УДК 519.624.2

О ТОЧНОМ РАВЕНСТВЕ ДЛЯ ВЕРОЯТНОСТИ ПРОПУСКА ОШИБКИ ПРИ НАБЛЮДЕНИИ ВЕКТОРОВ ПЕРЕХОДОВ

И.П. Кобяк
доцент кафедры ЭВМ
канд. техн. наук

Белорусский государственный университет информатики и радиоэлектроники, Республика Беларусь
E-mail: IPKobyak2012@mail.ru

И.П.Кобяк
Работает в Белорусском государственном университете с 1982 г. Занимаемые должности инженер, ассистент, доцент кафедры ЭВМ. Защитил кандидатскую диссертацию в 1993 г. Область научных интересов: методы идентификации сообщений, проектирование спецпроцессоров.

Аннотация. В статье рассмотрен метод идентификации случайных последовательностей оценками вероятности наблюдения векторов переходов. Определено точное равенство для вероятности пропуска ошибки, соответствующее данному алгоритму при регистрации заданных сложных событий в электронных системах наблюдения. Рассмотрен вопрос о моде распределения вероятностей данного параметра, что позволило выполнить сравнение уровней ошибки, порождаемых методом регистрации векторов переходов с известными алгоритмами контрольного кодирования, такими как сигнатурный анализ и статистика векторов состояний.

Ключевые слова: блок векторов, субдинамические объекты, идентификация последовательностей, вероятность пропуска ошибки, сигнатурный анализ.

Введение.

Моделирование процессов идентификации случайных сообщений в большинстве технических систем основывается на использовании эмпирических функций в качестве контрольных кодов, в определении их качества с точки зрения вероятности пропуска ошибки [1,2], в выборе оптимального алгоритма для конкретного приложения. Подобная задача решается при наблюдении сигналов в криптографических каналах связи, контроле и диагностике цифровых устройств, при пересылках информации между блоками вычислительных систем.

Методология идентификации последовательностей выборочной вероятностью наблюдения векторов переходов (ВП) заданного вида (или лебеговской мерой ВП) представляет собой известный интерес в силу квазиминимальности вероятности пропуска ошибки, определяемой интегральным сравнением рассматриваемого и других алгоритмов идентификации [3,4]. Однако детальное исследование полученных результатов показало, что соотношения в выше указанных источниках характеризуют только верхнюю границу для вероятностных показателей ошибки. Реальный уровень параметров в асимптотике значительно ниже и позволяет достаточно точно, используя принципы моделирования в теории разбиений, определить количественные показатели предлагаемого в данной работе метода идентификации. Таким образом, вопрос получения вероятностных параметров

выборки в задачах наблюдения векторов переходов является актуальным и рассматривается в представляемой статье.

Вероятность пропуска ошибки для сложных событий с произвольным числом - последовательных ВП.

При формировании формулы для числа различных перестановок с повторениями для меры Лебега сложных событий на n местах размещения, рассмотрим вероятность наблюдения ВП, в которых каждый субдинамический объект с μ единичными битами может быть учтен соотношением вида:

$$p = \frac{3^{r-\mu}}{m^2}. \tag{1}$$

При этом пространство между блоками ВП или число вариантов постобъектов размера $i \times r$ определяется перестановками всевозможных ВС (не содержащих ВП), число которых может быть определено классическим методом включения и исключения.

При формировании соотношения (1) предполагается, что источник последовательностей является идеальным, а в качестве опорного генератора для анализа используется имитатор псевдослучайных чисел, приведенный в [5].

Из представления сложных объектов случайного процесса в виде [4] следует, что каждая пара событий из суммы всех событий k_ω будет объединена с набором векторов функции правдоподобия или частных вариантов постобъектов длиной $i \geq 0$, а соотношение для вероятности пропуска ошибки (P_{ifc} - *interaction function correlation*) при произвольном $\mu \leq r$, может быть определено утверждением 1.

Утверждение 1. Произведение вида:

$$P_{ifc} = \sum_g \pi(g) \sum_{n-g} \frac{1}{m^n} \left(3^{r-\mu}\right)^{\sum_{i=1}^{n-2} k_{1,i}} m^{\sum_{i=1}^{n-2} i k_{1,i}} \prod_{i=2}^{n-2} \left(\frac{1}{2^{i+1}} \beta_i\right)^{k_{1,i}} \frac{(n-g)!}{k_{1,1}! k_{1,2}! \ldots k_{1,n-2}!},$$
$$\beta_i = \frac{1}{\sqrt{1-4p}} \left[\left(1+\sqrt{1-4p}\right)^{i+1} - \left(1-\sqrt{1-4p}\right)^{i+1} \right], \tag{2}$$

где $\pi(g)$-композиция разбиений числа $g = \sum_{i=1}^{n-2} (1+i) k_{1,i}$, является вероятностью пропуска ошибки для лебеговской меры векторов переходов со статистиками $k_{j,i}$ при $j = 1$ и всех $i = \overline{1, n-2}$, где n-четно.

Нечетность параметра n рассмотрена ниже в конце работы и не вносит существенных изменений в утверждение 1.

Доказательство. Анализ вариантов объектов, входящих в структурный состав некоторой многомерной последовательности (замкнутой от конца к началу), показывает, что монообъекты длиной $2j+i$ с параметрами $j = 1,\ i = 1$ могут быть представлены в функции (2) комбинаторными соотношениями вида $Q_{j,i} = Q_{1,1} = 3^{r-\mu} m$. Для значений $j = 1,\ i = 2$ приведенное соотношение трансформируется в равенство $Q_{1,2} = 3^{r-\mu}\left(m^2 - 3^{r-\mu}\right)$, что предполагает использование частного случая включения и исключения и преобразования данного произведения в соотношение:

$$Q_{1,2} = 3^{r-\mu} m^2 \sum_{s=0}^{1} C_{2-s}^{s} (-p)^s,$$

где C_{2-s}^{s} биномиальный коэффициент.

При $j=1,\ i=3$ соответственно имеем:

$$Q_{1,3}=3^{r-\mu}m^{3}\sum_{s=0}^{1}C_{3-s}^{s}\left(-p\right)^{s}\cdot$$

В общем случае, можно показать, что для $i=\overline{2,n-2}$ комбинаторный момент $Q_{1,i}$ имеет вид:

$$Q_{1,i}=3^{r-\mu}m^{i}\left[\sum_{s=0}^{z}C_{i-s}^{s}\left(-p\right)^{s}\right], \tag{3}$$

где z - это целая часть от $0{,}5i$.

Преобразуем сумму биномиальных коэффициентов в (3), с использованием результатов [6], в следующее равенство:

$$\sum_{s=0}^{z}C_{i-s}^{s}\left(-p\right)^{s}=\frac{1}{2^{i+1}}\beta_{i},$$

где параметр β_i определен выше в соотношении (2).

Учитывая теперь все перестановки с повторениями объектов $k_{1,i}\in k_{\omega},\ i>0$, можно записать соотношение для нормированной мощности классов эквивалентностей (МКЭ), соответствующей вероятности пропуска ошибки при наблюдении лебеговской меры ВП в форме суммы произведений:

$$P_{ifc}=\sum_{g}\pi\left(g\right)\sum_{n-g}\frac{1}{m^{n}}\left(3^{r-\mu}\right)^{k_{1,1}}m^{k_{1,1}}\prod_{i=2}^{n-2}\left[3^{r-\mu}m^{i}\left(\sum_{s=0}^{z}C_{i-s}^{s}\left(-p\right)^{s}\right)\right]^{k_{1,i}}\frac{\left(n-g\right)!}{k_{1,1}!k_{1,2}!...k_{1,n-2}!}, \tag{4}$$

где $\pi\left(g\right)$, как уж говорилось выше это композиция разбиений числа

$$g=2k_{1,1}+3k_{1,2}+...+\left(n-1\right)k_{1,n-2}$$

на целые части вида $\left(1+i\right)k_{1,i}$.

Необходимость получения параметра $\pi\left(g\right)$ в (4) следует из неоднозначности представления некоторого текущего константного значения g в виде функции $f\left(k_{1,1},k_{1,2},...,k_{1,n-2}\right)$.

Преобразуем произведение факториальных моментов (4), используя аппроксимацию β_i. При этом получим равенство:

$$P_{ifc}=\sum_{g}\pi\left(g\right)\sum_{n-g}\frac{1}{m^{n}}\left(3^{r-\mu}\right)^{k_{1,1}}m^{k_{1,1}}\prod_{i=2}^{n-2}\left[3^{r-\mu}m^{i}\left(\frac{1}{2^{i+1}}\beta_{i}\right)\right]^{k_{1,i}}\frac{\left(n-g\right)!}{k_{1,1}!k_{1,2}!...k_{1,n-2}!}.$$

Объединяя в данном соотношении степени при параметрах m и $3^{r-\mu}$, переходим к соотношению, представленному в утверждении 1.

Утверждение 2. Произведение вида:

$$P_{ifc}=\sum_{g}\pi(g)\sum_{n-g}\frac{1}{m^n}\left(3^{r-\mu}\right)^{\sum_{i=1}^{n-2}k_{1,i}+\sum_{i=1}^{n-4}2k_{2,i}}m^{\sum_{i=1}^{n-2}ik_{1,i}+\sum_{i=1}^{n-4}ik_{2,i}}\prod_{i=2}^{n-2}\left(\frac{1}{2^{i+1}}\beta_i\right)^{k_{1,i}}\prod_{i=2}^{n-4}\left(\frac{1}{2^{i+1}}\beta_i\right)^{k_{2,i}}\frac{(n-g)!}{k_{1,1}!k_{1,2}!...k_{1,n-2}!k_{2,1}!k_{2,2}!...k_{1,n-4}!},\quad(5)$$

n-четно, является вероятностью пропуска ошибки при наблюдении ВП с параметрами $j=1,2$ и $i=\overline{1,n-2j}$, $\pi(g)$ - это композиция разбиений числа

$$g=2k_{1,1}+3k_{1,2}+...+(n-1)k_{1,n-2}+4k_{2,1}+5k_{2,2}+...+(n-1)k_{2,n-4}$$

на целые части $(1+i)k_{1,i}+(3+i)k_{2,i}$.

Доказательство. Аналогично случаю с $j=1$ в утверждении 2 необходимо дополнительно рассмотреть монообъекты с параметром $j=2$. При этом соответствующие факториальные моменты могут быть представлены в функции (5) комбинаторными соотношениями: $Q_{2,1}=\left(3^{r-\mu}\right)^2 m$ для $i=1$, $Q_{2,2}=\left(3^{r-\mu}\right)^2\left(m^2-3^{r-\mu}\right)$ для значений $i=2$, а также

$$Q_{2,3}=\left(3^{r-\mu}\right)^2 m^3\sum_{s=0}^{1}C_{3-s}^{s}(-p)^s$$

при $i=3$. Очевидно, что общем случае для $i=\overline{2,n-2}$ комбинаторный момент $Q_{2,i}$ будет иметь вид:

$$Q_{2,i}=\left(3^{r-\mu}\right)^2 m^i\left[\sum_{s=0}^{z}C_{i-s}^{s}(-p)^s\right],\quad(6)$$

где z - также целая часть от значения $0{,}5i$.

Учитывая теперь, что функция правдоподобия для случая $j=2$ формируется совершенно аналогично случаю $j=1$, обобщая равенство (6) можем записать:

$$P_{ifc}=\sum_{g}\pi(g)\sum_{n-g}\frac{1}{m^n}\left(3^{r-\mu}\right)^{k_{1,1}}m^{k_{1,1}}\left(3^{r-\mu}\right)^{2k_{2,1}}m^{k_{2,1}}\prod_{i=2}^{n-2}\left[3^{r-\mu}m^i\left(\sum_{s=0}^{z}C_{i-s}^{s}(-p)^s\right)\right]^{k_{1,i}}\prod_{i=2}^{n-4}\left[\left(3^{r-\mu}\right)^2 m^i\left(\sum_{s=0}^{z}C_{i-s}^{s}(-p)^s\right)\right]^{k_{2,i}}\times$$
$$\times\frac{(n-g)!}{k_{1,1}!k_{1,2}!...k_{1,n-2}!k_{2,1}!k_{2,2}!...k_{1,n-4}!},\quad(7)$$

В данном равенстве композиция разбиений $\pi(g)$ формируется для представления параметра g в двумерном виде:

$$g=\sum_{j=1}^{2}\sum_{i=1}^{n-2j}(2j+i-1)k_{j,i}.$$

Далее, в соотношении (7) можем объединить степени параметров m и $3^{r-\mu}$. При этом получим:

$$\left(3^{r-\mu}\right)\deg\left(\sum_{i=1}^{n-2}k_{1,i}+\sum_{i=1}^{n-4}2k_{2,i}\right)\cdot m\deg\left(\sum_{i=1}^{n-2}ik_{1,i}+\sum_{i=1}^{n-4}ik_{2,i}\right)\cdot$$

Используя в очередной раз параметр β_i для аппроксимации функции правдоподобия, получаем результат, приведенный в (5).

Общий случай формирования вероятности ошибки P_{ifc} сформулируем в виде теоремы в форме суммы комбинаторных моментов.

Теорема. Произведение моментов

$$P_{ifc}=\sum_{g}\pi(g)\sum_{n-g}\frac{1}{m^n}\left(3^{r-\mu}\right)^{\sum\limits_{j=1}^{0,5n-1}\sum\limits_{i=1}^{n-2j}k_{j,i}+\frac{n}{2}k_{\frac{n}{2},0}}m^{\sum\limits_{j=1}^{0,5n-1}\sum\limits_{i=1}^{n-2j}ik_{j,i}}\prod_{j=1}^{0,5n-1}\prod_{i=2}^{n-2j}\left(\frac{1}{2^{i+1}}\beta_i\right)^{k_{j,i}}\frac{(n-g)!}{\prod\limits_{j=1}^{0,5n-1}\prod\limits_{i=1}^{n-2j}k_{j,i}!\ldots k_{\frac{n}{2},0}!},\quad(8)$$

где n-четно, является вероятностью пропуска ошибки при наблюдении лебеговской меры векторов переходов с параметрами $j=\overline{1,0,5n-1}$ и $i=\overline{1,n-2j}$, а также $j=0,5n,\ i=0$ по всей композиции $\pi(g)$ разбиений числа

$$g=\sum_{j=1}^{0,5n-1}\sum_{i=1}^{n-2j}(2j+i-1)k_{j,i}+(n-1)k_{\frac{n}{2},0}$$

на целые части вида $(2j+i-1)k_{j,i}+(n-1)k_{\frac{n}{2},0}$.

Доказательство. Анализ вариантов произвольных объектов числом $k_{j,i}\geq 1$, входящих в структурный состав последовательности достаточно большой длины, показывает, что монообъекты с параметрами $j\geq 1,\ i=1,$ могут быть представлены в функции (8) комбинаторными соотношениями вида $Q_{j,1}=\left(3^{r-\mu}\right)^j m$. Для значений $j\geq 1,\ i=2$ соответственно имеем равенство

$$Q_{j,2}=\left(3^{r-\mu}\right)^j\left(m^2-3^{r-\mu}\right),$$

а для $j\geq 1,\ i=3$ параметр

$$Q_{j,3}=\left(3^{r-\mu}\right)^j m^3\sum_{s=0}^{1}C_{3-s}^{s}(-p)^s.$$

В общем случае для произвольного $0,5n>j\geq 1$ можем записать:

$$Q_{j,i}=\left(3^{r-\mu}\right)^j m^i\left[\sum_{s=0}^{z}C_{i-s}^{s}(-p)^s\right],\quad z=]0,5i[.$$

Тогда произведение всех реализаций параметра $Q_{j,i}$ дает составляющие в (8) вида:

$$\deg\left(3^{r-\mu}\right)=\sum_{i=1}^{n-2}k_{1,i}+\sum_{i=1}^{n-4}2k_{2,i}+\sum_{i=1}^{n-6}3k_{3,i}+\ldots+\sum_{i=1}^{4}\frac{n-4}{2}k_{\frac{n-4}{2},i}+\sum_{i=1}^{2}\frac{n-2}{2}k_{\frac{n-2}{2},i}+\sum_{i=0}^{0}\frac{n}{2}k_{\frac{n}{2},i},$$

$$\deg m=\sum_{i=1}^{n-2}ik_{1,i}+\sum_{i=1}^{n-4}ik_{2,i}+\sum_{i=1}^{n-6}ik_{3,i}+\ldots+\sum_{i=1}^{4}ik_{\frac{n-4}{2},i}+\sum_{i=1}^{2}ik_{\frac{n-2}{2},i},$$

а множитель, характеризующий все реализации функции правдоподобия - это произведение произведений:

$$\prod_{i=2}^{n-2}\left(\frac{1}{2^{i+1}}\beta_i\right)^{k_{1,i}}\prod_{i=2}^{n-4}\left(\frac{1}{2^{i+1}}\beta_i\right)^{k_{2,i}}\prod_{i=2}^{n-6}\left(\frac{1}{2^{i+1}}\beta_i\right)^{k_{3,i}}\times\ldots\times\prod_{i=2}^{4}\left(\frac{1}{2^{i+1}}\beta_i\right)^{k_{\frac{n-4}{2},i}}\prod_{i=2}^{2}\left(\frac{1}{2^{i+1}}\beta_i\right)^{k_{\frac{n-2}{2},i}}.$$

Таким образом, объединяя все факториальные моменты при $k_{j,i}\geq 0$ по всей композиции $\pi(g)$ разбиений g на указанные в теореме целые части приходим к соотношению (8).

Исключение составляет случай с $j=0,5n$ при котором функция правдоподобия с учетом $i=0$ отсутствует. Данный факт учитывается показателем степени параметра $3^{r-\mu}$ вида $\frac{n}{2}k_{\frac{n}{2},0}$. Итак, теорема доказана.

Аналогичное соотношение при нечетном значении длины выборки n имеет несущественное отличие от вероятности (8). При этом:

$$P_{ifc}=\sum_{g}\pi(g)\sum_{n-g}\frac{1}{m^n}\left(3^{r-\mu}\right)^{\sum_{j=1}^{0,5(n-1)}\sum_{i=1}^{n-2j}k_{j,i}}m^{\sum_{j=1}^{0,5(n-1)}\sum_{i=1}^{n-2j}ik_{j,i}}\prod_{j=1}^{0,5(n-3)}\prod_{i=2}^{n-2j}\left(\frac{1}{2^{i+1}}\beta_i\right)^{k_{j,i}}\frac{(n-g)!}{\sum_{j=1}^{0,5(n-1)}\prod_{i=1}^{n-2j}k_{j,i}!}. \quad (9)$$

В данном равенстве отличиями являются пределы суммирования в степенях параметров m, $3^{r-\mu}$, g и в функции правдоподобия.

В случае, если n-четно, а $n+1$ - нечетно, вероятность (9) по отношению к параметру (8) при граничном значении j возрастает в m раз.

Сравнительный анализ методов идентификации выборки путем наблюдения ВП с другими известными алгоритмами свертки.

Анализ комбинаторной вариативности моментов, входящих в состав вероятностей (8) и (9), показывает, что МКЭ существенно зависит от числа ВП в составе многомерной выборки. При этом учитывая, что пара векторов, образующих ВП, имеет число вариантов $3^{r-\mu}$, а пара векторов состояний не образующих ВП - $m^2-3^{r-\mu}$, с учетом производной

$$\frac{\partial}{\partial r}3^{r-\mu}<\frac{\partial}{\partial r}\left(m^2-3^{r-\mu}\right), \quad (10)$$

можем заключить, что максимальное значение или мода для функции распределения вероятностей ошибки (8) или (9) (при $k_{i,j}=0$), имеет вид:

$$P_{ifc}=\frac{1}{m^n}\left(m^n\frac{1}{2^{n+1}}\beta_n\right)=\frac{1}{2^{n+1}}\frac{1}{\sqrt{1-4p}}\left[\left(1+\sqrt{1-4p}\right)^{n+1}-\left(1-\sqrt{1-4p}\right)^{n+1}\right]. \quad (11)$$

Для максимального значения вероятности (1) равного $p=\frac{3}{16}$, из (11) имеем и максимальное значение моды:

$$P_{ifc}=\frac{3}{2}\left(\frac{3}{4}\right)^n.$$

Максимальное значение аналогичного параметра при построении графика

вероятности пропуска ошибки многоразрядным сигнатурным анализатором равно его классической вероятности ошибки $P_{msa} = \frac{1}{rn}$. Таким образом, сравнивая оба метода в точках максимума, имеем соотношение:

$$\frac{P_{ifc}}{P_{msa}} = \frac{3}{2}\left(\frac{3}{4}\right)^n rn,$$

что говорит об абсолютном превосходстве метода наблюдения лебеговской меры ВП над алгоритмом синтеза сигнатур.

При сравнении методов наблюдения ВП и ВС (с вероятностью P_{cvc}, *cvc – condition vector counting*) также в точке моды [7] можем записать соотношение:

$$\frac{P_{ifc}}{P_{cvc}} = \frac{3}{2}\left(\frac{3}{4}\right)^n \sqrt{0,5\pi n},$$

что также указывает на превосходство метода наблюдения ВП.

Таким образом, для вероятностей пропуска ошибки в точке моды приведенными тремя методами получаем ряд неравенств: P_{ifc} $P_{msa} < P_{cvc}$. Иными словами, в точке максимума представляемый в данных исследованиях метод является наиболее предпочтительным.

Выводы.

В представленной работе получены следующие результаты:

1) сформировано соотношение для вероятности пропуска ошибки при наблюдении заданных векторов переходов с произвольным параметром j;

2) исследования позволяют сделать вывод, что увеличение выборочного числа событий $\hat{k}_{\omega}$ наблюдения лебеговской меры ВП приводит к уменьшению значения вероятности пропуска ошибки за счет уменьшения числа перестановок в ВС на местах расположения регистрируемых ВП (10);

3) сравнительный анализ метода наблюдения заданных ВП с алгоритмами линейной свертки и СВС показал, что в точке моды принцип идентификации случайных процессов вероятностью наблюдения ВП имеет явное преимущество перед известными алгоритмами синтеза точечных оценок; этот же вывод можно сделать и относительно асимптотики рассматриваемых алгоритмов [8];

4) используя алгоритм перекодирования состояний в выборке при $r = \mu$, можно заключить, что полученные результаты будут справедливы и для любых пар различных соседних упорядоченных векторов состояний априори заданного вида.

Список использованных источников

[1] *Кобяк И.П.* Сравнительная оценка достоверности методов сигнатурного анализа и счета состояний // Электрон. Моделирование. – 1996. – Т.18. – №1. – С. 58-62.

[2] *Гордон Г., Надиг Х.* Локализация неисправностей в микропроцессорных системах при помощи шестнадцатиричных ключевых кодов // Электроника. – 1977. – 5 №5. – С.89–96.

[3] *Кобяк И.П.* Производящая функция для распределения вероятностей наблюдения векторов переходов // Автоматика и вычислительная техника. – 2006. – № 6. – С. 60–67.

[4] *Кобяк И.П.* Теория внутрисхемного наблюдения СБИС с использованием автокорреляционных функций // Автоматика и вычислительная техника. – 2009. – № 2. – С. 37-46.

[5] *Tootill J.P.R., Robinson W.D., Eagle D.J.* An Asymptotically Random Tausworthe Sequences // Journal of the Association Computing Machinery. July 1973. – Vol. 30. –No. 3. – P. 469–481.

[6] *Риордан Дж.* Комбинаторные тождества. М.: Наука. – 1982. – 255 с.

[7] *Ширяев А.Н.* Вероятность. –М.: Наука. Глав. ред. физ.-мат. литературы, 1980.–575 с.

[8] *Кобяк И.П.* Точное значение вероятности пропуска ошибки при наблюдении векторов переходов в асимптотике. В кн.: Информационные технологии и системы 2021 (ИТС 2021), Information Technologies and Systems 2021 (ITS 2021): материалы международной научной конференции, Минск, 24 ноября 2021 г. / Белорусский государственный университет информатики и радиоэлектроники, Минск, 2021. – С. 200–201.

ABOUT THE EXACT EQUALITY FOR THE PROBABILITY OF MISSING AN ERROR WHEN OBSERVING TRANSITION VECTORS

I.P. KABIAK,
PhD, Associate Professor, Chair of ECM, BSUIR.

Belarusian State University of Informatics and Radioelectronics, Republic of Belarus
E-mail: IPKobyak2012@mail.ru

Abstract. The article considers a method for identifying random sequences by estimating the probability of observing transition vectors. The exact equality for the probability of missing an error corresponding to this algorithm when registering specified complex events in electronic surveillance systems is determined. The question of the mode of probability distribution of this parameter is considered, which made it possible to compare the error levels generated by the method of registering transition vectors with known control coding algorithms, such as signature analysis and statistics of state vectors.

Keywords: block of vectors, subdynamic objects, sequence identification, probability of missing an error, signature analysis.