УДК 519.92

ДИСПЕРСИЯ РАСПРЕДЕЛЕНИЯ ВЕРОЯТНОСТЕЙ ОШИБКИ ПРИ НАБЛЮДЕНИИ ВЕКТОРОВ ПЕРЕХОДОВ

И.П. Кобяк

Учреждение образования «Белорусский государственный университет информатики и радиоэлектроники», Минск, Беларусь

Аннотация. В работе исследован метод идентификации последовательностей случайных событий с помощью оценок наблюдения векторов переходов заданного вида. Рассмотрена дисперсия плотности распределения вероятностей ошибки, требуемая для задач сравнения статистических показателей предлагаемого и известных алгоритмов свертки. Основой для расчетов послужила производящая функция, позволяющая представить произведение комбинаторных моментов, характеризующих вероятности пропуска ошибки, в виде суммы произведений статистик на соответствующие моменты. Полученные в работе соотношения характеризуют метод наблюдения векторов переходов как наиболее эффективный алгоритм формирования оценок для идентификации сообщений со случайной или псевдослучайной природой.

Ключевые слова: сложные вероятностные события; дисперсия; идентификация последовательностей; вероятность пропуска ошибки; сигнатурный анализ; производящая функция.

THE VARIANCE OF THE ERROR PROBABILITY DISTRIBUTION WHEN OBSERVING TRANSITION VECTORS

I.P. Kobiak

Educational Institution "Belarusian State University of Informatics and Radioelectronics", Minsk, Belarus

Annotation. The paper investigates a method for identifying sequences of random events using estimates of the observation of transition vectors of a given type. The variance of the error probability distribution density required for the task of comparing statistical indicators of the proposed and known convolution algorithms is considered. The basis for the calculations was the generating function, which makes it possible to represent the product of combinatorial moments characterizing the probability of missing an error in the form of the sum of the products of statistics for the corresponding moments. The relations obtained in the work characterize the method of observing transition vectors as the most effective algorithm for generating estimates for identifying messages with a random or pseudorandom nature.

Keywords: complex probabilistic events; variance; sequence identification; probability of missing an error; signature analysis; generating function.

Введение

Шифрование сообщений с помощью шумоподобных последовательностей ставит под вопрос задачу обнаружения детерминизма в наблюдаемых каналах связи. При этом проблема обнаружения самих каналов оказывается весьма трудоемкой даже для самых современных компьютеров.

Классический поиск сообщений циркулирующих в криптосетях, как правило, осуществляется с помощью теоретических и эвристических алгоритмов анализа. Соответственно к теоретическим относят математические методы, основанные на решении задач, на основе теории вероятностей, а к эвристическим - методы свертки по модулю два и другие аппаратные или программно-аппаратные подходы. Недостатком всех известных теоретических алгоритмов является факт равенства единице нормированной интегральной суммы моментов под кривой функции распределения вероятностей пропуска ошибки.

На сегодняшний день одним из самых современных методов обнаружения детерминизма в r-разрядных случайных процессах следует считать методологию наблюдения событий на основе двух и более элементарных векторов [1]. При этом вероятность пропуска ошибки статистикой сложных событий оказывается существенно ниже, чем при регистрации элементарных событий.

В целом, вопросы, связанные с применением новых методов, таких, например, как наблюдение векторов переходов заданного вида (ВПЗВ), требуют детального анализа известных параметров, а именно: математического ожидания и дисперсии производящей функции (ПФ) или энумератора для заданного числа переменных. Соответственно в представляемой работе рассмотрена дисперсия распределения вероятностей пропуска ошибки при наблюдении ВПЗВ, используемая как база для определения свойств оценок данного параметра.

Дисперсия распределения вероятностей ошибки

Для определения дисперсии распределения вероятностей пропуска ошибки при наблюдении ВПЗВ используем классическое соотношение вида:

$$D_{ifc} = P''_{ifc} - \left(P'_{ifc}\right)^2, \tag{1}$$

где ifc – функция корреляции, P_{ifc} – перечисляющая ПФ [1] для вероятностей пропуска ошибки при наблюдении ВПЗВ, соответственно P'_{ifc}, P''_{ifc} две производные ПФ.

Первая производная вероятности в функции (1) равна:

$$P'_{ifc} = \sum_{g} \pi(g)(n-g)\left[\sum_{j=1}^{0.5(n-5)} \xi_j k_{j,1} p^j e^{jt} x_j^1 + \sum_{j=1}^{0.5(n-5)} \xi_j \sum_{i=2}^{n-2j-2} k_{j,i} p^j e^{jt} x_j^i \frac{1}{2^{i+1}} \beta_{j,i} + \right.$$

$$\left. + \sum_{j=1}^{0.5(n-3)} \xi_j k_{j,n-2j} p^j e^{jt} x_j^{n-2j} \frac{1}{2^{n-2j+1}} \beta_{j,n-2j} + \xi_{0.5(n-1)} k_{\frac{n-1}{2},1} p^{\frac{n-1}{2}} e^{\frac{n-1}{2}t} x_{\frac{n-1}{2}}^1 \right]^{n-g-1} \times \tag{2}$$

$$\times \left[\sum_{j=1}^{0,5(n-5)} \xi_j k_{j,1} p^j \left(\frac{\partial}{\partial t} e^{jt}\right) x_j^1 + \sum_{j=1}^{0,5(n-5)} \xi_j \sum_{i=2}^{n-2j-2} k_{j,i} p^j \left(\frac{\partial}{\partial t} e^{jt}\right) x_j^i \frac{1}{2^{i+1}} \beta_{j,i} + \right.$$

$$+\sum_{j=1}^{0,5(n-3)}\xi_j k_{j,n-2j}p^j\left(\frac{\partial}{\partial t}e^{jt}\right)x_j^{n-2j}\frac{1}{2^{n-2j+1}}\beta_{j,n-2j}+\xi_{0,5(n-1)}k_{\frac{n-1}{2},1}p^{\frac{n-1}{2}}\left(\frac{\partial}{\partial t}e^{\frac{n-1}{2}t}\right)x_{\frac{n-1}{2}}^1\Bigg]^{n-g-1}.$$

Практическое моделирование данной задачи показало, что ВПЗВ с параметром $j = 3$ встречаются достаточно редко, а параметр $j = 4$, например, при $r = 8$ отсутствует в течение весьма длительного времени наблюдения. Следовательно, для верхней границы вероятности p требуемые множители ξ_1, ξ_2, ξ_3 в (2) могут быть определены в соответствии с равенствами:

$$\xi_1 = \frac{1}{1,146}\left(1-\frac{3}{16}\right),\quad \xi_2 = \frac{1}{1,146}\cdot\frac{3}{16},\quad \xi_3 = \frac{0,146}{1,146}.$$

Подставляя данные значения в многочлен (2) получаем:

$$P'_{ifc} \approx n\left[\frac{33}{8}\left(\xi_1 p+\xi_2 p^2+\xi_3 p^3\right)\right]^{n-1}\left[\frac{33}{8}\left(\xi_1 p+2\xi_2 p^2+3\xi_3 p^3\right)\right] = 0,606n\left(0,576\right)^{n-1}. \quad (3)$$

Повторное дифференцирование равенства (2), с использованием вспомогательного параметра e^t ПФ, приводит к соотношению:

$$P''_{ifc} = \sum_g \pi(g)(n-g)(n-g-1)\left(k_{1,1}pe^t x_1^1+\sum_{i=2}^{n-4}k_{1,i}pe^t x_1^i\frac{1}{2^{i-1}}\beta_{1,i}+k_{1,n-2}pe^t x_1^{n-2}\frac{1}{2^{n-1}}\beta_{1,n-2}\right)^{n-g-2}\times$$

$$\times\left(k_{1,1}pe^t x_1^1+\sum_{i=2}^{n-4}k_{1,i}pe^t x_1^i\frac{1}{2^{i-1}}\beta_{1,i}+k_{1,n-2}pe^t x_1^{n-2}\frac{1}{2^{n-1}}\beta_{1,n-2}\right)^2+$$

$$+\sum_g \pi(g)(n-g)\left(k_{1,1}pe^t x_1^1+\sum_{i=2}^{n-4}k_{1,i}pe^t x_1^i\frac{1}{2^{i+1}}\beta_{1,i}+k_{1,n-2}pe^t x_1^{n-2}\frac{1}{2^{n-1}}\beta_{1,n-2}\right)^{n-g-1}.$$

При $t = 0, x_1^i = 1$ и $\pi(g) = 1$ полученное равенство приводится к виду:

$$P''_{ifc} = n(n-1)\left(k_{1,1}p+\sum_{i=2}^{n_1-4}k_{1,i}p\frac{1}{2^{i+1}}\beta_{1,i}+k_{1,n_1-2}p\frac{1}{2^{n_1-1}}\beta_{1,n_1-2}\right)^{n-2}\times \quad (4)$$

$$\times\left[k_{1,1}p+\sum_{i=2}^{n_1-4}k_{1,i}p\frac{1}{2^{i+1}}\beta_{1,i}+k_{1,n_1-2}p\frac{1}{2^{n_1-1}}\beta_{1,n_1-2}\right]^2+n\left(k_{1,1}p+\sum_{i=2}^{n_1-4}k_{1,i}p\frac{1}{2^{i-1}}\beta_{1,i}+k_{1,n_1-2}p\frac{1}{2^{n_1-1}}\beta_{1,n_1-2}\right)^{n-1}.$$

Формирование суммы моментов в функции (4) выполним с учетом $k_{1,i} = 1$ определив вспомогательную статистику $n_1 >> 1$ при $n \ge n_1$. Тогда:

$$P''_{ifc} = n(n-1)\left(p+\sum_{i=2}^{n_1-4}p\frac{1}{2^{i-1}}\beta_{1,i}+p\frac{1}{2^{n_1-1}}\beta_{1,n_1-2}\right)^{n-2}\times$$

$$\times\left(p+\sum_{i=2}^{n_1-4}p\frac{1}{2^{i-1}}\beta_{1,i}+p\frac{1}{2^{n_1-1}}\beta_{1,n_1-2}\right)^2+n\left(p+\sum_{i=2}^{n_1-4}p\frac{1}{2^{i-1}}\beta_{1,i}+p\frac{1}{2^{n_1-1}}\beta_{1,n_1-2}\right)^{n-1}.$$

С учетом результатов работы [2] то есть для максимального значения p можем записать:

$$P''_{ifc} = n(n-1)\left(p\frac{33}{8}\right)^{n-2}\left(p\frac{33}{8}\right)^{2} + n\left(p\frac{33}{8}\right)^{n-1} \approx n^2\left(\frac{99}{128}\right)^{n}.$$

Таким образом, для минимальной суммы моментов энумератора, дисперсия плотности распределения вероятностей ошибки при $j=1$ и $i=var$ имеет численное значение:

$$D_{ifc} \approx n^2\left(\frac{99}{128}\right)^{n} - \left[n\left(p\frac{33}{8}\right)^{n}\right]^2 \approx n^2\left(\frac{99}{128}\right)^{n}\left[1-\left(\frac{99}{128}\right)^{n}\right] \xrightarrow{n\to\infty} n^2\left(\frac{99}{128}\right)^{n}.$$

Усложним вид производящей функции, используя параметр $j=1,2$. Тогда вторая производная ПФ может быть представлена в виде:

$$P''_{ifc} = P''_{ifc}(1) + P''_{ifc}(2).$$

Для первой части функции имеем:

$$P''_{ifc}(1) = \left\{\sum_{g}\pi(g)(n-g)(n-g-1)\left[\xi_1\left(k_{1,1}pe^{t}x_1^{1} + \sum_{i=2}^{n-4}k_{1,i}pe^{t}x_1^{i}\frac{1}{2^{i+1}}\beta_{1,i} + k_{1,n-2}pe^{t}x_1^{n-2}\frac{1}{2^{n-1}}\beta_{1,n-2}\right) + \right.\right.$$

$$\left.\left. + \xi_2\left(k_{2,1}p^2e^{2t}x_2^{1} + \sum_{i=2}^{n-6}k_{2,i}p^2e^{2t}x_2^{i}\frac{1}{2^{i+1}}\beta_{2,i} + k_{2,n-4}p^2e^{2t}x_2^{n-4}\frac{1}{2^{n-3}}\beta_{2,n-4}\right)\right]^{n-g-2}\right\} \times$$

$$\times\left[\xi_1\left(k_{1,1}pe^{t}x_1^{1} + \sum_{i=2}^{n-4}k_{1,i}pe^{t}x_1^{i}\frac{1}{2^{i+1}}\beta_{1,i} + k_{1,n-2}pe^{t}x_1^{n-2}\frac{1}{2^{n-1}}\beta_{1,n-2}\right) + \right.$$

$$\left. + \xi_2\left(k_{2,1}2p^2e^{2t}x_2^{1} + \sum_{i=2}^{n-6}k_{2,i}2p^2e^{2t}x_2^{i}\frac{1}{2^{i-1}}\beta_{2,i} + k_{2,n-4}2p^2e^{2t}x_2^{n-4}\frac{1}{2^{n-3}}\beta_{2,n-4}\right)\right]^{2}.$$

Для второй части функции получаем:

$$P''_{ifc}(2) = \sum_{g}\pi(g)(n-g)\left[\xi_1\left(k_{1,1}pe^{t}x_1^{1} + \sum_{i=2}^{n-4}k_{1,i}pe^{t}x_1^{i}\frac{1}{2^{i-1}}\beta_{1,i} + k_{1,n-2}pe^{t}x_1^{n-2}\frac{1}{2^{n-1}}\beta_{1,n-2}\right) + \right.$$

$$\left. + \xi_2\left(k_{2,1}p^2e^{2t}x_2^{1} + \sum_{i=2}^{n-6}k_{2,i}p^2e^{2t}x_2^{i}\frac{1}{2^{i+1}}\beta_{2,i} + k_{2,n-4}p^2e^{2t}x_2^{n-4}\frac{1}{2^{n-3}}\beta_{2,n-4}\right)\right]^{n-g-1} \times$$

$$\times\left[\xi_1\left(k_{1,1}pe^{t}x_1^{1} + \sum_{i=2}^{n-4}k_{1,i}pe^{t}x_1^{i}\frac{1}{2^{i+1}}\beta_{1,i} + k_{1,n-2}pe^{t}x_1^{n-2}\frac{1}{2^{n-1}}\beta_{1,n-2}\right) + \right.$$

$$\left. + \xi_2\left(k_{2,1}p^2 4e^{2t}x_2^{1} + \sum_{i=2}^{n-6}k_{2,i}p^2 4e^{2t}x_2^{i}\frac{1}{2^{i+1}}\beta_{2,i} + k_{2,n-4}p^2 4e^{2t}x_2^{n-4}\frac{1}{2^{n-3}}\beta_{2,n-4}\right)\right].$$

Упростим полученное равенство, полагая, что значения $k_{1,i} = k_{2,i} = 1$ и $\pi(g) = 1$, $n_1 >> 1$, но $n \geq n_1$. Тогда, при $t = 0$ и $r = 2$ имеем:

$$P_{ifc}^{n} = n(n-1)\left[\xi_1\left(p\frac{33}{8}\right)+\xi_2\left(p^2\frac{33}{8}\right)\right]^{n-2}\left[\xi_1\left(p\frac{33}{8}\right)+2\xi_2\left(p^2\frac{33}{8}\right)\right]^2 +$$

$$+n\left[\xi_1\left(p\frac{33}{8}\right)+\xi_2\left(p^2\frac{33}{8}\right)\right]^{n-1}\left[\xi_1\left(p\frac{33}{8}\right)+\xi_2\left(4p^2\frac{33}{8}\right)\right].$$

С учетом данного равенства и соотношения для математического ожидания, а также с учетом ряда допущений при округлении, имеем:

$$D_{ifc} \approx 1{,}085n^2\left(0{,}65561\right)^n \approx n^2\left(0{,}65561\right)^n. \quad (5)$$

Для минимальной вероятности p дисперсия распределения вероятностей ошибки при наблюдении ВПЗВ (5) принимает вид:

$$D_{ifc} = n^2\left(\frac{1}{2}pe^t\right)^n - n^2\left(\frac{1}{2}pe^t\right)^{2n}.$$

Очевидно, что при t=0 отсюда следует примерное равенство: $D_{ifc} < n^2\left(0{,}5p\right)^n$.

Учитывая существенное уменьшение дробной функции при возведении в степень, можем сделать вывод о стремлении к нулю дисперсии распределения вероятностей ошибки при наблюдении ВПЗВ.

Заключение

В представленной работе приведены результаты расчетов дисперсии энумератора распределения вероятностей ошибки при наблюдении и синтезе теоретических параметров для ВПЗВ. Дисперсия рассчитана для двух граничных случаев, а именно: для минимального значения интеграла вероятностей ошибки при максимальном значении вероятности p и для максимального значения указанного интеграла при минимальном p.

Показано, что дисперсия распределения для достаточно больших значений выборки в соответствии со второй производной для степенной функции образует параметр эквивалентный математическому ожиданию, но дополнительно умноженный на длину последовательности n.

Список использованных источников

1. Кобяк И.П. (2009) Теория внутрисхемного наблюдения СБИС с использованием автокорреляционных функций. Автоматика и вычислительная техника. (2) С.37-46.

2. Кобяк И.П. (2023) Производящая функция для вероятности пропуска ошибки при наблюдении векторов переходов. В кн.: BIG DATA и анализ высокого уровня 2023. сборник материалов 9-й междунар.науч.-практ. конф., часть 2, С. 16-23.

References

1. Kobyak I.P. (2009) Theory of in-circuit VLSI monitoring using autocorrelation functions. Automation and computer technology. (2), pp. 37-46.

2. Kobyak I.P. (2023) Generating function for the probability of missing an error when observing transition vectors. In: BIG DATA and High-level Analysis 2023. BIG DATA and advanced analytics 2023: proc. of the 9th Int. Scientific and Practical Conference. Part 2. pp. 16-23.

Сведения об авторе

Кобяк И.П., канд. техн. наук, доцент, доцент кафедры ЭВМ, учреждение образования «Белорусский государственный университет информатики и радиоэлектроники», IPKobyak2012@mail.ru.

Information about the author

Kobiak I.P., PhD, Associate Professor, Educational Institution "Belarusian State University of Informatics and Radioelectronics", Chair of Electronic Computing Machines.