с каскадной генерацией случайных чисел.

Самым простым и прозрачным методом сравнения для ПРА-кривых [9] может служить коэффициент детерминации R^2, определяющий точность совпадения двух кривых (или точность фитинга при сравнении эксперимента с гипотезой, отвечающей некоторой модельной функции с параметрами). По изложенной выше схеме мы получили ПРА исходных данных с АЦП (целые числа от –75 до 75) для квантового генератора случайных чисел на основе гомодинного детектирования. Предполагалось, что ПРА-распределение данных с АЦП имело пуассоновский характер [10, 11], то есть $x_n = a - b^*\ln(n + 1)$.

В результате фитинга ПРА [3, 4] пуассоновским распределением (сравнение с пуассоновской статистикой) для $N = 1000$ точек мы получили $R^2 = 0.82$, что определяет в нашем случае меру качества экспериментальных данных по отношению к идеальным для ГСЧ данным (см. рис. 2). То есть около 82% сырых данных с АЦП могут считаться “хорошими” (надежными) для конвертации в конечную битовую последовательность, что говорит о достаточно хорошем качестве квантового ГСЧ, основанного на гомодинном детектировании.

ЗАКЛЮЧЕНИЕ

ПРА-анализ и многократное дискретное интегрирование являются удобными робастными инструментами для быстрого прикладного анализа качества первичной случайности, заложенной в физический ГСЧ. По сравнению с известными информационными энтропийными мерами Шеннона–Реньи, предложенные техники работают на естественном фазовом пространстве физической измеряемой системы для набора дискретных данных и являются гладкими функциями. Более

того, тренды ПРА и интегральных кривых допускают количественную параметризацию, являющуюся уникальным идентификатором физического ГСЧ.

При анализе экспериментальных данных для квантового ГСЧ на основе гомодинного детектирования нами выявлены новые дискретно-интегральные характеристики шумов, показывающие наличие уникальных трендов, что может быть использовано в качестве аналогов хэш-функций для первичных данных физических ГСЧ. Предложена робастная оценка качества случайности на основе сравнения ПРА-кривой с идеальной ПРА-кривой для пуассоновского процесса, которая дополняет известные оценки качества случайности. Изученные методы также могут быть использованы в квантовых коммуникациях в целях мониторинга и для обеспечения большей безопасности.

Работа поддержана грантом молодых ученых РТ № 06-36-ц-Г 2018 “Безопасность оптических и квантовых коммуникаций” (рук. Перминов Н.С.) (основная идея и численное моделирование), а также грантом Правительства Российской Федерации, проект № 14.Z50.31.0040, 17 февраля 2017 года (эксперимент).

СПИСОК ЛИТЕРАТУРЫ

1. *Andersen U.L., Leuchs G., Silberhorn C.* // Las. and Photonics Rev. 2010. V. 4. No. 3. P. 337.
2. *Andersen U.L., Neergaard-Nielsen J.S., Van Loock P. et al.* // Nat. Phys. 2015. V. 11. No. 9. P. 713.
3. *Gabriel C., Wittmann C., Sych D. et al.* Nat. Photon. 2010. V. 4. No. 10. P. 711.
4. *Raffaelli F., Ferranti G., Mahler D.H. et al.* // Quant. Sci. Techn. 2018. V. 3. No. 2. Art. no. 025003.
5. *Trevisan L.* // J. of ACM. 2001. V. 48. No. 4. P. 860.
6. *Konig R., Renner R., Schaffner C.* // IEEE Trans. Inform. Theory. 2009. V. 55. P. 4337.
7. *Nigmatullin R.R., Smith G.* // Phys. A. 2003. V. 320. P. 291.
8. *Nigmatullin R.R.* // Commun. in Nonlin. Sci. Numer. Simulation. 2010. V. 15. No. 3. P. 637.
9. *Nigmatullin R.R., Budnikov H.K., Sidelnikov A.V. et al.* Comp. Commun. and Collaboration. 2017. V. 5. No. 3. P. 12.
10. *Smirnov M.A., Perminov N.S., Nigmatullin R.R. et al.* Appl. Opt. 2018. V. 57. No. 1. P. 57.
11. *Перминов Н.С., Смирнов М.А., Нигматуллин Р.Р. и др.* // Комп. оптика. 2018. V. 42. № 2. С. 338; *Perminov N.S., Smirnov M.A., Nigmatullin R.R. et al.* // Comp. Opt. 2018. V. 42. No 2. P. 338.