УДК 681.3.06

Ю.И. ГОРБЕНКО, канд. техн. наук, Р.И. МОРДВИНОВ

СУЩНОСТЬ И АНАЛИЗ КРИПТОГРАФИЧЕСКИХ ТРЕБОВАНИЙ СТАНДАРТА NIST SP 800-90B

К источникам энтропии и методам генерации ключей в криптографии уделяется особое внимание, требования к ним изложены в ряде стандартов [1 – 4]. В то же время указанные стандарты относительно криптографии носят приложенческий характер. Поэтому важной задачей является, с одной стороны, формализация их требований, а с другой – рассмотрение ряда новых требований как к источникам энтропии, так и непосредственно к самим средствам генерации ключей.

Решение указанных задач представлено в рабочей версии стандарте NIST DRAFT Special Publication 800-90B [5].

Цель настоящей статьи – рассмотрение и анализ основных требований к источникам энтропии и непосредственно генераторам случайных последовательностей, стратегии их применения, возможности нестандартного поведения источников энтропии и требования валидации, а также общей стратегии и частных вопросов тестирования источников энтропии и генераторов ключей. Дополнительной целью является анализ и обоснование возможностей гармонизации этого стандарта в Украине.

Требования к источнику энтропии

Ниже, на основе [5], приведены основные требования, выдвигаемые к источнику энтропии, как в целом, так и к отдельным его компонентам. Данные требования должны помочь разработчикам в проектировании источников энтропии, способных удовлетворить требованиям, выдвигаемым в документации относительно их валидации.

Основные функциональные требования к источнику энтропии следующие [5]:

1. Разработчик должен документировать конструкцию источника энтропии в целом, включая взаимодействие компонентов. Данная документация должна обосновывать, почему источник энтропии способен вырабатывать биты с заданной энтропией.

2. Источник энтропии должен иметь четкую и понятную границу безопасности, которая должна быть такой же, либо находится в пределах криптографических границ приведенных в FIPS 140. Данная граница безопасности должна быть задокументирована и включать в себя:

- описание содержания границы безопасности, причем граница безопасности может выходить за пределы энтропии самого источника, т.е. источник энтропии может содержаться в больших границах неопределенности, чем детерминированный генератор случайных бит (ДГСБ), а граница безопасности может быть логической, т.е. не физической;
- описание того, как граница безопасности гарантирует, что нарушитель не сможет уменьшить энтропию, если он находится за пределами границы, либо путем наблюдения либо манипуляции;
- любые предложения относительно поддержки функций, от которых зависит граница безопасности.

3. Разработчик должен задокументировать диапазон рабочих условий, при которых источник энтропии может генерировать приемлемые случайные данные, причем должны быть четко описаны меры, которые были приняты при проектировании системы, с целью работы источника энтропии в нормальных условиях.

4. Предусмотрена возможность проверки источника энтропии на соответствие FIPS 140, включая соответствующие интерфейсы для получения данных.

5. Предоставлена документация, которая описывает поведение источника шума и объясняет, почему считается, что значение энтропии не колеблется при нормальных операциях.

6. После выявления в процессе тестирования неисправности источник энтропии должен прекратить вывод данных и уведомить приложение о состоянии ошибки.

7. Для улучшения устойчивости по отношению к деградации или некорректному поведению, источник энтропии может содержать несколько источником шума. Если данная функция реализована, то требования применяются к каждому из источников энтропии.

К полному(законченному) источнику энтропии выдвигаются еще и такие требования:

1. Выходная строка бит, полученная из источника энтропии, должна обеспечивать, по крайней мере, $(1-\varepsilon)n$ бит энтропии, где n – длина каждой выходной строки и $0 \le \varepsilon \le 2^{-64}$;

2. Выход источника энтропии должен формироваться с использованием криптографической функции, по крайней мере, двойной размер блока базового криптографического примитива должен быть использован в качестве входных данных в функцию состояния

Анализ требований к источнику шума

К источнику шума выдвигаются следующие функциональные требования [5]:

1. Источник шума должен формировать случайные выходные данные, т.е. на выходе не должно быть последовательностей с известными правилами;

2. Разработчик должен задокументировать функционирование источника шума. Документация должна включать в себя описание того, как источник шума работает, и обоснование, почему обеспечивается приемлемая выходная энтропия, ссылаясь на соответствующие исследования и литературу.

3. Для обеспечения правильной работоспособности источник шума должен иметь средства тестирования правильной работоспособности. В частности, должна быть обеспечена возможность сбора от источника шума данных для тестирования его работоспособности и определения значения энтропии, а также соответствие теста работоспособности. Также выходные данные, полученные от источника шума, должны обеспечивать свойства не оборотности, однородности и не влиять на последующие выходные;

4. Должна обеспечиваться возможность обнаружения отказа или серьезной деградации источника шума. Методы, используемые в таких условиях, должны быть задокументированы.

5. В документации на источник шума должны быть описаны условия, если таковые имеются, при которых известны неисправности, включая описание различных сред, в которых источник шума может работать неправильно. В этом случае непрерывное тестирование или другие механизмы источника энтропии должны обнаруживать неисправности;

6. Источник шума должен быть защищен, насколько это возможно, от ознакомления или влияния нарушителя. Методы, используемые для этого, должны быть задокументированы, в том числе описано использование границ безопасности для защиты источника шума от наблюдения или влияния со стороны нарушителя.

Особенности требований к валидации

Под валидацией понимается подтверждение на основе представления объективных свидетельств того, что требования, предназначенные для конкретного использования или применения, точно и в полном объёме предопределены, а цель достигнута. Проверка источника энтропии, т.е. его валидация, необходима для того, чтобы получить уверенность в том, что все соответствующие требования рекомендации выполнены. Проверка состоит из тестирования аккредитованной лаборатории источника шума в соответствии с требованиями SP 800-90B.

Проверка энтропии источника является сложной проблемной задачей. Это связано с тем, что ни одна другая часть генератора случайных бит (ГСБ) не является настолько зависимой от технологических различий и различий окружающей среды. В то же время, правильная работа источника энтропии является важной для безопасности ГСБ, и как следствие безопасности криптографического приложения.

В этом разделе представлены требования высокого уровня как для разработчиков, так и тестировщиков, которые должны быть выполнены для проверки методов отображения больших выборок в более мелкие битовые строки в ситуациях, когда сложно собрать достаточно данных для проверки с учетом размера выборки. Также требования, представленные ниже, предназначены для обеспечения проверки источника энтропии и обоснования, почему источник энтропии при их выполнении можно использовать.

Средства создания энтропия источника должны состоять, как минимум из трех компонентов: источника шума, средства тестирования работоспособности и дополнительных компонентов анализа состояния. Источник энтропии должен обеспечивать энтропию, которая предусмотрена, причем ее значение должно подтверждаться во время тестирования.

Ниже приведены общие требования к валидации тестирования:

1. При сборе данных:

- сбор данных должен осуществляться одним из двух способов: разработчиком со свидетелем из тестовой лаборатории или путем лабораторных испытаний. Источник энтропии должен содержать интерфейс, который позволяет получать доступ к битам из источника шума и выходным состояниям из компонентов состояний. Этот интерфейс должен позволять получать выходные данные источника шума, то есть, эти результаты не должны использоваться для чего-либо еще хотя бы раз. Интерфейс должен быть доступен во время проверки тестирования, но может быть отключен, если тестирование не проводится;

- данные должны быть получены из источника шума и компонентов состояния, если доступ к таковым имеется, при нормальных условиях эксплуатации;

- данные, полученные из источника энтропии, должны быть не обработанными, а лишь оцифрованными. Может указываться необходимый формат данных для подачи последовательности для тестирования;

- из источника шума должна быть получена необработанная последовательность длины не меньше, чем 1000000 последовательных значений;

- если используется неутвержденный компонент состояния, то длина последовательности для проверки должна быть не меньше 1000000 последовательных значений. Необходимо учитывать, что данные, полученные из источника шума для проверки, могут быть использованы в качестве входных данных для компонентов состояния, чтобы получить выходные данные;

- для выборки значений, состоящих из более чем одного бита, разработчик должен предоставить тестер с упорядоченным ранжированием бит выборки.

2. При тестировании с целью валидации:

- Должно непрерывно осуществляться тестирование работоспособности, причем необходимо проверять, что реализованные тесты обнаруживают отказ с помощью теста счетчика повторов и адаптивного теста долей.

- Тесты должны быть применимы для всех образцов, представленных для тестирования.

- Разработчик должен указать, производит ли источник шума IID данные, либо не IID данные. Это должно использоваться во время определения алгоритма реализации теста валидации. При этом требование полной энтропии будет интерпретироваться, как IID требование. Сущность их состоит в следующем:

- минимальная оценка энтропии, получаемая путем тестирования, должна иметь значение, при котором источник энтропии является валидным. Эта оценка энтропии будет использоваться как минимальная энтропия в выборках;

- источник будет признан таким, что имеет полную энтропию, только после прохождения тестов на IID данные;

- тестировщик будет проверять все данные, а так же рассматривать всю документацию и теоретические обоснования, которые представленные разработчиком.

3. При документировании с целью валидации:

- разработчик должен предоставить документацию, которая описывает работу источника энтропии, а именно: как он работает, как создается энтропия и как получить данные их источника энтропии для их тестирования;

- документация должна быть представлена таким образом, чтобы в лаборатории или поставщик могли выполнять (или повторять) процесс сборки, если это необходимо. Процесс сборки не требует глубоких знаний источника или вмешательств, которые могут изменить поведение источника энтропии;

- документация должна обеспечивать техническую аргументацию, почему источник шума может поддерживать определенный уровень энтропии. Это может быть описание в общих чертах о том, откуда появляется непредсказуемость, а также описание поведения источника шума, чтобы показать его нормальное поведение;

- документация должна содержать описание условий, при которых источник энтропии должен работать корректно (температура, напряжение, системы деятельности). Анализ поведения источника энтропии на границах этих условий должен быть документально оформлен, так же, как и вероятность отказов и повреждений;

- описание тестов работоспособности и основания для этих тестов должны быть включены в документацию. Разработчик должен представить исходный код для любых тестов как альтернативу, либо в дополнение к данным рекомендациям;

- разработчик должен представить описание пространства выходных данных источника шума, в том числе его размер, а также указать размер выборки источника шума, являющегося фиксированной величиной для данного источника шума;

- для источников энтропии, содержащих компоненты состояния, предусматривается их описание, включающее спецификации размера выходного блока компонента состояния.

Общая стратегия тестирования источника энтропии

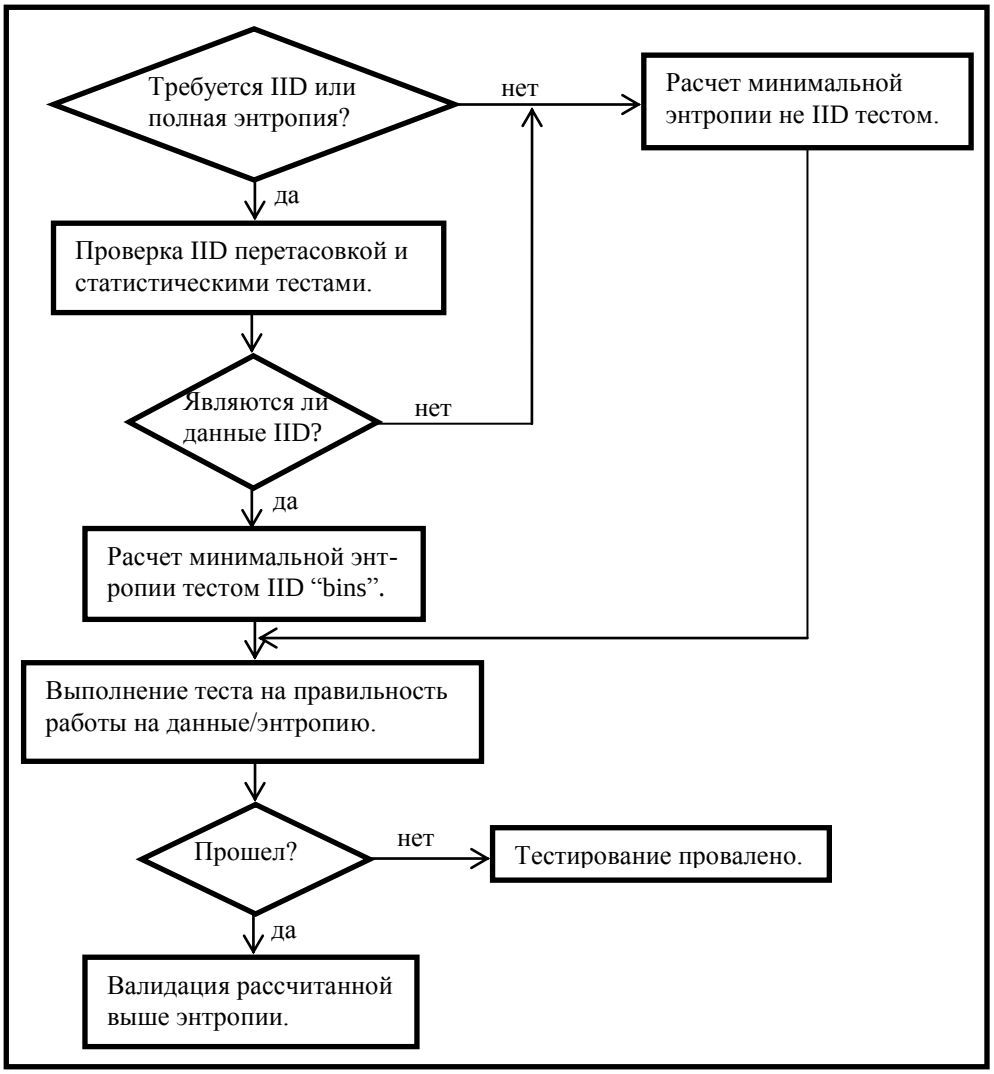
Основным требованием для любого источника энтропии является возможность правильной оценки минимальной энтропии источника. Оценка минимальной энтропии источника энтропии рассчитывается исходя из данных, представленных для тестирования. Как показано на рисунке, первым происходит выяснение, производятся ли IID данные, или нет. Оба результата приводят к оценке минимальной энтропии, но используют при этом для определения оценки энтропии разную информацию. Требования полной энтропии означают, что источник энтропии производит IID данные; любое требование о полной энтропии будет интерпретироваться, как требование IID. Их сущность заключается в следующем:

1. Если требуется, чтобы данные были IID, то полный набор проверок и статистических тестов выполняется для подтверждения того, что данные есть IID.

2. Если результаты тестирования доказали, что данные являются IID, то минимальная энтропия оценивается для источника шума с помощью тестов. Эта оценка должна использоваться для подтверждения минимальной энтропии источника шума, который производит IID данные.

3. Альтернативой, если нет требования IID, или если тесты не поддерживают такого требования, является набор из пяти тестов. Далее, данные пяти тестов приводят к пяти различным оценкам энтропии. При выборе минимальной энтропии выбирается худшая оценка, полученная по тестам, которая и будет в качестве оценки минимальной энтропии источника шума, который не производит IID данные.

4. После получения оценки энтропии данные подвергаются проверке. Эти тесты предназначены для обнаружения существенных сбоев или завышенных значений тестов энтропии. Неспособность пройти такую проверку означает, что источник энтропии не удовлетворяет требования. Если источник энтропии не использует компоненты состояния, но компоненты источника шума прошли тестирование, то считается, что источник энтропии имеет минимальную энтропию.



Стратегия тестирования энтропии

Анализ основных тестов

Одним из основных требований к источнику энтропии является генерирование случайных выходных даных. При этом, для обеспечения достаточной энтропии, подающейся на вход ДГСБ, количество энтропии, образующейся за счет источника шума, должно быть определенным. Далее рассматриваются основные тесты, которые используются при тестировании источников шума, в том числе, когда используются неутвержденные методы, компонент состояний. Эти тесты являются универсальными, но они предназначены в основном для выявления серьезных проблем, которые могут возникать. Тесты предназначены для применения в аккредитованных испытательных лабораториях, но они могут также использоваться и разработчиками источников энтропии.

Тест проверки на независимость и стабильность (перемешивание). Тест позволяет выявить отклонения источника шума или компонент состояния от распределения при независимом и стабильном поведении. Набор данных длины N делится на 10 неперекрывающихся подмножеств равной длины $\left\lfloor \frac{N}{10} \right\rfloor$. Каждое из этих подмножеств тестируется, чтобы определить, что выборки относятся к IID распределению.

В процессе тестирования рассчитываются следующие оценки:

- оценка сжатия – один балл за подмножество данных;

- оценка больших/малых серий – два балла за подмножество данных;
- оценка посещения – один балл за подмножество данных;
- оценка направленных серий – три балла за подмножество данных;
- оценка ковариации – один балл за подмножество данных;
- оценка коллизий – три балла за подмножество данных.

В последующем на каждом из 10 подмножество данных выполняются:

- оценки по оригинальному подмножеству данных, что дает вектор J оценок, где J является количеством баллов для данного теста на каждое подмножество данных;
- следующие шаги повторяются 1000 раз для подмножества данных, которые перемешиваются с помощью генератора псевдослучайных чисел;
– оценки рассчитываются по преобразованным данным, что дает вектор оценок J, где J является количеством баллов для данного теста на одно подмножество данных.
- оценки запоминаются и хранятся в списке;
- для каждого из полученных соответствующих баллов за каждое подмножество данных, оценке оригинальных подмножеств данных присваивается ранг. При этом, если значение оценки оригинального подмножества данных ниже, чем оценка всех преобразованных подмножеств данных, то ранг оригинального подмножества составит 1, и оценки всех других преобразований подмножеств будут иметь более высокий ранг. Если значение оценки оригинального подмножества выше, чем оценка всех преобразованных подмножеств данных, то он будет иметь ранг 1000, а все остальные оценки будут иметь ранг < 1000.

Возможно, что многие преобразованные подмножества данных будут иметь одинаковые оценки. Если оригинал подмножества имеет такую же оценку, как одно или несколько преобразованных подмножеств данных, то она берется ближе к средней. Далее, с учетом отсортированного списка $L[0...1001]$ из оценок преобразованных подмножеств данных и оценки оригинального подмножества S ранг определяется согласно следующему правилу (1):

$$Rank(S) = \begin{cases} \max(j) \textit{ такое, что } L[j] \le S & \textit{если } L[500] > S \\ 500 & \textit{если } L[500] = S \\ \min(j) \textit{ такое, что } L[j] \ge S & \textit{если } L[500] < S \end{cases} \quad (1)$$

Правило (1) гарантирует, что только исключительно высокие или низкие показатели отображаются как аномальные в тесте преобразований (перемешивания).

Далее ранги рассматриваются как p-значения в двухстороннем тесте. С использованием 10 рассматриваемых подмножеств данных формируется набор из $10 \times J$ p-значений для каждого из шести тестов. При этом P-значения объединяются в соответствии со следующими правилами:

- для интервала $50 \le rank \le 950$ вероятность 10 %;
- если 8 или более оригинальных подмножеств данных имеют значение $50 \le rank \le 950$, то источник не проходит испытание;
- если источник шума, подвергнувшись тестированию, не проходит испытание на независимость и стабильность, то он не считается IID. Далее тестирование источника шума необходимо испытывать по альтернативному пути;
- если неутвержденный компонент состояния проходит тест, источник энтропии не проходит валидацию.

К примеру, если S = 20, L[500] = 22, L[299] = 19, L[300] = 20, L[301] = 20 и L[302] = 22, тогда Rank(S) = 301. Если S = 20, L[500] = 18, L[599] = 19, L[600] = 20 и L[601] = 20, тогда Rank(S) = 600.

Тест по оценке сжатия. Универсальные алгоритмы сжатия эффективно применяются для удаления избыточности в строке символов, особенно при наличии повторяющихся в последовательностях символов. Оценка сжатия подмножества данных, оригинальных или

преобразованных, изменяет длину подмножества. Баллы за сжатие вычисляются следующим образом:

- выборка в подгруппе данных кодируется в виде строки символов, содержащих список значений, разделенных запятыми;
- символьная строка обрабатывается при помощи, например, [BZ2] алгоритма сжатия;
- определяется длина сжатой строки в байтах.

Оценка больших/малых серий. В процессе этой оценки вычисляется среднее значение подмножеств данных и все значения, не равные ему либо большие, либо ниже среднего. Длинные серии, которые выше либо ниже серии, должны появляться сравнительно редко, если выходные данные являются независимыми и стабильными. Кроме того, число серий выборок, которые находятся выше или ниже среднего, являются мерой, показывающей, является ли данная выборка зависимой от соседних и устойчиво ли распределение.

Этот тест применим для любого подмножества данных, выборки берутся в числовом или порядковом значении. Тем не менее, тест должен использоваться, даже если данные не являются числовым или порядковым номером.

Каждое подмножество данных (оригинальное или преобразованное), в дальнейшем используется для получения временного подмножества данных, состоящих только из значений -1 и +1. Значениям, которые меньше среднего, присваиваются -1 во временном подмножестве данных, а тем, что больше среднего, присваивается +1. Если значение равно среднему, то оно упускается из временного подмножества. Самая длинная серия и количество серий во временном подмножестве отмечаются, ей дается на 2 балла меньше для каждого из оригинального или преобразованного подмножества данных.

Баллы рассчитываются следующим образом:

1) подмножества данных используются для вычисления среднего значения, если данное подмножество является двоичным, то среднее будет 0.5;

2) временное подмножество строится следующим образом – для каждого из оригинальных и преобразованных подмножеств данных:

- если элемент больше, чем среднее, то добавляем +1 к временному подмножеству,
- если элемент меньше, чем среднее, то добавляем -1 к временному подмножеству,
- если элемент такой же, как и средний, то опускаем его во временном подмножестве;

3) определяется самая длинная серия -1 или +1 во временном подмножестве данных и длинна этой серии будет первой оценкой;

4) число серий из -1 и +1 во временном подмножестве используется как вторая оценка.

К примеру, пусть подмножество данных состоит из 7 значений {5, 15, 12, 1, 13, 9, 4}. Средним значением для этого подмножества будет 9. В таком случае временным подмножеством будет {-1, +1, +1, -1, +1, -1}. Получим серии (-1), (+1, +1), (-1), (+1), (-1). Таким образом, длина наибольшей серии составляет 2 (первая оценка), а количество серий составляет 5 (вторая оценка).

Оценка «посещений». Определяется, когда имеется последовательность высоких или низких значений оцено(кластеров). Это факт указывает на то, что распределение может не быть стабильным и независимым. Потому, как высокие, так и низкие баллы «посещений» представляют интерес. Высокие показатели представляют собой оценки, которые принимают значение из области высокого/низкого значения. Низкие указывают на некоторый процесс, который предотвращает необычно высокие/низкие значения от кластеризации вместе.

Оценки «посещения» имеют смысл, если среднее значение из выборки имеет смысл, например, если среднее может быть вычислено по набору значений данных. В большинстве случаев, в которых выборка значений является счетчиком или оцифрованным значением, среднее значение выборки данных может быть вычислено. Тем не менее, тестирование должно выполняться, если даже среднее не может быть вычислено.

Оценка «посещений» является мерой того, насколько текущая сумма значений отсчетов отклоняется от ожидаемого значения в каждой точке множества данных. Если подмножество данных, $s_0, s_1, s_2,...,$ а среднее из выборки значений является μ, то «посещ ение» в позиции i будет $s_0 + s_1 + ... + s_i - i \times \mu$. Возвращаемая оценка является абсолютным максимальным значением любого «посещения» в подмножестве данных.

В этом случае оценка вычисляется следующим образом:

- для $j = 1 \, до \left\lfloor \frac{N}{10} \right\rfloor$:

d_j _ абсолютное значение (сумма первых j выборок – $j \times \mu$);

- возвращаемая оценка является наибольшим d_j значением.

К примеру, если подмножество данных {2, 15, 4, 10, 9}, тогда $\mu = 8$. $d_1 = |2-8| = 6;$ $d_2 = |(2+15)-(2\times 8)| = 1;$ $d_3 = |(2+15+4)-(3\times 8)| = 3;$ $d_4 = |(2+15+4+10)-(4\times 8)| = 1;$ $d_5 = |(2+15+4+10+9)-(5\times 8)| = 0.$ Получим, что наибольшее значение $d_j = 6$.

Оценка направленных серий. Многие физические процессы лучше всего понимаются в терминах их производных. Если в первой производной знаки чередуются между выборками, то в последовательности будет присутствовать большое количество коротких направленных серий. Каждое подмножество данных (оригинал или преобразование) используется для создания временного подмножества данных, такого, что один элемент короче, чем оригинал подмножества данных. Временное подмножество содержит указания о том, что первое значение в каждой паре элементов меньше, равно или больше второго элемента в паре. Если первый элемент меньше, то добавляется +1 к временному подмножеству, если элементы равно, то добавляется 0 к временному подмножеству, если первый элемент больше, то добавляется -1 к временному подмножеству. Оценки рассчитываются по алгоритму, описанному ниже, где s_i – *i*-й элемент подмножества данных, а вес Хэмминга определяется как число единиц в последовательности.

Временное подмножество данных формируется из подмножества данных:

1) Если источник является недвоичным, то :

для $i = 0$ до (длина оригинального подмножества данных) – 2:

- если $s_i < s_{i+1}$, тогда $temp_i = 1$;

- если $s_i > s_{i+1}$, тогда $temp_i = -1$;

- в противном случае $temp_i = 0$.

2) Если источник последовательности является двоичным, то подмножество битов объединяется в байты, и выполняется:

- для $i = 0$ до (длина оригинального подмножества данных)/8 – 1:

$$W_i = вес\_Хэмминга(s_i,..., s_{i+7}) \ ;$$

- для $i = 0$ до (длина последовательности W) – 2:

-если $W_i < W_{i+1}$, $тогда \ \ temp_i = 1$;

- если $W_i > W_{i+1}$, $тогда \ \ temp_i = -1$;

- в противном случае $temp_i = 0$.

Затем выполняется расчет баллов по временному подмножеству данных.

Пример с недвоичным подмножеством данных {2, 2, 2, 5, 7, 7, 9, 3, 1, 4, 4}. Получим временное подмножество {0, 0, +1, +1, 0, +1, -1, -1, +1, 0}. Таким образом получаем три серии: (+1, +1, 0, +1), (-1, -1), (+1, 0), следовательно, первая оценка равна 3. Длина самой длинной серии равна 4 (вторая оценка). С учетом того, что во временном подмножестве

четыре “+1” и две “-1”, то третья оценка равна 4.

Пример с двоичным подмножеством данных. Биты объединяются в байты и последовательность имеет вид {A3, 57, 3F, 42, BD}. Вес Хэмминга соответственно {4, 5, 6, 2, 6}. Временным подмножеством будет {+1, +1, -1, +1}. Получим три серии (+1, +1), (-1), (+1), и первая оценка равна 3. Самая длинная серия равна 2 (вторая оценка). Мы имеем три значения “+1” и одно значение “-1”, значит третья оценка равна 3.

Оценка ковариации. Есть случаи, когда источник шума не связан с IID зависимостями между соседними выборками. Тест на независимость Хи – квадрат и оценка сжатия являются эффективными для обнаружения повторяющихся значений. Но он не будет выявлять связи между числовыми значениями последовательных выборок.

Поведение ковариационной оценки рассмотрим для двумерного нормального распределения. В этом случае любая линейная зависимость между последовательными парами значений повлияет на этот результат, и поэтому ковариация будет различной для исходного подмножества данных и для преобразованного подмножества данных, если такая связь существует. Ковариация двух переменных определяется как ожидаемое значение их результата, за вычетом результата их ожидаемого значения. Ковариация является общим статистическим показателем наличия линейной зависимости между двумя переменными. Ковариационный «балл» вычисляется для каждой пары подмножества данных S, но таким образом, что s_0 является парой для s_1, s_1 для s_2, и т.д.

Оценка вычисляется следующим образом:

- счетчик $count = 0$;
- вычисляется из $s_0, s_1, \ldots s_{N/10-1}$ μ – среднее значение;
- для $i = 0$ *до* $\left\lfloor \frac{N}{10} \right\rfloor$:

$count = count + (s_i - \mu)(s_{i-1} - \mu)$.

Баллы вычисляются как $\frac{count}{\left\lfloor \frac{N}{10} \right\rfloor - 1}$.

К примеру, если подмножество данных будет {15, 2, 6, 10, 12}, тогда $\mu = 9$. Для

$i = 1,\ count = (2-9)(15-9) = -42;$

$i = 2,\ count = -42 + (6-9)(2-9) = -21;$

$i = 3,\ count = -21 + (10-9)(6-9) = -24;$

$i = 4,\ count = -24 + (12-9)(10-9) = -21.$

Тогда оценка будет равна $(-21/4) = -5$.

Оценка коллизий. Энтропия зависит от того, сколько раз от источника получены случайные данные перед первым их повторением, т.е. коллизией. От источника IID можно ожидать одинаковое поведение относительно коллизии для преобразованных данных. При этом, оценка коллизий сводится к определению появления выборок из двух повторяющихся комбинаций. Однако возможны различные коллизии, в том числе и многократные.

Начиная с начала подмножества данных рассматриваются выборки, пока не будет найдено повторение. Число исследованных выборок фиксируется, и новый поиск начинается со следующего элемента подмножества. Это продолжается, пока не будет проверено все подмножество данных. Набор чисел из исследованных образцов затем используется для получения трех оценок: наименьшего числа, среднего числа и наибольшего числа.

В случае, если данные двоичные, подмножество, состоящее из двоичных данных, преобразуется в последовательность 8-битных байт таким образом, что каждая последовательность из 8 бит в оригинальном подмножестве становится одним байтом в модифицирован-

ном подмножестве данных. Затем измененное подмножество подвергают тестированию для вычисления баллов коллизии. Необходимо заметить, что оригинальное и преобразованное подмножества модифицируются отдельно. Необходимо отметить, что, что длина модифицированного подмножества равна $\left\lfloor \frac{N}{80} \right\rfloor$, в то время как длина исходного двоичного и недвоичного подмножества равны $\left\lfloor \frac{N}{10} \right\rfloor$.

Баллы рассчитываются следующим образом:

- *Counts* – это список счетчиков, который используется для нахождения повторения(коллизий) в выборках, изначально пуст – $Counts = 0$;

- $pos = 0$.

- пока $pos <$ (длина подмножества данных):

- находим наименьшее j такое, что $s_{pos} ... s_{pos+j}$ содержат одно повторяющееся значение выборки;

- если такого j не существует – прерываем цикл;

- добавить j в список $Counts$;

– $pos = pos + j + 1$.

В заключение вернуть следующие значения как оценок: минимальное значение в списке $Counts$, среднее значение в списке $Counts$, максимальное значение в списке $Counts$.

К примеру, пусть подмножество данных будет {2, 1, 1, 2, 0, 1, 0, 1, 1, 2} длиной, равной 10. Первое значение “2” находится на нулевой позиции, тогда первая коллизия происходит, когда $j = 2$ (единицы на первой и второй позиции). Не обращая внимания на первые три значения подмножества, необходимо рассмотреть коллизии в подмножестве {2, 0, 1, 0, 1, 1, 2}. Они возникают при $j = 3$(нули на первой и третьей позиции). Третьим подмножеством будет {1, 1, 2}, в котором коллизия будет при $j = 1$ (нулевая и первая позиции равны 1). Последнее подмножество {2}, для которого нет коллизий. Список счетчиков для этого теста будет $j = 2, j = 3, j = 1$. Первая оценка будет 1 (минимальное значение счетчика), вторая оценка 2 (среднее значение счетчиков), третья оценка 3 (максимальное значение счетчика).

Выводы

В стандарте NIST SP 800-90b приведены требования, которые выдвигаются в процессе создания и применения источника энтропии, обоснование этих требований и методики тестирования. При этом тестирование направлено на проверку корректной работоспособности и удовлетворения критериям источника энтропии, минимальной энтропии, полного источника энтропии, генератора шума, а также проверки компонентов состояния.

На сегодняшний день в Украине нет аналогичного национального стандарта, который содержал бы требования и методики тестирования источников энтропии, источников шума, определения минимальной энтропии, полной энтропии источника и компонентов состояния. Можно сделать вывод, что данный стандарт является мощным инструментом в создании и тестировании генераторов случайных последовательностей и является хорошим кандидатом на гармонизацию с последующим принятием его в качестве государственного стандарта Украины.

Список литературы: 1. *Federal* Information Processing Standards Publication (FIPS PUB) 140-2. Security requirements for cryptographic modules. NIST, 1999. 2. *National* Institute of Standards and Technology, FIPS 140-3 (DRAFT), Security for cryptographis modules: [Електронний ресурс]. Режим доступу: http://www. nist. gov/cmvp 3. *ISO/IEC* FCD 19790: Information technology- Security requirements for cryptographicmodules. Proect: 1.27.40. 4. *NIST SP 800-22*. A Statistical Test Suite for Random and Pseudorandom Number Generators for Cryptographic Applications: [Електронний ресурс]. April 2000. 5. *NIST* DRAFT Special Publication 800-90B. Recommendation for the Entropy Sources Used for Random Bit Generation / E. Barker, J. Kelse. 2012.

Харьковский национальный
университет радиоэлектроники

Поступила в редколлегию 15.09.2012