2. Критерий оказывается нечувствительным к очередности распределения нулей и единиц внутри группы, а также расположения начала группы относительно условного начала ДВП.

3. По мере увеличения ширины группы при T=const значения χ^2 увеличиваются. Отклонения гистограмм также увеличиваются, но их асимметричный характер не нарушается.

4. По мере увеличения интервала T при неизменной ширине группы значения χ^2 почти не изменяются.

Таким образом, методика выявления в ДВП детерминированной составляющей должна содержать следующие операции.

Операция 1. Составление из исходной ДВП (1) в соответствии с (4) массива значений случайной величины *B*.

Операция 2. Расчет осредненной гистограммы случайной величины *B,* определение направления ее асимметрии относительно теоретической плотности распределения (5).

Операция 3. Установление по полученной асимметрии типа детерминированной составляющей: если асимметрия характеризуется положительными приращениями в сторону уменьшения значений случайной величины, то детерминированная составляющая состоит из единиц, в противном случае – из нулей.

Операция 4. Вычисление критерия χ^2 для осредненной гистограммы и сравнение его с табличным значением для заданного объема выборки и числа степеней свободы. Отклонение полученного значения от теоретического более чем в 2 раза, особенно при наличии определенной на операции 2 асимметрии, является необходимым условием наличия в исходной ДВП детерминированной составляющей.

III Выводы

1. Предложена методика выявления в ДВП детерминированных составляющих, основанная на известном методе биномиального преобразования.

2. Методика дает необходимые доказательства наличия в ДВП детерминированных составляющих и может быть применена как составная часть системы информационной безопасности.

Литература: 1.Зегжда Д. П., Ивашко А. М. Основы безопасности информационных систем. – М.: Горячая линия – Телеком, 2000. – 452 с., ил. 2. Д. Кнут. Искусство программирования для ЭВМ.– Т. 2. – Получисленные алгоритмы. – М.: Мир. – 1977. – 482 с. 3. P. L'Ecuyer. Uniform random number generation.// Annals of Operations Research. – 1994. – V. 53. – pp. 77-120. 4. H. Niederreiter. Pseudo-random numbers and optimal coefficients. // Advances in Mathematics. – 1977. – V. 26. – pp. 99-181. 5. S. K. Park and K. W. Miller. Random number generators: good ones are hard to find. // Communs of the ACM. – 1988. – V. 31. – pp. 1192-1201. 6. N. S. Altman. Bit-wise behavior of random number generators.// SIAM Journal of Sci. Stat. Computing. – 1988. – V9(5). – pp. 941-949. 7. J. Eichenauer-Herrmann. Inversive congruential pseudorandom numbers: a tutorial. // International Statistical Review. – 1992. – V. 60. – pp. 167-176. 8. J. Walker. HotBits: Genuine random numbers, generated by radioactive decay. // http://www.fourmilab.ch/hotbits/ 9. Т. Левченко. Тестирование двоичных вероятностных последовательностей методом биномиального преобразования. – В сб.: Правове, нормативне та метрологічне забезпечення системи захисту інформації в Україні. – Науково-технічний збірник. – Випуск 2. – К.:НДЦ "Тезіс" НТУУ "КПІ". – 2001. – 273с.– С. 152. 10. Справочник по математике (для научных работников и инженеров). – Г. Корн, Т. Корн. – Изд. 4.– М.: Наука, 1978.– 831 с.

УДК 681.3.067:681.3.016

ВЫЯВЛЕНИЕ В ДВОИЧНЫХ ВЕРОЯТНОСТНЫХ ПОСЛЕДОВАТЕЛЬНОСТЯХ ДЕТЕРМИНИРОВАННЫХ СОСТАВЛЯЮЩИХ С ИСПОЛЬЗОВАНИЕМ ЛИНЕЙНОГО БУЛЕВА ПРОГРАММИРОВАНИЯ

Тарас Левченко

Научно-технический комплекс "Импульс", г. Киев

Аннотация: **Двоичная вероятностная последовательность (ДВП) разбивается на фрагменты длиной *m* бит, рассматриваемых как совокупность из *n* образцов, обладающих некоторым числом признаков из заданного множества *m* признаков. Показано, что периодическую компоненту можно выявить в результате решения задачи покрытия методом линейного булева программирования. Приведен алгоритм выявления.**

Summary: **Binary probabilistic sequence (BPS) is fragmented in *n* patterns by *m* bits length, each one**

considered as number of tags from given set of *m* tags. The conclusion based on the solution of a cover problem is given about a feasibility of linear Boolean programming for detection of the determined components. There is given algorithm used received results.
Ключевые слова: **Информационная безопасность, двоичные вероятностные последовательности, детерминированная составляющая, задача покрытия, линейное булево программирование.**

I Постановка задачи

Широкое применение двоичных вероятностных последовательностей (ДВП)

$$\{b_N b_{N-1} \mathrm{K}\, b_1 b_0\}, \tag{1}$$

где

$$b_i = \begin{cases} 1, \ p_1 = 0.5, \\ 0, \ p_0 = 0.5, \end{cases} \tag{2}$$

– некоррелированный бит в позиции номер i с дискретной плотностью распределения p, для защиты информационных ресурсов в интегрированных информационных системах [1] делает актуальной проблему выявления в ДВП детерминированных составляющих [2]. Критерием использования генератора ДВП является отсутствие детерминированной составляющей и неповторяемость фрагментов определенной длины при достаточно большом N.

ДВП, разбитую на фрагменты длиной m бит, можно рассматривать как совокупность из n образцов, каждый из которых обладает некоторым числом признаков из заданного множества m признаков, а вместе эти n образцов обладают всеми m признаками.

Цель работы – определение минимального числа объектов, которые в совокупности обладали бы всеми m признаками влияния на него искусственно внесенной в исходную ДВП детерминированной составляющей и выработка рекомендаций по выявлению детерминированной составляющей в ДВП.

I Основная часть

Предложенный подход позволяет свести задачу выявления в ДВП детерминированной составляющей к известной комбинаторной экстремальной задаче о покрытии [3]. В этом случае составленная из исходной последовательности (1) при условии (2) матрица

$$\left(\left(b_{ij}\right)\right) = \begin{pmatrix} b_{11} & b_{12} & \Lambda & b_{1m} \\ b_{21} & b_{22} & \Lambda & b_{2m} \\ \Lambda & \Lambda & \Lambda & \Lambda \\ b_{n1} & b_{n2} & \Lambda & b_{nm} \end{pmatrix}, \tag{3}$$

где nxm=N, является матрицей инцидентности, в которой j-й образец обладает i-м признаком при b_{ij} =1 и не обладает при b_{ij} =0.

Введем булевы переменные x_1, ..., x_n и будем полагать, что x_j = 1, если j-й образец выбран нами в искомую совокупность, и x_j = 0 – в противном случае. Тогда задача о покрытии может быть представлена в виде одновременного выполнения для всех i=1, 2, …, m условий

$$\sum_{j=1}^{n} x_j \to \min, \tag{4a}$$

$$\sum_{j=1}^{n} b_{ij} x_j \geq 1. \tag{4b}$$

Отметим, что условие булевости коэффициентов (3) позволяет свести в (4b) умножение к сложению.

Для численного анализа (4) была исследована плотность вероятности $p(\xi)$ псевдослучайной величины

$$\xi = \sum_{j=1}^{n} x_j \tag{5}$$

при наложении на каждую из ее реализаций, обусловленных исходной ДВП случайных ограничений (4b).

В качестве исходных ДВП были выбраны несколько двоичных реализаций выходной последовательности линейного конгруэнтного генератора Borland C++ с параметрами разбивки {N=16, n=4}, {N=36, n=6}, {N=64, n=8}. Детерминированными составляющими служили: повторяющаяся с интервалом L группа шириной в l

единиц или нулей, а также повторяющаяся с интервалом L группа из последовательности $l/2$ единиц и $l/2$ нулей в разных сочетаниях. Булевы переменные x_j изменялись по алгоритму полного перебора.

На рис. 1 – 3 приведены результаты численных расчетов $p(\xi)$ для перечисленных параметров разбивки при отсутствии (гистограмма 1) и наличии (гистограмма 2) детерминированной составляющей, которые для улучшения сходимости результатов были осреднены по 12 реализациям.

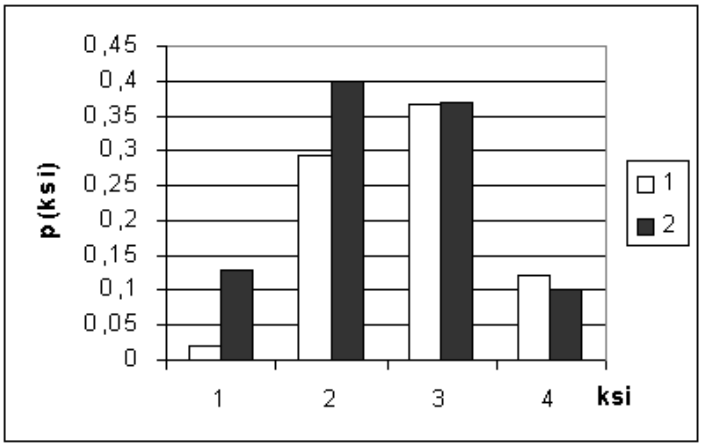

Рисунок 1 – Плотность распределения псевдослучайной величины ξ с параметрами разбивки ДВП {N=16, n=4}

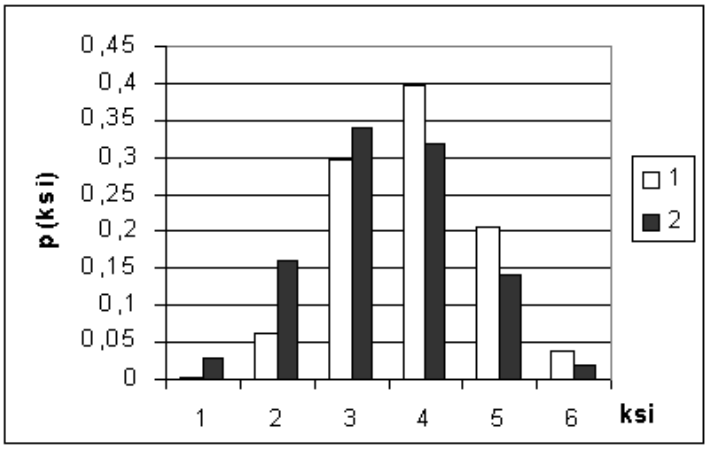

Рисунок 2 – Плотность распределения псевдослучайной величины ξ с параметрами разбивки ДВП {N=36, n=6}

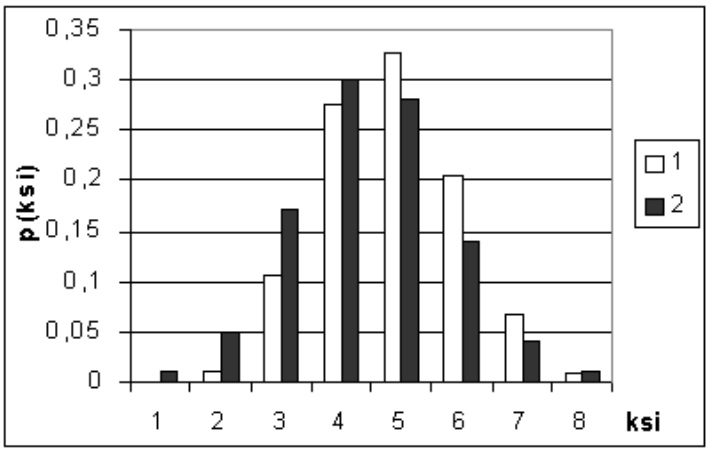

Рисунок 3 – Плотность распределения псевдослучайной величины ξ с параметрами разбивки ДВП {N=64, n=8}

Можно отметить следующие результаты.

1. $p(\xi)$ имеет выраженную асимметрию, обусловленную влиянием ограничений (4b), следствием которой малые значения ξ оказываются менее вероятными, чем большие. Указанное обстоятельство является причиной того, что в исходных ДВП количество булевых переменных, найденных при помощи алгоритма полного перебора и удовлетворяющих условиям (4), составляет от 1–2 для n=4 до 2–3 для других n. Другими словами, при достаточно большой длине образца в *любой* из исследованных ДВП всегда существует 2–3 фрагмента длиной m, в совокупности которых *все* биты принимают значение 1.

2. При наложении на исходные ДВП ненулевых детерминированных составляющих при $L/n=k$, где k целое, отмечалась четкая тенденция к уменьшению асимметрии $p(\xi)$, что свидетельствует об уменьшении количества указанных булевых переменных до 1. Тем самым в преобразованных ДВП всегда возникал фрагмент длиной m, состоящий только из единиц.

3. Во всех других случаях детерминированная составляющая на $p(\xi)$ и указанные булевы переменные не влияла.

4. Трудоемкость использованного алгоритма полного перебора пропорциональна 2^m.

Ввиду значительной трудоемкости алгоритма полного перебора при увеличении n для определения x_1, ..., x_n целесообразно использовать приближенные алгоритмы, например, так называемый жадный алгоритм. На каждом шаге алгоритма одна из переменных x_j полагается равной единице и выбирается так, чтобы максимизировать число выполненных к данному шагу неравенств. После этого все выполненные неравенства вычеркиваются, число переменных сокращается на единицу и повторяется очередной шаг. Алгоритм заканчивает работу, если все неравенства выполнены. Очевидно, что трудоемкость жадного алгоритма по сравнению с алгоритмами полного перебора невелика. Имеется гарантированная верхняя оценка максимального отклонения получаемого с его помощью решения от оптимума и доказана его асимптотическая оптимальность для широкого класса матриц. Отметим также, что на результаты расчетов выбор алгоритма нахождения минимального покрытия не влияет.

Основываясь на полученных выводах, можно рекомендовать следующий алгоритм нахождения в ДВП детерминированной составляющей.

Шаг 1. Из исходной ДВП сформировать матрицу (3), имеющую m столбцов, при этом начальное значение m должно быть заведомо меньше ожидаемого периода детерминированной составляющей.

Шаг 2. При помощи жадного алгоритма найти минимальное число образцов, удовлетворяющих условиям (4).

Шаг 3. Увеличить m на единицу.

Шаг 4. При помощи жадного алгоритма снова найти минимальное число образцов, удовлетворяющих условиям (4).

Шаг 5. Если полученное число меньше определенного на шаге 2, можно сделать вывод, что в исходной ДВП присутствует детерминированная составляющая, в которой имеются единицы, следующие с периодом $L=m$, и завершить работу. В противном случае перейти на шаг 3. Выход из цикла происходит также после достижения m заданного предельного значения.

Алгоритм применим и для ДВП, детерминированная составляющая в которых состоит из нулей. В этом случае исходную ДВП следует инвертировать.

В отличие от других методов выявления детерминированной составляющей, предложенный алгоритм позволяет определить *точное* значение периода.

Выводы

3. Предложен алгоритм выявления в ДВП детерминированных составляющих, основанный на решении известной задачи о покрытии при помощи линейного булева программирования.

4. Алгоритм дает необходимые и достаточные доказательства наличия в ДВП детерминированной составляющей и может быть применен как составная часть системы информационной безопасности.

Литература: 1. Зегжда Д. П., Ивашко А. М. Основы безопасности информационных систем – М.: Горячая линия – Телеком. – 2000. – 452 с. 2. Т. Левченко. Тестирование двоичных вероятностных последовательностей методом биномиального преобразования. – В сб.: Правове, нормативне та метрологічне забезпечення системи захисту інформації в Україні. – Науково-технічний збірник.– Випуск 2. – К.: НДЦ "Тезіс" НТУУ "КПІ". – 2001. – 273 с. – С. 152. 3. Пападимитриу Х., Стайглиц К. Комбинаторная оптимизация.– М.: Мир. – 1985. – 512 с.