Министерство образования Республики Беларусь
Учреждение образования
Белорусский государственный университет
информатики и радиоэлектроники

УДК 681.7.078

Тэт Наунг

Тестирование физического квантового генераторов случайных чисел

АВТОРЕФЕРАТ

на соискание степени магистра наук
по специальности 7-06-0713-03 «Радиосистемы и радиотехнологии»

Научный руководитель Михневич Светлана Юрьевна
канд. физ.-мат. наук, доцент

Минск 2025

ВВЕДЕНИЕ

Генераторы случайных чисел играют критически важную роль в современной цифровой экосистеме. Их использование охватывает широкий спектр приложений: от криптографической защиты и аутентификации до статистического моделирования, научных симуляций и генерации тестов. Как указывается в рекомендациях NIST SP 800-90B, надёжная генерация случайных битов является базовой основой для построения стойких криптографических систем, особенно в условиях активного противостояния с потенциальным злоумышленником, способным воспользоваться предсказуемостью генератора для компрометации ключей и протоколов. [1]

Квантовая механика предлагает интересные новые протоколы на стыке компьютерной науки, телекоммуникаций, теории информации и физики. Такие результаты, как протоколы для квантового распределения ключей BB84 и эффективные алгоритмы для задач, которые считаются или известны как сложные для классических, показывают, что квантовая физика может оказать большое влияние на криптографию.

Важной и хорошо зарекомендовавшей себя квантовой технологией является квантовая генерация случайных чисел. Квантовые генераторы случайных чисел (квантовые ГСЧ) — это устройства, которые используют квантово-механические эффекты для генерации случайных чисел и имеют области применения от моделирования до криптографии. [2]

В рамках магистерской диссертации проведено исследование физического КГСЧ, соотнесение его с классом ГСЧ по типу энтропии (IID или не-IID). Проведен коллизионный тест источника энтропии рассматриваемого генератора.

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы магистерской диссертации

Исследование физических генераторов бинарных случайных последовательностей актуально, поскольку псевдослучайные (алгоритмические) генераторы случайных чисел работают по определенным алгоритмам, которые потенциально могут быть обращены. Вместе с тем, согласно рекомендациям NIST, при работе с физическими ГСЧ необходимо проводить тестирование не только выходной последовательности, но и энтропии источника случайности. Тесты для энтропии отличаются для

различных физических ГСЧ в зависимости от того какой источник случайности использует данный ГСЧ. По типу источника энтропии различаются ГСЧ с IID или к не-IID источниками энтропии.

Целью диссертации является проведение тестов для источника энтропии физического КГСЧ.

Задачей исследования было разработка тестов на отнесение физического квантового ГСЧ к IID или к не-IID источникам. Разработка тестов и тестирование энтропии физического КГСЧ.

Результаты данной работы докладывались на конференциях и отражены в тезисах докладов.

Полученные результаты могут быть применены для тестирования физических ГСЧ.

Связь работы с приоритетными научными направлениями

Работа соответствует пункту 1 «Цифровые информационно-коммуникационные и междисциплинарные технологии, основанные на них производства» Указа Президента Республики Беларусь от 7 мая 2020 г. № 156 «О приоритетных направлениях научной, научно-технической и инновационной деятельности на 2021-2025 годы».

Личный вклад соискателя

Соискателем опубликованы два тезиса и представлены доклады на двух конференциях, готовиться к печати статья.

КРАТКОЕ СОДЕРЖАНИЕ РАБОТЫ

Глава 1

Литературный обзор

Эта глава закладывает концептуальную основу моей диссертации. Я начинаю с обсуждения роли и важности генераторов случайных чисел (ГСЧ), особенно в безопасных вычислениях и криптографии.

1.1 Важность ГСЧ и области применения

В этом разделе я рассматриваю, почему случайность необходима в:

- *Криптография*: для генерации ключей, идентификаторов сеансов, защищенных токенов.

- *Симуляции и моделирование:* особенно методы Монте-Карло.
- Протоколы аутентификации и генерация цифровой подписи.

Я также изучаю уязвимости генераторов псевдослучайных чисел (ГПСЧ), которые могут быть использованы, если начальное значение известно или предсказуемо, что приводит к воспроизводимости и потенциальным векторам атак.

1.2 PRNGs vs. TRNGs

Здесь я описываю архитектурные и функциональные различия между PRNG (основанными на детерминированных алгоритмах) и генераторами истинных случайных чисел (TRNG), которые выводят энтропию из физических явлений, таких как тепловой или квантовый шум. Сравнительная таблица приведена для демонстрации основных различий:

Особенность	PRNG	TRNG
Источник энтропии	Алгоритмический (на основе семени)	Физический (квантовый/тепловой шум)
Воспроизводимость	Да	Нет
Предсказуемость	Высокий, если известно семя	Очень низкий
Скорость	Быстрый	Помедленнее
Проверка энтропии	Необязательный	Обязательный

1.3 Тестирование генераторов случайных чисел

Я представляю подробное сравнение того, как тестируются PRNG и TRNG:

- Генераторы псевдослучайных чисел обычно тестируются с использованием статистических пакетов, таких как NIST SP 800-22, DieHard или TestU01.
- Для TRNG требуется физическая проверка источника шума, тестирование работоспособности и оценка энтропии, как указано в NIST SP 800-90B.

Я также обсуждаю, как квантовые ГСЧ (QRNG) становятся наиболее перспективным типом TRNG, поскольку они основаны на присущей квантовым измерениям неопределенности.

Глава 2

Тестирование источника случайности и энтропии IID

Эта глава переходит от теории к практике. Я документирую свою реализацию и оценку физического QRNG, особенно проверку того, можно ли считать его источник энтропии IID, и оценку его минимальной энтропии, что имеет решающее значение для криптографической валидности.

2.1 Источники энтропии в физических QRNG

Я объясняю, как физическая случайность возникает из квантовых явлений и чем она отличается от псевдослучайности. Энтропия определяется как мера неопределенности: чем непредсказуемее поток битов, тем выше энтропия.

Я представлю типичные компоненты TRNG:

- Источник энтропии (квантовое событие)
- Дигитайзер (преобразует аналоговый сигнал в двоичный)
- Функции кондиционирования (например, криптографический хэш)
- Тесты работоспособности (для надежности выполнения)

2.2 Понимание предположения IID

Здесь я обсуждаю, что означает, что данные являются IID:

- *Независимый:* ни одно значение не зависит от предыдущих.
- *Одинаково распределено:* все значения имеют одинаковое распределение.

Это предположение упрощает оценку энтропии и требуется для большинства криптографических протоколов для обеспечения теоретических гарантий безопасности.

2.3 Статистическое тестирование для IID

В этом разделе я провел подробное статистическое тестирование выходных данных QRNG, чтобы оценить, можно ли классифицировать образцы как независимые и одинаково распределенные (IID) — критическое требование для точной оценки энтропии и безопасного использования в криптографических системах. Эти тесты основаны на процедурах, описанных в NIST SP 800-90B, который включает в себя полный набор тестов как на основе перестановок, так и на основе хи-квадрат.

Тестирование перестановок — это статистический метод, используемый для оценки того, является ли наблюдаемый результат значимым или просто случайным явлением. Он работает путем многократного перемешивания набора данных — обычно 10 000 раз — для создания новых, рандомизированных версий. Для каждого перемешанного набора данных вычисляется тестовая статистика (например, средняя разность или корреляция). Затем реальная тестовая статистика, основанная на исходном наборе данных, сравнивается с распределением тестовой статистики из перемешанных наборов данных. Если исходное значение является экстремальным по сравнению со случайными. это говорит о том, что наблюдаемая закономерность имеет статистический смысл, а не является следствием случайности.

2.3.1 Permutation-Based Tests

1. Экскурсионная тестовая статистика

Измеряет максимальное отклонение текущей суммы значений выборки от среднего. Большое отклонение указывает на тенденции или смещения в выходных данных.

2. Количество направленных пробегов

Подсчитывает количество увеличивающихся или уменьшающихся переходов в последовательности. Слишком мало или слишком много направленных пробегов могут указывать на структурированные шаблоны или зависимость.

3. Длина направленных пробегов

Оценивает самый длинный ряд строго возрастающих или убывающих значений. Длинные направленные ряды могут подразумевать эффекты памяти или систематические тенденции.

4. Количество увеличений и уменьшений

Отслеживает, как часто значение увеличивается или уменьшается относительно предыдущего. Значительный дисбаланс может быть признаком неслучайного поведения.

5. Количество пробегов на основе медианы

Создает бинарную последовательность на основе того, находятся ли значения выше или ниже медианы, и подсчитывает количество запусков. Этот тест оценивает баланс и изменчивость вокруг центрального значения.

6. Длины пробегов на основе медианы

Как и предыдущий тест, но измеряет длину самого длинного такого прогона. Длинные прогоны могут подразумевать перекос или отсутствие случайности в распределении.

7. Средняя статистика теста на столкновение

Определяет, сколько образцов обычно требуется, прежде чем то же самое значение будет наблюдаться снова. Частые ранние столкновения могут указывать на низкую энтропию или небольшое внутреннее пространство состояний.

8. Максимальная статистика теста на столкновение

Регистрирует самый длинный интервал между дублирующимися значениями, наблюдаемыми в данных. Очень короткие интервалы могут отражать плохое рассеивание энтропии.

9. Статистика теста на периодичность

Проверяет повторяющиеся значения на определенных фиксированных расстояниях (лагах), например, 1, 2, 8, 16, 32 образца. Регулярное повторение может сигнализировать о цикличном или детерминированном поведении.

10. Статистика теста на ковариацию

Измеряет статистическую корреляцию между значениями, разделенными фиксированным лагом. Ненулевая ковариация предполагает временную зависимость или структурное смещение.

11. Статистика теста на сжатие

Применяет алгоритм сжатия общего назначения (например, bzip2) к данным выборки. Если последовательность сжимаема, она содержит шаблоны или избыточность, что нежелательно в действительно случайных данных.

12. Проверка независимости для двоичных данных

Оценивает, встречаются ли соседние пары битов (00, 01, 10, 11) с ожидаемой частотой.

13. Тестирование соответствия для двоичных данных

Разделяет двоичные данные на 10 фрагментов и проверяет, что каждый фрагмент имеет одинаковую пропорцию 0 и 1. Большие различия указывают на нестабильность или непоследовательность в производстве энтропии.

14. Длина самого длинного повторяющегося подстрочного теста

Поиск самой длинной последовательности, которая встречается в наборе данных более одного раза. Наличие длинных повторяющихся подстрок является сильным индикатором детерминированного или шаблонного вывода.

Каждый из этих тестов предназначен для обнаружения различных видов неслучайности — будь то тенденции, смещение, периодичность, память или избыточность. Я реализовал эти тесты на Python, применил их к собранным данным битового потока и задокументировал статус прохождения/неудачи каждого теста вместе с числовыми значениями достоверности, где это применимо.

Результаты этих тестов сыграли решающую роль в определении того, можно ли рассматривать выходные данные QRNG как IID, что, в свою очередь, повлияло на выбор оценок энтропии, используемых в последующих разделах диссертации.

2.4 Реализация теста на основе Python

Я реализовал фреймворк тестирования IID на Python:

- Мои скрипты считывают битовые потоки из QRNG.
- Данные предварительно обрабатываются и очищаются (например, удаляются недвоичные символы).
- Каждый тест выполняется с выводом результатов, включающих в себя уровень прохождения/непрохождения и уровень статистической значимости.

Я продемонстрировал это с помощью данных, сгенерированных QRNG, и задокументировал, как результаты испытаний помогают проверить качество генератора.

2.5 Оценка столкновений в энтропийном анализе

Оценка столкновений — один из рекомендуемых методов в стандарте NIST SP 800-90B для оценки энтропии генератора случайных чисел, когда выходные данные не гарантированно независимы или равномерно распределены.

Этот метод основан на простой, но мощной идее: если я случайным образом извлекаю значения из источника данных, сколько времени пройдет, прежде чем я снова увижу то же самое значение? При оценке энтропии, чем раньше повторяется значение, тем меньше энтропия (непредсказуемость) у источника.

- Я взял длинную последовательность выборок из источника энтропии (например, блоки по 8 бит каждый).
- Оценщик отслеживает каждое уникальное значение, полученное на данный момент.
- Он измеряет количество образцов, собранных до того, как произойдет первый дубликат — это называется «столкновение».
- Я повторил этот процесс несколько раз, чтобы получить набор расстояний столкновений.
- Из этого набора я вычислил статистическое среднее значение и использовал его в консервативной формуле, которая оценивает минимальную энтропию.

Оценка столкновений предполагает, что источник с высокой энтропией будет производить много уникальных значений перед повторением, в то время как источник с низкой энтропией будет быстро генерировать повторяющиеся значения. Важно отметить, что этот метод не предполагает, что данные являются совершенно случайными или следуют известному распределению, что делает его хорошо подходящим для реальных физических источников, таких как QRNG.

Я применил эту технику с помощью скриптов Python для анализа потоков битов, сгенерированных QRNG. Каждый образец рассматривался как блок фиксированной длины (например, 8 бит), а расстояния столкновений рассчитывались по нескольким неперекрывающимся сегментам набора данных.

Используя оценку столкновений, эта часть моей диссертации предлагает практичный и статистически обоснованный способ оценки непредсказуемости квантово-генерируемых двоичных данных, даже если данные не

соответствуют строгим критериям IID. Результаты дают убедительные доказательства прочности и надежности использованного мной источника энтропии.

ЗАКЛЮЧЕНИЕ

Тестирование физических генераторов случайных чисел – важное и актуальное направление в криптографии. Тестирование физических генераторов случайных чисел значительно отличается от тестирования псевдослучайных генераторов. Для физических генераторов большое значение имеет тестирование источника случайности на энтропию. Но эти набора тестов отличаются для IID и не IID последовательностей, снятых с генератора.

В работе реализован набор тестов на языке программирования Python на отнесение последовательностей случайных чисел к IID или не IID типу. Расчеты продемонстрировали. Что исследуемые последовательности не IID типа, поэтому для оценки энтропии источника случайности нужно применять соответствующие тесты.

Для тестирования источника случайности для исследуемых последовательностей программно реализован не IID тест оценки столкновений. Полученное расчетное значение мин-энтропии составило 6.4. Для максимального значения мин-энтропии теоретическое значение равно 8, для 8-битной системы. Таким образом расчетное значение близко к теоретическому пределу и источник энтропии исследуемого генератора случайных чисел можно считать достаточно надежным.

Полученные результаты могут быть использованы при разработке новых видов физических генераторов случайных чисел, а набор разработанных тестов применим для тестирования физических генераторов случайных чисел.

СПИСОК ОПУБЛИКОВАННЫХ РАБОТ

1. Наунг, Т Тестирование физического квантового генераторов случайных чисел / Тэт Наунг // 61-ая научная конференция аспирантов, магистрантов и студентов учреждения образования «Белорусский государственный университет информатики и радиоэлектроники». Секции «Информационные радиотехнологии», 21 – 25 апр. 2025 г. / Белорусский

государственный университет информатики и радиоэлектроники. – Минск, 2025.

2. Наунг, Т Тестирование физического квантового генераторов случайных чисел / Тэт Наунг // Материалы XXI Научно-технической конференции «Новые информационные технологии в телекоммуникациях и почтовой связи», (Республика Беларусь, Минск, 13-14 мая 2025 г.) /редкол.: А.О.Зеневич [и др.]. – Минск, Белорусская государственная академия связи, 2025. – С.85.