УДК 621.391:519.72

Н. Ф. Казакова, Ю. В. Щербина

Одесский национальный экономический университет, г. Одесса

ПРИМЕНЕНИЕ ТЕСТА АППРОКСИМАЦИОННОЙ ЭНТРОПИИ ДЛЯ АНАЛИЗА КРИПТОГРАФИЧЕСКИХ ГЕНЕРАТОРОВ ПСП

Приведено описание теста аппроксимационной энтропии из пакета STS NIST, и методика подбора оптимальных значений его входных параметров при испытании псевдослучайных последовательностей для нужд криптографии.

Ключевые слова: *аппроксимация, энтропия, тест, криптография, генератор ПСП, параметр, STS NIST.*

Постановка проблемы в общем виде и ее связь с важными научными и практическими задачами. Интерес специалистов, работающих в области криптографии, к совершенствованию средств тестирования генераторов псевдослучайных последовательностей (ПСП), объясняется стремлением к повышению стойкости создаваемых ими симметричных криптографических систем. Основным принципом, на котором базируются критерии стойкости современных шифров, является статистическая оценка равномерности распределения символов в гаммирующих последовательностях и в последовательностях, из которых формируются сеансовые ключи. К сожалению, опубликованные в этой области научные работы еще не сложились в самостоятельный раздел криптоанализа, базирующийся на фундаментальных методах математической статистики. К настоящему времени известно несколько тестовых пакетов [1-3], позволяющих выполнять ориентировочную оценку криптографической стойкости ПСП, предназначенных для нужд криптографии. Окончательный же вывод об их надежности делается путем применения различных методов взлома, что требует огромных временных, интеллектуальных и вычислительных ресурсов.

Анализ исследований и публикаций. В работе «Искусство программирования на ЭВМ» [4], сформулировано правило, которое сводится к тому, что число тестов, применяемых для анализа последовательности, не должно быть строго оговорено. Чем их больше, тем выше уровень доверия к полученному результату. Отчасти это так. По этой причине тесты, включаемые в состав предлагаемых пакетов, базируются на статистической обработке различных показателей ПСП, с последующей оценкой их отклонения от ожидаемых теоретических значений.

Пакет тестов STS NIST [1], рекомендуемых Институтом стандартизации США, принято считать основным. Он содержит шестнадцать тестов, которые позволяют оценивать степень соответствия распределения вероятностей символов равномерному закону в двоичных последовательностях. Несмотря на то, что в руководстве, распространяемом NIST, содержится методика тестирования и упрощенные примеры, при попытке применения этого пакета возникает много вопросов, связанных с выбором таких параметров статистического материала, как размеры тестируемых сегментов последовательностей и некоторых иных индивидуальных показателей для каждого рекомендуемого теста. Эти показатели не могут быть универсальными. В каждом тесте их следует определять опытным путем. Дополнительная проблема состоит в том, что приводимые теоретические обоснования некоторых тестов явно недостаточны. В частности, это относится к тесту аппроксимационой энтропии, описанию и результатам работы с которыми посвящена данная статья. Для достижения указанной **цели** авторами использованы их публикации [5-15]. Учитывая это, перейдём к изложению основного материала.

Тест аппроксимационной энтропии

Понятие аппроксимационной энтропии было введено в 1991 году в [16] применительно к задаче проверки строки двоичных символов на случайность. Оно отличается от понятия информационной энтропии, введенного К. Шенноном [17]. Шенноновское понятие энтропии входит в него как составная часть. Суть метода исследования, предложенного Пинкусом, заключается в проверке частот однотипных m-разрядных «скользящих» фрагментов тестируемой n-разрядной последовательности

$$\varepsilon=\varepsilon_1,\varepsilon_2,\ldots,\varepsilon_n.$$

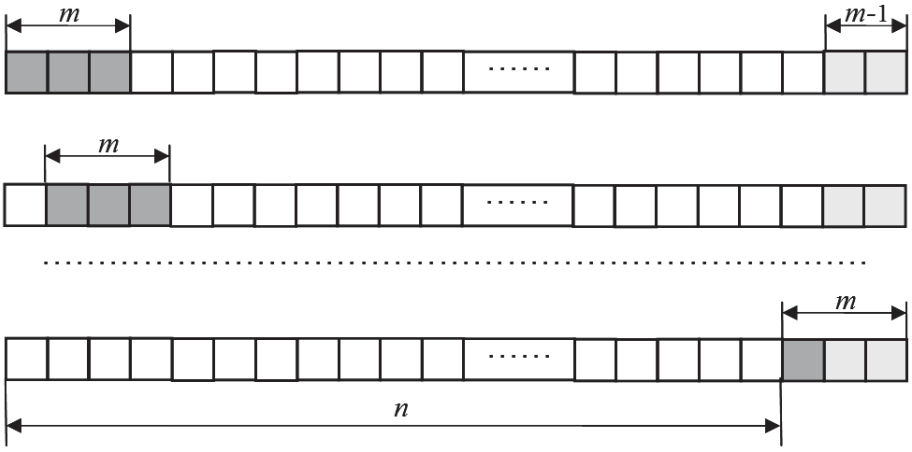

Рисунок 1 – Выделение фрагментов в тестируемом сегменте ПСП

Эти блоки (рис. 1) формируются с начала каждого символа тестируемого сегмента последовательности. Общее их число должно быть равно n. Для этого в конец n-разрядного сегмента добавляют $m-1$ символов следующего сегмента [1]. В этом случае общее число m-разрядных двоичных блоков равно 2^m. При высокой степени равномерности распределения вероятностей двоичных символов следует ожидать, что количества однотипных фрагментов будут также мало отличаться, а вероятности их появления в тестируемом сегменте p_i, $i = 1,2,\ldots 2^m$ будут стремиться к величине $1/2^m$. Однако в реальном случае эти величины отличаются, и степень этого отличия дает возможность оценить, насколько испытуемая последовательность «случайна» или «неслучайна». С этой целью, в процессе тестирования, предполагается выполнять оценку вероятностей p_i по правилу $p_i = k_i / 2^m$, где k_i – количество фрагментов i-го типа, встречающихся в тестируемом сегменте.

Рассматривая тестируемый сегмент последовательности как некоторый текст, составленный символами алфавита размером 2^m, и определив оценки величин p_i, можно посчитать значение шенноновской энтропии

$$\Phi^{(m)} = \sum_{i=1}^{2^m} p_i \log p_i ,$$

выражающее среднее количество информации на один символ 2^m-алфавита.

Значение введенной Пинкусом в [16] аппроксимационной энтропии определяется как

$$ApEn(m) = \Phi^{(m)} - \Phi^{(m+1)} ,$$

что представляет собой разницу энтропий, посчитанных для фрагментов с длинами m и $m+1$. В [18], показано, что для нерегулярных последовательностей величина ее приблизительной энтропии $ApEn(m)$ принимает наибольшее возможное значение. Напротив, если ее значение мало, испытуемая последовательность признается «неслучайной». Также в [18] показано, что при фиксированной длине блока m, следует ожидать, что в длинных «случайных» (нерегулярных) сегментах последовательности, $ApEn(m) \sim \log 2$, а предельное распределение $n[\log 2 - ApEn(m)]$ совпадает с распределением χ^2 с 2^m степенями свободы [19]. На основании этого утверждения, в этом же источнике, был предложен и обоснован количественный показатель, на основании которого может быть сделан вывод о качестве тестируемого сегмента псевдослучайной последовательности

$$P_v = igamc\left(2^{m-1}, \chi^2(n\,[\log 2 - ApEn(m)])/2\right) ,$$

где $igamc(m,x)$ – дополнительная неполная гамма-функция от аргументов m и x.

Руководство STS NIST [1] рекомендует признавать тестируемый сегмент последовательности «случайным», если полученное значение $P_v \geq 0.01$. В противном случае, он признается «неслучайным».

Если величина ошибки первого рода α, задается равной 0.01, тестируемая последовательность разделяется на более чем 100 сегментов и после тестирования определяется доля тех из них, которые не прошли тестирование. Если она оказывается меньше α, принимается решение о положительном завершении тестирования ПСП. При иных значениях α, поступают аналогично. Руководство также предлагает выбирать величину фрагмента m из условия

$$m < \lfloor \log_2 n \rfloor - 5 , \qquad (1)$$

однако в приводимых далее примерах это условие не всегда дает хороший результат. Кроме того, не существует и единого мнения о размерах тестируемых сегментов испытуемой последовательности. Этому вопросу посвящена публикация [20], в которой в результате анализа соответствующих нормативных документов, делается выводы о том, что к рекомендациям четвертого раздела руководства [1] следует относится с осторожностью, поскольку очевидно, что не может быть универсальных значений входных параметров для всех шестнадцати тестов, предлагаемых STS NIST. По результатам испытаний, приводимых в данной работе, делается вывод о том, что общая длина тестируемой последовательности должна колебаться от 20000 до 10^6 символов.

Скорее всего, при поиске ответа на вопрос о соотношении величин тестируемого сегмента n и фрагмента m, на основе которого рассчитывается аппроксимационная энтропия $ApEn(m)$, следует исходить из выражения (1), не забывая о том, что оно не является строгим. В работе [4] указывается на тот факт, что для достоверной оценки вероятностей p_i следует обеспечить величины k_i не менее 5. Это значит, что должно выполняться соотношение $n/2^m > 5$ и $n/2^{(m+1)} > 5$. Например, если n=1000, $m = \lfloor \log_2 1000 \rfloor - 5 = 4$, $1000/2^4 \approx 62$ и $1000/2^{(4+1)} \approx 31 > 5$, то значение m=4 с теоретической точки зрения подходит для тестирования при n=1000.

Однако опыт показывает, что при попытке создания программных продуктов, использование расчетных параметров не всегда обеспечивает желаемое качество тестирования. Так, например, тест, реализованный с параметрами приведенного выше примера и α=0.01, дает результаты тестирования, приведенные на рис. 2.

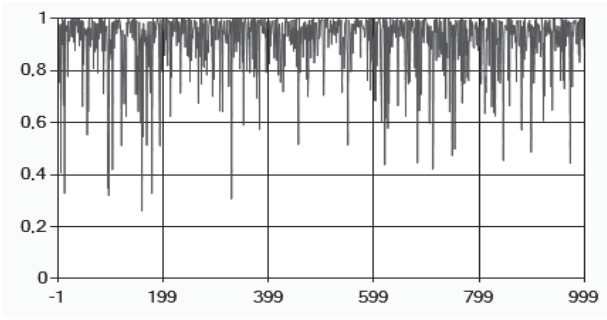

Рисунок 2 – Результаты тестирования ПСП, порождаемой генератором SHA-1. Используемые параметры теста: n=1000, m=4. Бракованных сегментов нет

Из приведенного графика видно, что все 1000 проверяемых 1000 битных сегментов дают положительный результат тестирования, а это говорит о том, чувствительность такого теста невысока. При правильной настройке теста и выбранном значении α, до одного процента сегментов должны быть забракованы (это логически вытекает из того факта, что в хорошей последовательности все сегменты, в том числе и те «случайные» сегменты, которые выглядят как «неслучайные», равновероятны). При увеличении значения m на единицу ($1000/2^{(5)} \approx 31 > 5$ и $1000/2^{(6)} \approx 15 > 5$), результаты тестирования изменятся, и будут иметь вид, приведенный на рис. 3.

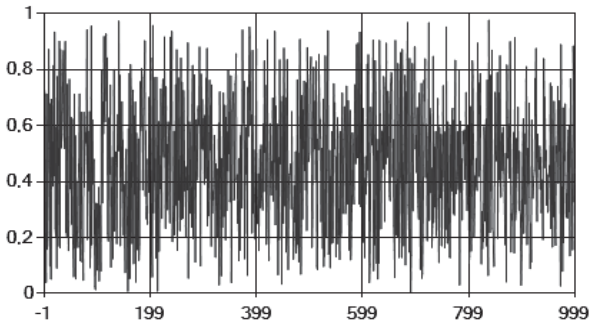

Рисунок 3 – Результаты тестирования ПСП, порождаемой генератором SHA-1. Используемые параметры теста: n=1000, m=5. Забраковано три сегмента

Как видно из приведенного графика, тест стал более чувствителен к выявлению «неслучайных» последовательностей (три сегмента из тысячи было забраковано, что укладывается в допуск, равный одному проценту). Дальнейшее увеличение числа m до 6 хотя и не нарушает требования, сформулированного в [4] ($1000/2^{(6)} \approx 15 > 5$) и условия (1) ($1000/2^{(7)} \approx 8 > 5$), приводит к увеличению вероятности ошибки второго рода – «случайные» сегменты последовательности признаются «неслучайными» (рис. 4.).

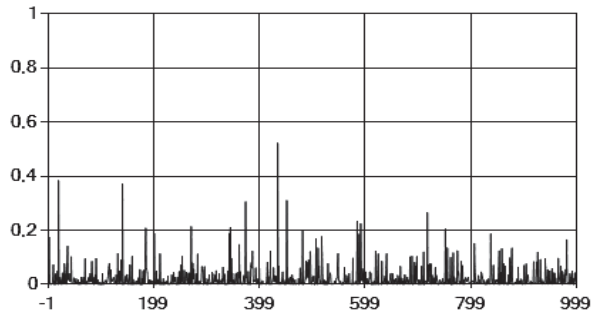

Рисунок 4 – Результаты тестирования ПСП, порождаемой генератором SHA-1. Используемые параметры теста: n=1000, m=6. Забраковано 542 сегмента

Выводы. При выборе тестов STS NIST, разработчики ориентировались на принцип их независимости. Все 16 тестов основаны на различных математических задачах из теории математической статистики. Естественно, что при этом трудно было выработать общие подходы к выбору для них входных параметров. В каждом конкретном случае попытка практической реализации того или иного теста требует некоторой доводки разрабатываемого программного обеспечения путем испытаний с применением заведомо качественных генераторов ПСП. В частности, что касается рассмотренного в статье теста аппроксимацтонной энтропии, то при увеличении размера статистического материала, например до 10^6 символов, для сохранения одинаковой чувствительности к ошибкам первого и второго рода, оптимальный размер фрагмента *m* придется определять заново.

Список использованных источников

1. A Statistical Test Suite for the Validation of Random Number Generators and Pseudo Random Number Generators for Cryptographic Applications. – NIST Special Publication 800-22. – May 15, 2001.
2. The Marsaglia Random Number CDROM including the Diehard Battery of Tests of Randomness [Електронний ресурс] // Портал : stat.fsu.edu. – Режим доступу \www/ URL: http:// www.stat.fsu.edu/pub/diehard**.** -- Заголовок з екрана, доступ вільний, 30.03.2015.
3. Statistical test suite Crypt-X [Електронний ресурс] // Портал : isi. – Режим доступу \www/ URL: http://www.isi.qut.edu.au/resources/cryptx**.** -- Заголовок з екрана, доступ вільний, 30.03.2015.
4. Кнут Д. Искусство программирования для ЭВМ : монография, Т. 2. – М. : Мир, 1977. – 727 с.
5. Казакова Н. Ф. аналитическое обоснование использования gfsr-генераторов в задачах криптографии [Текст] / Н. Ф. Казакова // Технологічний аудит та резерви виробництва. – 2013. – № 5/5(13). – С. 45-47.
6. Казакова Н. Ф. Аналіз стану розвитку інструментальних засобів для тестування вихідних послідовностей симетричних шифрувальних пристроїв [Текст] / Н. Ф. Казакова, Ю. В. Щербина // Метрологія та прилади. – 2011. – № 4(30). – С. 27-30.
7. Казакова Н. Ф. Застосування програмно реалізованого прогностичного контролю для вирішення практичних завдань забезпечення якості надання послуг у захищених інформаційних мережах [Текст] / Н. Ф. Казакова // Сучасна спеціальна техніка. – 2012. – № 2(29). – С. 86-95.
8. Казакова Н. Ф. Инструментальные средства для предварительного анализа криптографических программных генераторов ПСП [Текст] / Н. Ф. Казакова, Ю. В. Щербина // Сучасний захист інформації. – К. : ДУІКТ. – 2010. – № 1. – С. 31-35.
9. Казакова Н. Ф. Поэтапное тестирование и подбор составных элементов генераторов псевдослучайных последовательностей [Текст] / Н. Ф. Казакова // Восточно-европейский журнал передовых технологий. – 2010. № 2/8(44). – С. 44-48.
10. Казакова Н. Ф. Проблемы оценки качества работы современных линейных генераторов псевдослучайных последовательностей [Текст] / Н. Ф. Казакова, Ю. В. Щербина // Збірник наукових праць Одеської державної академії технічного регулювання та якості. – 2013. – № 1(2). – С 32-36.
11. Казакова Н. Ф. Проблемы построения комбинированных линейных генераторов псевдослучайных чисел [Текст] / Н. Ф. Казакова, Ю. В. Щербина // Інформаційна безпека. – 2013. – № 2(10). – С. 58-64.
12. Казакова Н. Ф. Программная реализация универсального статистического теста Маурера для анализа псевдослучайных последовательностей [Текст] / Н. Ф. Казакова, Ю. В. Щербина // Вісник Східноукраїнського національного університету імені Володимира Даля. – 2011. – № 7(161). – Т. 1. – С. 289-296.
13. Казакова Н. Ф. Статистическое тестирование криптографических генераторов [Текст] / Н. Ф. Казакова, Ю. В. Щербина // Вісник Східноукраїнського національного університету імені Володимира Даля. – 2010. – № 9(151). – Т. 1. – С. 29-36.
14. Казакова Н. Ф. Частотное тестирование криптографических генераторов псевдослучайных последовательностей [Текст] / Н. Ф. Казакова, Ю. В. Щербина // Сучасний захист інформації. – 2010. – № 3. – С. 51-57.
15. Щербина Ю. В. Обобщение результатов тестирования генераторов псевдослучайных оследовательностей [Текст] / Ю. В. Щербина, Н. Ф. Казакова // Інформаційна безпека. – 2012. – № 1(7). – С. 90-95.
16. Pincus S. Approximate entropy as a measure of system complexity [Текст] / S. Pincus // Proceedings of the National Academy of Sciences of the USA. – 1991. – P. 2297-2301.
17. Шеннон К. Э. Математическая теория связи [Текст] / К. Э. Шеннон // Работы по теории информации и кибернетике. – 1963. – С. 243-322.
18. Pincus S. Not all (possibly) «random» sequences are created equal. [Текст] / S. Pincus, R. E.

Kalman // Proc. Natl. Acad. Sci. USA. – 1997. – Vol. 94. – P. 3513-3518.

19. Rukhin A. Approximate entropy for testing randomness [Текст] / A. Rukhin // Journal of Applied Probability. – 2000. – Vol. 37. – P. 24-29.

20. Hill J. ApEn Test Parameter Selection [Електронний ресурс] / J. Hill // Портал : untruth.org. – Режим доступу \www/ URL: http://www.untruth.org/~josh/papers/ApEn%20test%20parameter%20selection.pdf**.** -- Заголовок з контейнера, доступ вільний, 11.09.2014.

Поступила в редакцию 20.05.2015

Рецензент: д.т.н., проф. Скопа А. А., Одеський національний економічний університет, м. Одеса.

Н. Ф. Казакова, Ю. В. Щербина

ЗАСТОСУВАННЯ ТЕСТУ АПРОКСИМАЦІЙНОЇ ЕНТРОПІЇ, ДЛЯ АНАЛІЗУ КРИПТОГРАФІЧНИХ ГЕНЕРАТОРІВ ПВП

Наведено опис тесту апроксимаційної ентропії з пакету STS NIST, і методика підбору оптимальних значень його вхідних параметрів під час випробування псевдовипадкових послідовностей для потреб криптографії.

Ключові слова: *апроксимація, ентропія, тест, криптографія, генератор ПВП, параметр, STS NIST.*

N. F. Kazakova, Ju. V. Shherbina

APPLICATION TEST ANALYSIS APPROXIMATION ENTROPY FOR CRYPTOGRAPHIC GENERATOR OF PSEUDO-RANDOM SEQUENCE

The description of the test approximate entropy of packet STS NIST, and methods of selection of the optimal values of its input parameters in the test pseudorandom sequences for the needs of cryptography.

Keywords: *approximation, entropy, test, cryptography, pseudo-random sequence, parameter, STS NIST.*

УДК 004.45:004.057.02

Є. В. Вавілов

Одеська державна академія технічного регулювання та якості, м. Одеса

СЕРІЯ СТАНДАРТІВ SQuaRE ЯК ОСНОВА ЗАБЕЗПЕЧЕННЯ ВИМОГ ДО ЯКОСТІ ТА ОЦІНКИ ПРОГРАМНИХ ЗАСОБІВ

Розглянуто серію перспективних міжнародних стандартів ISO/IEC 25000 – ISO/IEC 25099. Зазначені стандарти регламентують якість програмних продуктів, які відносяться до особливих прикладних областей застосування. Показано, що стандарти, які розглядаються, доповнюють існуючі стандарти якості програмних засобів, об'єднують вже діючі стандарти або уточняють їх.

Ключові слова. *Стандарт, SQuaRE, якість, оцінка, міра, ISO/IEC, програма.*

Постановка проблеми у загальному вигляді та її зв'язок із важливими науковими та практичними завданнями. Застосування стандартів в області інженерії програмного забезпечення (ПЗ) відіграє велику роль. За останні роки спостерігається бурхливий ріст кількості різноманітного ПЗ. Це ставить гостре питання про його якість [1-10]. На сьогоднішній день одним з найефективніших методів розв'язку даної проблеми є розробка та національна адаптація стандартів, що регламентують процес проектування якісного ПЗ. Існує безліч визначень цього поняття. Так, наприклад, згідно до стандарту 1061-1998 IEEE «Standard for Software Quality Metrics Methodology», якість програмного забезпечення визначено як ступінь, у якому ПЗ має необхідну комбінацію властивостей. У міжнародному стандарті ISO 8402:1994 «Quality management and quality assurance» якість ПЗ трактується як сукупність характеристик ПЗ, що ставляться до його