Введение

Большинство явлений, процессов, объектов, с которыми человек встречается в жизни, имеют случайную природу. Для их адекватного описания, изучения и моделирования недостаточно применять детерминированные методы (полностью определенные некоторым алгоритмом), поэтому последовательности случайных чисел и производящие их устройства и алгоритмы (генераторы случайных последовательностей) находят широкое применение в науке, технике, связи, информационных технологиях. Особую роль случайные последовательности играют в такой области, как обеспечение информационной безопасности. Одним из наиболее эффективных и перспективных подходов к решению проблемы защиты информации является применение криптографических методов, в которых генераторы случайных последовательностей часто являются ключевыми компонентами, во многом определяя их надежность. В предлагаемом пособии изложены основные сведения о случайных последовательностях, их использовании, получении и тестировании.

Пособие содержит 5 разделов, приложение и список рекомендуемой литературы. Каждый раздел сопровождается вопросами для контроля усвоения изученного материала. Первый раздел содержит сведения о сферах применения случайных чисел и способах их получения. Второй и третий разделы посвящены соответственно генераторам истинно случайных и псевдослучайных последовательностей. Рассматриваются общие принципы работы генераторов, приводятся их классификация, основные методы генерации, характеристики, а также примеры реализаций. Для генераторов истинно случайных последовательностей обсуждается процесс постобработки. Четвертый раздел включает определения криптостойких генераторов псевдослучайных последовательностей, требования к таким генераторам, их основные типы. В пятом разделе показаны основные подходы к проверке качества генераторов случайных последовательностей. Кратко описаны наиболее распространенные наборы тестов. Подробно рассмотрен известный пакет статистических тестов NIST STS. Приведены сведения о некоторых новых средствах тестирования. В приложении приводится необходимый материал из теории вероятностей, математической статистики и теории конечных полей.

Пособие предназначено для обучающихся по направлению «Информационная безопасность» (10.03.01 и 10.04.01), изучающих дисциплины, связанные с информационной безопасностью и, в частности, с применением криптографических методов защиты информации.

Освоение изложенного в пособии материала требует знания основ теории вероятностей, теории конечных полей, а также базового представления об 5алгоритмах шифрования. Разделы 1 –3 предназначены для получения начальных сведений о случайных последовательностях и методах их генерации. Они адресованы в первую очередь обучающимся бакалавриата. Разделы 4 и 5 болеецелесообразно использовать при обучении магистрантов, специализирующихся в области криптографии.

**Раздел 1. Случайные последовательности. Области применения и способы генерации**

* 1. **Случайные последовательности и их применение**

Большинство из окружающих нас объектов, явлений и происходящих процессов имеют случайную природу. Для адекватного описания, изучения и моделирования часто оказывается недостаточно детерминированных подходов, поэтому закономерно привлечение стохастических (т.е. имеющих случайный характер) методов решения разнообразных задач. В связи с этим случайные числа, последовательности таких чисел и производящие их генераторы находят все более широкое применение в науке, технике, связи, различных информационных технологиях, а также во многих аспектах повседневной жизни [1-13].

Исторически случайные числа начали использоваться для проведения выборочных наблюдений вместо непрерывных. Случайные числа применяются при решении сложных вычислительных задач и реализации вычислительных методов (например, метод Монте-Карло). Развитие ЭВМ, с одной стороны, расширило круг задач, использующих случайные числа, а с другой – предъявило высокие требования к качеству их генерации. Со временем случайные числа стали играть важнейшую роль в информатике, распределенных вычислениях, криптографии и других областях.

Последовательность чисел называется случайной, если воспроизвести ее, зная алгоритм и все исходные данные, не представляется возможным (дважды запустив генератор в тех же условиях, мы получим разные последовательности). Но компьютерные системы детерминированы, т.е. для них характерен строго определенный набор состояний (количество таких состояний может быть весьма большим, но конечным). Это приводит к тому, что генерируемые ими последовательности будут периодичны и воспроизводимы – такие последовательности называются псевдослучайными. Как известно, все периодическое является в той или иной степени предсказуемым, т.е. неслучайным. Получение истинно случайных последовательностей достаточно трудоемко. К тому же для их создания подходит далеко не каждый физический или информационный процесс.

Случайные последовательности находят применение в самых разных сферах человеческой деятельности. Ниже приведен перечень самых известных направлений, в которых случайные последовательности применяются наиболее интенсивно.

1. **Криптография.** Криптографические методы являются базовыми в обеспечении информационной безопасности. В криптографии случайные последовательности играют определяющую роль. Они используются, в частности, для получения ключевой последовательности используемого алгоритма шифрования, для генерации гаммы поточных шифров, а также для выработки векторов инициализации (блочных шифров).
2. **Другие направления защиты информации.** Случайные последовательности незаменимы при формировании паролей и пользовательских ключей (хороший пароль представляет собой короткую последовательность случайных символов). Кроме этого, они могут использоваться для внесения неопределенности в результаты работы различных алгоритмов защиты информации, а также в длительность выполнения шагов алгоритмов для защиты от утечек по побочным каналам. Они также необходимы при формировании случайных запросов при аутентификации и решении многих других задач.
3. **Тестирование алгоритмов.** Важной задачей является проверка правильности работы программ. Тестирование – достаточно долгий и трудоемкий процесс. Для его осуществления требуется большой объем входных данных. Использование генераторов случайных чисел повышает эффективность тестирования и позволяет экономить время.
4. **Сетевые протоколы.** Случайные последовательности могут использоваться, например, в качестве сессионных ключей, а также для выработки случайных параметров протокола, что обеспечивает уникальность его различных реализаций.
5. **Математическое и имитационное моделирование.** При моделировании сложных физических, технологических и социально-экономических систем и процессов обойтись без применения источников случайности не представляется возможным.
6. **Математическая статистика.** Математическая статистика изучает приближенные методы сбора и анализа данных по результатам эксперимента для выявления существующих закономерностей, т.е. нахождение законов распределения случайных величин и их числовых характеристик. Необходимой составляющей выборочных методов является формирование представительных выборок из генеральной совокупности с использованием случайных чисел.

Также случайные последовательности часто используются в теории чисел; статистической физике; прогнозировании; методах вычислений (в том числе методе Монте-Карло); теории управления; информационных технологиях для банковских, платежных, торговых систем; помехоустойчивом кодировании; автономном и встроенном диагностировании компонентов компьютерных систем; модуляции радиосигналов; в контроле хода выполнения программ с использованием сторожевых процессоров; индустрии игр и лотереях.

Одной из сфер деятельности, в которых случайные последовательности играют важную роль, является защита информации.

Случайные последовательности, применяемые в различных аспектах защиты информации, используются для решения следующих задач:

* генерация гаммирующих последовательностей при поточном шифровании информации по схеме, наиболее близкой к схеме абсолютно стойкого шифра;
* формирование векторов инициализации для блочных шифров, работающих в режиме обратной связи;
* получение начальных значений для программ генерации некоторых параметров в асимметричных криптосистемах;
* формирование пользовательских ключей и паролей;
* формирование ключевой информации, на секретности и качестве которой основывается стойкость большинства криптоалгоритмов;
* формирование случайных запросов при реализации большого числа криптографических протоколов, например, протоколов выработки общего секретного ключа, разделения секрета, привязки к биту, аутентификации, электронной подписи и др.;
* внесение неопределенности в работу средств защиты, например, при реализации концепции вероятностного шифрования, при котором одному и тому же исходному тексту при одном и том же ключе соответствует огромное множество шифротекстов;
* выполнение статистического тестирования;
* формирование затемняющих множителей при слепом шифровании (протокол слепой подписи).

Особое место занимает использование случайных чисел в криптографии – одном из самых мощных и эффективных методов защиты информации. В криптографии ключевую роль играют последовательности битов – двоичные последовательности, состоящих из случайно разделенных значений «0» и «1». Именно такие последовательности и будут называться в этом пособии «случайными».

Для ряда криптографических преобразований используют случайные первичные состояния либо целые последовательности. Следовательно, стойкость криптоалгоритма, использующего такие состояния или последовательности, напрямую зависит от алгоритма генерации случайных чисел и последовательностей, точнее от степени случайности выходных последовательностей.

Широта и важность областей применения случайных последовательностей, их определяющая роль в обеспечении высокого уровня защиты информации обуславливают актуальность их изучения.

Процесс генерации случайных чисел является основной частью многих криптографических операций. Например, криптографические ключи должны выбираться настолько случайно, насколько это возможно, чтобы на практике нельзя было воспроизвести их значения. Криптографически стойкие генераторы случайных чисел должны выдавать данные, которые невозможно предугадать с вероятностью выше 0,5; это означает, что любой метод предсказания очередного выходного бита не должен действовать эффективнее, чем просто случайное угадывание.

**1.2 Pseudo-tasodifiy ketma-ketliklarni yaratish usullari**

В настоящее время существует большое количество способов генерации последовательностей, обладающих той или иной степенью случайности [1-4, 7-9, 14-15]. Однако на практике большинство из таких генераторов производят последовательности, свойства которых не удовлетворяют требованиям случайности. Один из самых распространенных примеров этого – генераторы псевдослучайных чисел, встроенные в стандартные библиотеки многих языков программирования (такие, как, например, функция стандартной библиотеки языка C rand()). Часто в числах, сгенерированных с помощью подобных функций, прослеживаются явные закономерности. Например, полученные числа в одном и том же сеансе с течением времени монотонно возрастают, что прямо противоречит требованиям, предъявляемым к свойствам случайных (и псевдослучайных) последовательностей. Для широко известных и распространенных линейных конгруэнтных генераторов по четырем известным сгенерированным числам также можно предсказать дальнейшие значения.

Большинство криптографических приложений используют генераторы случайных чисел для создания ключей, с помощью которых шифруется и расшифровывается нужная информация. Однако часто именно применяемые в них генераторы являются самым слабым местом в системах шифрования. Дело в том, что программные генераторы полностью детерминированы. Обычно они используют различные сложные функции для вычисления псевдослучайных чисел. Соответственно, последовательности, полученные в результате работы таких генераторов, являются в той или иной степени предсказуемыми и воспроизводимыми и не подходят, например, для использования в криптографических приложениях. Необходимо отметить, что в некоторых случаях возможность воспроизвести случайную последовательность является полезной (например, при тестировании алгоритмов разработчиком). Тем не менее, последовательность не должна обладать свойствами, которые позволили бы злонамеренному криптоаналитику восстановить ее в процессе анализа работы защищенного приложения или протокола.

Существует табличный способ генерации случайных последовательностей. Он заключается в том, что случайные числа оформлены в виде таблицы, бумажной или электронной, которая хранится в оперативной памяти или на внешнем носителе. Один из вариантов таблицы случайных чисел и способа их выбора описан в ГОСТ Р ИСО 24153-2012 (Статистические методы. Процедуры рандомизации и отбора случайной выборки). Достоинство этого способа состоит в том, что с его помощью можно воспроизводить неоднократно одну и ту же последовательность псевдослучайных чисел. Однако серьезным недостатком, фактически не допускающим применения таких генераторов при решении практических задач, является то, что запас доступных чисел ограничен. Также при таком подходе возможно неэффективное использование вычислительных ресурсов компьютера (например, из-за необходимости хранить таблицу или ее части в оперативной памяти или обращаться к внешней памяти). В настоящее время такой способ генерации используется достаточно редко.

Среди генераторов псевдослучайных последовательностей (ГПСЧ), получивших широкое распространение и применимых при решении задач с серьезными требованиями к качеству сгенерированной последовательности, различают аппаратные, программные и программно-аппаратные (смешанные).

Аппаратный генератор случайных чисел – это устройство, которое генерирует последовательность случайных чисел на основе измеряемых, хаотически изменяющихся параметров протекающего физического процесса. При аппаратном способе генерации случайные числа являются прямым или побочным продуктом измерений некоторой физической величины, служащей надежным источником энтропии. Обычно это процессы, протекающие в неживой природе. Теоретически такие процессы абсолютно непредсказуемы, однако на практике полученные таким образом случайные числа приходится подвергать проверке с помощью специальных статистических тестов. Несмотря на лучшие статистические свойства и, соответственно, более высокую степень случайности, аппаратным генераторам присущи следующие недостатки:

-потенциально высокие временные и материальные затраты на конструирование, установку и настройку по сравнению с программными ГПСЧ;

-более низкая скорость генерации случайных чисел, чем при программной реализации ГПСЧ [14, 15];

-невозможность воспроизведения ранее сгенерированной последовательности чисел (что в некоторых случаях является нежелательным).

Программные (алгоритмические) генераторы (генераторы псевдослучайных последовательностей) основаны на детерминированных алгоритмах. У полученных таким образом последовательностей всегда существует период (пусть иногда и очень большой), а также наблюдаются и другие отклонения от случайности. Любой ГПСЧ с ограниченными ресурсами рано или поздно зацикливается – начинает повторять одну и ту же последовательность чисел. Период ГПСЧ зависит от типа генератора и его параметров [1, 4, 7, 9]. Если порождаемая последовательность ГПСЧ имеет слишком короткий период, то такой ГПСЧ становится непригодным для многих практических приложений.

Большинство простых арифметических генераторов хотя и обладают большой скоростью, но страдают от многих серьезных недостатков:

-слишком короткий период;

-последовательные значения не являются независимыми;

-некоторые биты «менее случайны», чем другие;

-неравномерное распределение;

-обратимость.

Фактически, результат работы таких генераторов не является случайной последовательностью. Тем не менее, к последовательностям, производимым программными генераторами, предъявляются определенные требования, поскольку они должны в какой-то степени имитировать случайные последовательности. В частности, период таких последовательностей должен быть достаточно большим, чтобы при генерации последовательности требуемой длины не возникало повторений. В отличие от аппаратных генераторов, программные генераторы способны воспроизвести ранее сгенерированную последовательность, что в некоторых случаях является бесспорным преимуществом.

**Программно-аппаратные генераторы.** Такой генератор может формировать поток случайных шумов, которые затем преобразуются в числа. Также возможен вариант, когда «зерно» (т.е. некие входные данные алгоритма шифрования) генерируется с помощью аппаратного генератора (поскольку ее размер достаточно небольшой, и, соответственно, ее получение не требует больших затрат времени и ресурсов), а итоговая последовательность – с помощью программного.

К программно-аппаратным генераторам случайных чисел могут относиться, например, устройства компьютера. В частности, источником случайной последовательности могут быть шумы устройств компьютера (например, процессора), системное время, временные интервалы между нажатиями клавиш, движения мыши и так далее. Как правило, последовательности, получившиеся в результате таких процессов, нуждаются в постобработке. К тому же скорость их получения является достаточно низкой (особенно при генерации последовательностей достаточно больших объемов). К таким генераторам можно отнести, в частности, псевдоустройства /dev/random и /dev/urandom ОС Linux.

В последующих разделах описанные выше подходы к генерации случайных последовательностей, а также процесс статистического тестирования генераторов случайных последовательностей будут рассмотрены подробнее.

Последовательность называется истинно случайной (ИСП), если ее нельзя воспроизвести. Это означает, что если запустить генератор истинно случайных последовательностей дважды при одном и том же входе, то на его выходе получатся разные случайные последовательности. Основная трудность состоит в том, чтобы суметь отличить случайную последовательность от неслучайной.

Однако на практике далеко не всегда можно непосредственно использовать выходные данные источников истинно случайных чисел. Поэтому обычно приходится использовать так называемые псевдослучайные последовательности. Псевдослучайная последовательность (ПСП) – это последовательность, состоящая из псевдослучайных двоичных чисел, получаемых с помощью заданного детерминированного алгоритма, но применяемых в качестве случайных. При этом обычно алгоритмы получения ПСП используют специальное случайное начальное значение, или «зерно» (seed). Для того чтобы ПСП могли использоваться в качестве случайных последовательностей, они должны по статистическим свойствам быть близки к ИСП.

Мы знаем, что на микроуровне случайность существует (квантовая механика), но неизвестно, сохраняется ли эта случайность при переходе на макроуровень. Дополнительное свойство случайной последовательности заключается в том, что случайная последовательность не может быть сжата.

Требования к качественному генератору случайных чисел:

1. Непредсказуемость результатов работы: при неизвестном ключе/начальном состоянии генератора на основе известной конечной части ПСП невозможно определить как ее последующий элемент (прямая непредсказуемость, или непредсказуемость вправо), так и предыдущий (обратная непредсказуемость, непредсказуемость влево);
2. Неотличимость статистических свойств генерируемых ПСП от аналогичных свойств истинно случайной последовательности;
3. Большой период последовательности;
4. Возможность эффективной аппаратной и программной реализации.

На практике добиться выполнения всех этих условий, как правило, не представляется возможным. Более того, часто эти условия являются взаимоисключающими. Поэтому приходится искать баланс между ними и в первую очередь стремиться к выполнению того, что является наиболее важным в контексте решаемой задачи.

Часто наилучшие результаты получаются в случае комбинирования разных способов генерации случайных последовательностей. Например, начальная информация может быть получена при помощи аппаратного генератора, а сама итоговая последовательность – с помощью программного, получившего на вход начальные данные с аппаратного генератора.

**Раздел 2. Генераторы истинно случайных последовательностей**

Истинно случайные последовательности (ИСП) – это представленные в виде последовательности случайные числа. Случайные числа являются реализацией некоторой случайной величины. Случайная величина – это функция над пространством элементарных событий, принимающая вещественные значения с некоторыми вероятностями. Таким образом, истинно случайные последовательности – это последовательности статистически независимых друг от друга величин, принимающих в результате опыта одно из множества заранее непредсказуемых значений. Связь между значениями случайной величины и соответствующими вероятностями описывается законом распределения, который задается с помощью функции распределения или связанной с ней функцией плотности вероятности. Истинно случайные последовательности должны обладать равномерным распределением.

Генераторы истинно случайных последовательностей – это устройства, которые для получения случайных последовательностей используют объективно существующую случайность физических процессов, происходящих как на макроуровне, так и на микроуровне. Как правило, такие генераторы вляются аппаратными, либо программно-аппаратными.

Потребность в ГСП возникает в тех случаях, когда требуется очень высокая степень случайности, недостижимая при использовании ГПСП (квантовая криптография, начальное состояние ключа для блочных шифров).

**2.1 Принципы работы**

В аппаратных генераторах случайных последовательностей для генерации случайных чисел используются такие источники энтропии, как:

-тепловой и электрический шум;

-квантовые процессы;

-радиоактивный распад;

-космическое излучение;

-различные механические, оптические и фотоэлектрические явления.

Конкретными примерами источников энтропии, подходящих для генерации истинно случайных последовательностей, могут быть:

-временные интервалы между выбросами частиц при радиоактивном распаде;

-тепловой шум полупроводникового диода или резистора;

-состояния спутанности фотонов;

-квантовый шум лазеров;

-дробовой шум;

-нестабильности частоты осцилляторов.

Программно-аппаратные генераторы истинно случайных последовательностей основаны на случайностях, присущих работе компьютерного оборудования, таких, как показания системных часов, уровень загрузки процессора, задержки между прибытиями сетевых пакетов, интервалы времени между срабатываниями мыши или клавиатуры, содержимое буферов ввода/вывода, шумы процессоров или других устройств.

Реальное применение ГСП может быть затруднено такими их особенностями, как:

-низкая скорость работы;

-сложность повторного воспроизведения, дублирования и взаимодействия с процессором;

-отклонения и корреляции в получаемых последовательностях, связанные с систематическими ошибками в ходе измерений или наличием волновых или других периодических (неслучайных) составляющих, выявляемые при статистическом тестировании.

Результат работы ГСП может потребовать дополнительной обработки (так называемая постобработка).

Полученная случайная последовательность может использоваться непосредственно или быть входной для генератора псевдослучайных последовательностей.

Особое место среди источников случайных данных занимают процессы, описываемые в рамках квантовой физики. Их вероятностная природа делает теоретически возможным получение истинно случайных последовательностей. Существуют и разрабатываются квантовые генераторы истинно случайных чисел, основанные на явлениях радиоактивного распада, запутанных квантовых состояниях, лазерном квантовом шуме, квантовых флуктуациях в вакууме, процессах эмиссии и детектирования фотонов.

Физические явления различной природы принято называть «случайным шумом» (белым шумом), если они представляют собой беспорядочные колебания.

Известный пример случайного шума – тепловой шум, или шум Джонсона. Это колебания напряжения, измеренного для любого материала, обладающего электрическим сопротивлением и находящегося при температуре выше абсолютного нуля. Причиной таких колебаний является тепловое движение носителей электрического заряда, имеющее случайный характер. Следует, однако, заметить, что указанное напряжение в реальности не является полностью случайным, поскольку существуют определенные корреляции носителей в проводниках, вызывающие корреляции в движениях электрических зарядов.

Туннельный эффект Зенера, наблюдаемый в полупроводниковых стабилитронах (специальных диодах Зенера, которые способны работать в условиях обратного смещения в зоне пробоя), вызывает случайные скачки напряжения при переходах носителей через квантовый барьер (так называемый «розовый шум»). При этом эффект Зенера не изолирован полностью в физических устройствах от других эффектов. К тому же для названных процессов в сопротивлениях и стабилитронах характерен эффект памяти: мгновенное напряжение на устройстве зависит от напряжения в недавнем прошлом. Это приводит к корреляции полученных таким образом чисел и не дает права назвать такую последовательность истинно случайной.

Можно назвать и другие популярные источники шума, например, пробой база – эмиттер в биполярных транзисторах, фазовый шум лазера, хаотический шум и др. Однако общая проблема для всех этих источников шума состоит в том, что порождаемую ими случайность невозможно абсолютно точно проконтролировать при изготовлении соответствующего устройства или измерить. Значения напряжения (например, для шума Джонсона) могут быть очень малы, что требует существенного усиления перед преобразованием в цифровую форму. Это добавляет дополнительные отклонения из-за ограниченной полосы пропускания усилителя и нелинейности коэффициента усиления. При быстром переключении двоичной логики, которое используется в схеме генератора случайных последовательностей, возникают сильные электромагнитные помехи, из-за которых находящиеся вблизи генераторы (особенно расположенные на одном чипе), обычно взаимно синхронизируются, что приводит к резкому падению общей энтропии. Существует также опасность криптографических атак на шумовые генераторы истинно случайных последовательностей путем воздействия на них высокочувствительных усилителей.

Основная идея построения генератора истинно случайных последовательностей, базирующихся на источниках шума, состоит в следующем. Случайное аналоговое напряжение, поступающее от источника шума, периодически дискретизируется, усиливается и подается на компаратор для сравнивается с заранее выбранным порогом. При превышении этого порога генерируется значение «1», в обратном случае генерируется «0». Порог может быть установлен таким образом, чтобы вероятности появления «1» и «0» будут примерно равны. Процедура настройки порога оказывается сложной, отнимающей много времени и может вызвать заметное искажение показаний генератора шума.

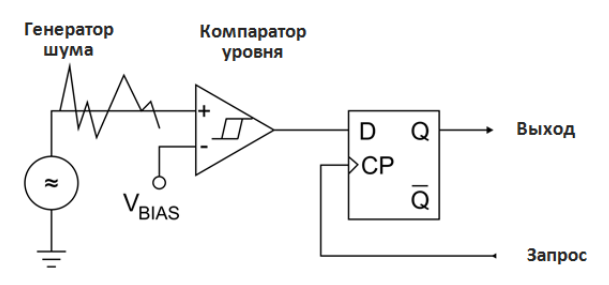


Рисунок 2.1 – Схема генератора истинно случайных последовательностей, основанного на источниках шума.

Для приведенной базовой схемы предлагаются различные модификации, направленные на улучшение степени случайности выходных данных генератора, и в особенности – на уменьшение смещения, которое содержится в необработанном потоке шума. Примером одного из таких решений служит генератор Баджини-Буччи.

В целом надежность любого генератора случайных последовательностей, основанных на шуме, зависит от следующих факторов:

-степень случайности используемого источника шума;

-влияние процедур выборки и оцифровки сигналов;

-необходимость использования детерминированной постобработки.

В итоге по указанным причинам доказательство надежности ГСП на основе шума становится практически невозможным.