УДК 621. 3.06

О МЕТОДИКЕ ОЦЕНКИ ЗАКОНОВ РАСПРЕДЕЛЕНИЯ ВЕРОЯТНОСТЕЙ МАКСИМУМОВ ПОЛНЫХ ДИФФЕРЕНЦИАЛОВ И СМЕЩЕНИЙ ЛИНЕЙНЫХ ОБОЛОЧЕК БЛОЧНЫХ СИММЕТРИЧНЫХ ШИФРОВ

К.Е. ЛИСИЦКИЙ

Представляется уточнённая (исправленная) версия известной методики определения законов распределения максимумов для множеств независимых переменных, имеющих одни и те же законы распределения, в частности для переменных с пуассоновским законом распределения вероятностей и нормальным законом распределения вероятностей.

Ключевые слова: закон распределения максимумов, независимые случайные переменные, максимумы полных дифференциалов, максимумы смещений линейных корпусов, переменные с пуассоновским законом распределения, переменные с нормальным законом распределения,

ВВЕДЕНИЕ

В этой статье речь будет идти об известной (опубликованной) методике определения законов распределения максимумов для множеств независимых переменных, имеющих одни и те же законы распределения, в частности для переменных с пуассоновским законом распределения вероятностей и нормальным законом распределения вероятностей.

Интерес к этой методике был вызван тем, что как оказалось множествами независимых переменных, имеющими одни и те же законы распределения, в частности переменными с пуассоновским законом распределения вероятностей и нормальным законом распределения вероятностей можно описать переходы таблиц полных дифференциалов и смещения таблиц линейных аппроксимаций блочных симметричных шифров, рассматриваемых как подстановочные преобразования.

Эта методика в работе [1] была успешно применена для выполнения вычислительных экспериментов по определению законов распределения максимумов переходов таблиц полных дифференциалов и максимумов смещений таблиц линейных аппроксимаций для уменьшенных моделей шифров и, в частности, шифра из нового белорусского стандарта. Были выполнены также оценки ожидаемых значений максимумов дифференциальных переходов и максимумов смещений для полномасштабного шифра Rijndael.

В соответствии с полученными для малых моделей шифров результатами сделан важный вывод о том, что все современные шифры имеют весьма малый диапазон изменения максимумов полных дифференциалов и максимумов смещений линейных корпусов. Это означает, что практически для оценки показателей доказуемой стойкости этих шифров можно пользоваться результатами оценок максимальных дифференциальных вероятностей и максимальных линейных вероятностей, вычисленных для произвольно взятого (одного) ключа зашифрования. Этот результат в соответствии с новой методологией оценки показателей стойкости блочных симметричных шифров к атакам дифференциального и линейного криптоанализа, развиваемой в работах И.В. Лисицкой [2, 3 и др.], может быть распространён и на полномасштабные версии этих шифров, что позволяет перейти к практическому вычислению показателей их стойкости к отмеченным атакам.

Математический аппарат, который успешно решает в общетеоретическом плане эту задачу, представлен в совместной работе учёных Joan Daemen, Vincent Rijmen [4].

Более глубокий анализ материалов этой работы, однако, показал, что по ходу изложения методики её авторами был допущен ряд неточностей (м.б. описок), вследствие чего излагаемые исходные положения не согласуются с результатами их применения к рассмотренным законам распределения вероятностей.

Следует отметить, что эти неточности не повлияли на правильное представление итоговых результатов. Соответственно правильными следует считать и результаты исследований, представленные в работе [1]. Тем не менее, мы посчитали важным навести порядок в описании этой важной на наш взгляд методики и представить её в исправленном виде.

В этой статье предлагается изложение этой методики вместе с исправлениями, позволяющими устранить несогласованность исходных положений и результатов их применения к конкретным законам распределения вероятностей.

1. ЗАКОН РАСПРЕДЕЛЕНИЯ МАКСИМУМОВ

Мы приведём здесь изложение материалов приложения из работы [4], в котором приводятся все необходимые теоретические сведения, составляющие сущность развиваемого подхода, тем более, что мы здесь имеем возможность исправить обнаруженные в представленном материале ошибки (описки).

Рассматривается случай, когда все значения множества независимых случайных переменных x имеют одни и те же распределения. Считается, что их плотности распределения убывают экспоненциально при больших значениях x. Обозначим, пишут авторы, число таких значений 2^Y и воспользуемся моделью интегрального распределения $D(X)$ в виде

$$D(X) = 1 - e^{-f(X)} \quad (1)$$

(на самом деле, как будет видно из дальнейшего, авторы цитируемой работы рассматривают интегральное распределение в виде $D(X) = 1 - f(X)$), с $f(X)$ – функцией, описывающей плотности распределения вероятностей независимых значений X.

Из порядковых статистик [5,6] известно, отмечают они, что это интегральное распределение максимального числа значений является произведением интегральных распределений этих значений. Поэтому мы имеем:

$$D_{\max}(X) = D(X)^{2^Y} = (1 - e^{-f(X)})^{2^Y} \approx \quad (2)$$

$$\approx e^{-2^Y e^{-f(X)}} = e^{-e^{\ln(2)Y - f(X)}}.$$

На самом деле здесь должен быть записан результат

$$D_{\max}(X) = D(X)^{2^Y} =$$

$$= (1 - f(X))^{2^Y} \approx e^{-2^Y \cdot f(X)} = e^{-e^{\ln(2)Y + \ln f(X)}}.$$

Мы можем аппроксимировать функцию $\ln(2)Y - f(X)$ (должно быть $\ln(2)Y + \ln f(X)$), говорят авторы работы [4], линейной функцией в окрестности точки, где функция близка к нулю. Пусть a будет решением уравнения

$$\ln(2)Y = f(X)$$

(должно быть уравнения $\ln(2)Y + \ln f(X) = 0$) и пусть b будет единицей, делённой на производную функции $f(x)$ (должно быть функции $\ln f(X)$) в точке a. Тогда справедливо выражение

$$D_{\max}(X) \approx e^{-e^{\frac{a-X}{b}}}. \quad (3)$$

Это распределение хорошо изучено в теории вероятностей, отмечается в [4], и известно как распределение экстремальных значений, Fisher-Tippett распределение или лог-Вейбулла распределение [5, 6]. Соответствующая плотность изображена на рис. 1, заимствованном из цитируемой работы [4]. Отмечается, что пик этой функции есть a, а ее ширина пропорциональна b. Это распределение имеет математическое ожидание $\mu(X) = a + b\gamma$ с $\gamma \approx 0{,}58$ и среднеквадратическое отклонение $\frac{\pi}{\sqrt{6}} b \approx 1{,}3b$. Авторы замечают, что справедливость выражения (3) зависит от качества линейной аппроксимации $f(x)$ вблизи точки $(a, 0)$.

Далее авторы работы [4] конкретизируют представление плотности вероятности (3) для переменных, подчиняющихся некоторым известным законам распределения. Нас будут интересовать два закона распределения: пуассоновский и нормальный. Первый соответствует закону распределения вероятностей переходов XOR таблицы случайной подстановки (шифра), второй – закону распределения вероятностей смещений случайной подстановки (шифра).

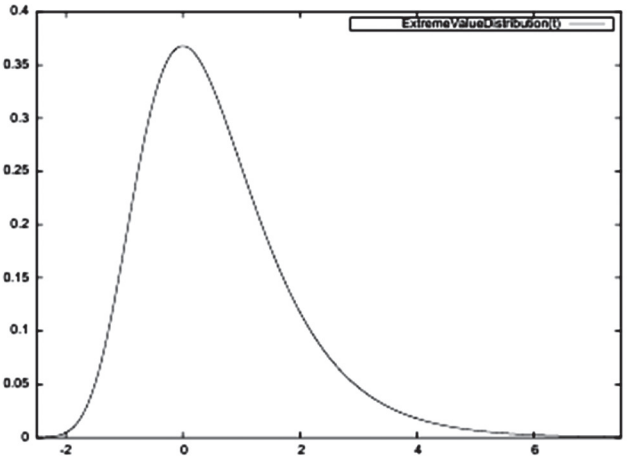

Рис. 1. Распределение экстремальных значений при $a = 0$, $b = 1$

1.1. МАКСИМУМЫ НАД X С ПУАССОНОВСКИМ РАСПРЕДЕЛЕНИЕМ

Дальнейший материал повторяет результаты, представленные в [1 и 4]. Если максимум берется по переменным с распределением Пуассона, мы, замечают авторы цитируемой работы, должен приниматься во внимание дискретный характер последнего. Однако, мы можем получить выражения для среднего значения и стандартного отклонения максимумов, если приблизим распределение Пуассона непрерывной функцией. Имеем:

$$\Phi(i;\lambda) = \sum_{x=0}^{i-1} \mathbf{Poisson}(x;\lambda) = 1 - \sum_{x \geq i} \mathbf{Poisson}(x;\lambda). \quad (4)$$

Для $i >> \lambda$, это выражение в соответствии с [5, 6] аппроксимируется так:

$$\Phi(i;\lambda) \approx 1 - \left(1 - \frac{\lambda}{i}\right) \cdot \mathbf{Poisson}(i;\lambda) \approx \quad (5)$$

$$\approx \mathbf{Poisson}(i,\lambda) = e^{-\lambda} \frac{\lambda^i}{i!}.$$

Далее, используя приближение Стирлинга для факториала [5,6], мы получим следующее выражение для функции $f(i)$:

$$f(i) = \frac{1}{2}\ln(2\pi) + \lambda + i\ln i - (1 + \ln\lambda)i + \frac{1}{2}\ln(i). \quad (6)$$

Конечно это выражение для логарифма функции $f(i)$, а не самой функции, как пишут авторы цитируемой работы, т.е. вместо (6)необходимо рассматривать

$$\ln f(i) = \frac{1}{2}\ln(2\pi) + \lambda + i\ln i - (1 + \ln\lambda)i + \frac{1}{2}\ln(i).$$

Если теперь абстрагироваться от факта, что i должно быть целым числом, пишут авторы ци-

тируемой работы, то мы можем вычислить параметр a путем решения уравнения:

$$\ln(2)y = \frac{1}{2}\ln(2\pi) + \lambda + i\ln i - (1+\ln\lambda)i + \frac{1}{2}\ln(i), \quad (7)$$

или, что эквивалентно:

$$i = \frac{\ln(2)y - \frac{1}{2}\ln(2\pi i) - \lambda}{\ln\left(\frac{i}{\lambda}\right) - 1}. \quad (8)$$

Последнее уравнение может быть решено итеративно. Производная функции $f(i)$ (должно быть функции $\ln f(i)$) определяется по формуле:

$$\ln\left(\frac{i}{\lambda}\right) + \frac{1}{2i}. \quad (9)$$

Определив a и используя условие $a >> \lambda$, мы имеем:

$$b = \frac{1}{\ln\left(\frac{a}{\lambda}\right)}. \quad (10)$$

Отсюда следует, заключают авторы цитируемой работы, что если a намного больше, чем λ, стандартное отклонение становится меньше 1.

Поскольку распределение максимума дискретное, то эта малая величина стандартного отклонения приводит к тому, что распределение сосредоточено в двух целочисленных значениях вблизи значения a.

1.2. МАКСИМУМЫ НАД X С НОРМАЛЬНЫМ РАСПРЕДЕЛЕНИЕМ

Здесь рассматривается частный случай, для переменной x со стандартным нормальным распределением. Мы имеем (цитируются результаты из работы [4]):

$$D(x) \approx \int_{-\infty}^{\infty} Z(u)du. \quad (11)$$

При больших x, это интегральное распределение, становится близким к [13, 14]:

$$D(x) \approx 1 - \frac{1}{x}Z(x) \approx 1 - \frac{1}{x\sqrt{2\pi}}e^{-\frac{x^2}{2}}. \quad (12)$$

Из этого результата мы можем получить следующее выражение для $f(x)$ (должно быть для $\ln f(x)$), и мы сразу запишем выражение так, как надо:

$$\ln f(x)):$$

$$= -\ln\left(\frac{1}{x}Z(x)\right) = \frac{1}{2}\left(\ln(2\pi) + x^2\right) + \ln(x) \quad (13)$$

Параметр, a_s (подстрочный индекс s для стандарта) является решением уравнения

$$a_s = \sqrt{2\ln(2)y - \ln(2\pi) - 2\ln(a_s)}, \quad (14)$$

которое может быть найдено итеративным путём, не обращая внимания на правый член в первой итерации. Производная функции $f(x)$ (должно быть $\ln f(x)$) определяется по формуле:

$$x + \frac{1}{x}, \quad (15)$$

и, следовательно,

$$b_s = \frac{a_s}{a_s^2 + 1} \approx \frac{1}{a_s}. \quad (16)$$

Грубо говоря, максимум имеет распределение со средним значением $1{,}17\sqrt{y}$ и стандартным отклонением $1{,}11/\sqrt{y}$. Мы, отмечают авторы работы [4], можем найти значения a и b для любого нормального закона распределения со средним $\mu(X)$ и стандартным отклонением σ, заменив x на $\frac{X - \mu(X)}{\sigma}$. Это дает:

$$\begin{aligned} a &= \sigma a_s + м(X), \\ b &= \sigma b_s. \end{aligned} \quad (17)$$

В общем, итоговые результаты работы [4] следует считать правильными.Они и были использованы при проведении экспериментов в работе [1].

Таким образом, изложенная методика оценки максимумов распределений, характерных для переходов таблиц полных дифференциалов и смещений таблиц линейных аппроксимаций шифров (случайных подстановок) с представленными правками вполне может быть применена для оценки максимумов полных дифференциалов и смещений блочных симметричных шифров.

Литература

[1] *Lisitskiy K.E.* On Maxima Distribution of Full Differentials and Linear Hulls of Block Symmetric Ciphers [Text] / K.E. Lisitskiy // I.J. Computer Network and Information Security, 2014, 1, 11-18 Published Online November 2013 in MECS (http://www.mecs-press.org/) DOI: 10.5815/ijcnis. 2014.01.02.

[2] *Лисицкая И.В.* Методология оценки стойкости блочных симметричных шифров / И.В. Лисицкая // Автоматизированные системы управления и приборы автоматики. – 2011. – № 163. – С. 123–133.

[3] Долгов Виктор Иванович. Методология оценки стойкости блочных симметричных шифров к атакам дифференциального и линейного криптоанализа: монография. / В.И. Долгов, И.В. Лисицкая. – Харьков. Издательство “Форт”, 2013. – 420 с.

[4] Joan Daemen, Vincent Rijmen Probability distributions of Correlation and Differentials in Block Ciphers. / Joan Daemen, Vincent Rijmen // April 13, 2006/ — P. 1–38.

[5] W. Feller An Introduction to Probability Theory and Its Applications, Vol.1. Wiley&Sons. 1968.

[6] Mathworld. http: //mathworld. wolfram.com.

Поступила в редколлегию 27.10.2015

Лисицкий Константин Евгеньевич, студент Харьковского национального университета им. В.Н. Каразина. Научные интересы: криптография, методы криптоанализа.

УДК 621. 3.06

Про методику оцінки законів розподілу ймовірностей максимумів повних диференціалів і зсувів лінійних оболонок блокових симетричних шифрів / К.Е. Лисицький // Прикладна радіоелектроніка: наук.-техн. журнал. – 2015. – Том 14. – № 4. – С. 335–338.

Дається уточнена (виправлена) версія відомої методики визначення законів розподілу максимумів для множин незалежних змінних, що мають одні й ті самі закони розподілу, зокрема для змінних з пуассонівським законом розподілу ймовірностей і нормальним законом розподілу ймовірностей.

Ключові слова: закон розподілу максимумів, незалежні випадкові змінні, максимуми повних диференціалів, максимуми зміщень лінійних корпусів, змінні з пуассонівським законом розподілу, змінні з нормальним законом розподілу.

Іл.: 1. Бібліогр.: 6 найм.

UDC 621. 3.06

About the methodology for assessing the probability distribution laws of maxima of total differentials and offsets of block symmetric ciphers linear hulls / K.E. Lisitsky // Applied Radio Electronics: Sci. Journ. — 2015. — Vol. 14. — № 4. — P. 335–338.

The paper provides a refined (corrected) version of the well-known methods for determining the maxima distribution laws for a set of independent variables with the same distribution laws, in particular for variables with Poisson probability distribution law and normal distribution law of probabilities.

Keywords: maxima distribution laws, independent random variables, total differentials maxima, maxima linear hulls offsets, variables with Poisson distribution law, variables with normal distribution law.

Fig. 1. Ref.: 6 items.