УДК 004.7.056.5 (0.45)

[1]**Виноградов Н. А.** д-р техн. наук
[2]**Коробко В. В.** канд. техн. наук
[2]**Задоя Г. М.**
[3]**Скоропаденко А. П.**
[3]**Вовк В. М.**

МЕТОДИКА ОЦЕНИВАНИЯ ПАРАМЕТРОВ ПОТОКОВ ЭКСТРЕМАЛЬНЫХ СИТУАЦИЙ ТЕЛЕКОММУНИКАЦИОННОЙ СЕТИ ДЛЯ ЗАДАЧ ОПТИМАЛЬНОГО РАСПРЕДЕЛЕНИЯ РЕСУРСОВ ЗАЩИТЫ

[1]Институт информатики Национального авиационного университета,
[2]ОАО «Укртелеком»,
[3]ОАО «Карат-комплекс»

Изучена проблема защиты телекоммуникационных сетей от угроз. Разработана методика канонического разложения функции распределения вероятности потока угроз. Результаты применяют для вычисления временных интервалов постоянного потока угроз для соответствующего назначения средств защиты.

Введение

Для эффективного функционирования телекоммуникационных сетей любого масштаба (локальных, региональных, глобальных), помимо применения высоконадежного оборудования, необходимо обеспечить защиту от угроз самой различной природы, как техногенных и природных, так и из-за человеческого фактора. В работе [1] рассмотрены именно угрозы последнего вида применительно к неохраняемым объектам связи. Решается задача интеграции системы сбора информации о состоянии телекоммуникационной сети в существующие фрагменты универсальной первичной сети и формируемые на ее основе вторичные (региональные) сети. В рамках теории массового обслуживания рассмотрен поток попыток несанкционированного доступа (НСД) на объект и предложена методика оптимального распределения ресурсов защиты объекта в зависимости от параметров входящего потока. В методике используется графический метод: строятся гистограммы функций распределения количества НСД по времени суток и определяются временные интервалы с примерно одинаковой интенсивностью попыток НСД.

Однако машинная реализация предложенной методики является весьма трудоемкой. Кроме того, трудно оценить зависимость ее эффективности от длины выборки.

В данной работе предложена методика рекуррентного оценивания функции распределения на основе канонического разложения В. С. Пугачева и вычисления интервалов методом Н. П. Бусленко.

Постановка задачи

Рассмотрим последовательность попыток НСД как случайную функцию $X(t)$, $t=t_i$, $i=1,2,\dots$, представленную своими дискретными отсчетами t_i. Объем выборки исходных данных (число отсчетов) растет в процессе функционирования системы сбора информации об экстремальных ситуациях, возникающих на участках телекоммуникационной сети. Поэтому необходимо применять некие рекуррентные алгоритмы текущего оценивания статистических характеристик выборки переменного объема и алгоритмы принятия решений по критериям последовательного анализа [2].

Рассмотрим сначала принципиальные возможности описания случайной функции, непрерывной или дискретной, аналитической моделью, достаточно простой и удобной для практической реализации на ЭВМ, и в то же время адекватно описывающей реальные физические процессы. Поставленная задача рассматривалась многими отечественными и зарубежными специалистами по математической статистике. При всем разнообразии методик и применяемого математического аппарата общий подход сводится к выделению в исследуемой случайной функции двух составляющих: зависящей от аргумента и не зависящей от аргумента. Эти составляющие в общем случае являются случайными, хотя иногда могут быть детерминированными.

Методика оценки параметров случайного потока

Один из наиболее универсальных методов построения аналитической модели случайной функции – канонические разложения В. С. Пугачева [3]. Рассмотрим сначала методику канонического разложения для выборки фиксированного объема.

Представим случайную функцию в виде линейной комбинации

$$X(t)=m_x(t)+\sum_{\nu=1}^{\infty}V_\nu\varphi_\nu(t), \qquad (1)$$

где $m_x(t)$ – математическое ожидание функции $X(t)$;

V_ν – произвольные некоррелированные между собой коэффициенты канонического разложения – случайные величины с нулевым математическим ожиданием и конечной дисперсией:

$$M[V_\nu]=0, \quad M[V_\nu^2]=\sigma_\nu^2, \quad M[V_\nu,V_\mu]=0, \quad \nu\neq\mu, \quad \nu=1,2,\ldots, \qquad (2)$$

$\varphi_\nu(t)$, $t=t_i, i=1,2,\ldots$, $\nu=1,2,\ldots$ – некоторые детерминированные функции (координатные функции), подлежащие определению.

Слагаемые $V_\nu\varphi_\nu(t)$ называют элементарными случайными функциями.

Если функция $X(t)$ задана своим каноническим разложением (1), то, в соответствии с определением корреляционной функции [4] и с учетом некоррелированности коэффициентов V_ν можно представить ее в следующем виде:

$$R_x(t,\tau)=\sum_{\nu=1}^{\infty}\sigma_\nu^2\varphi_\nu(t)\varphi_\nu(\tau). \qquad (3)$$

Такое представление называется каноническим разложением корреляционной функции. При $t=\tau$ получаем очевидное соотношение

$$R_x(t,t)=\sigma_x^2(t)=\sum_{\nu=1}^{\infty}\sigma_\nu^2\varphi_\nu^2(t). \qquad (4)$$

Для практической реализации данного метода необходимо выразить коэффициенты V_ν и функции $\varphi_\nu(t)$ через известные величины. В нашем случае это значения случайной функции $X(t)$ и ее статистические характеристики. По накопленным значениям можно определить математическое ожидание $m_x(t)$, после чего центрировать $X(t)$:

$$\overset{0}{X}(t)=X(t)-m_x(t)=\sum_{\nu=1}^{\infty}V_\nu\varphi_\nu(t). \qquad (5)$$

Домножим обе части (5) на V_μ и вычислим математическое ожидание:

$$M[V_\mu X_0(t)]=\sum_{\nu=1}^{\infty}M[V_\mu V_\nu]\varphi_\nu(t). \qquad (6)$$

Вследствие некоррелированности коэффициентов канонического разложения (см. (2)) в правой части (6) отлично от нуля лишь слагаемое с индексом $\nu=\mu$, поэтому выражение (6) после очевидных преобразований сводится к

$$\varphi_\mu(t)=\frac{1}{\sigma_\mu^2}M\left[V_\mu\overset{0}{X}(t)\right]. \qquad (7)$$

Для получения конечного результата суммирования ограничим число членов разложения (1), отбросив члены с индексами, большими некоторого номера $N<\infty$:

$$X_N(t)=m_x(t)+\sum_{\nu=1}^{N}V_\nu\varphi_\nu(t). \qquad (8)$$

Ясно, что при этом имеют место ошибки. В [3] показано, что при выборе координатных функций в соответствии с правилом (7) среднеквадратическая ошибка разложения (1) будет минимальной для любого N :

$$\sigma_N^2 = M\left[\left|\overset{0}{X}(t) - \sum_{\nu=1}^{N} V_\nu \varphi_\nu(t)\right|^2\right] = \\ = \sigma_x^2(t) - \sum_{\nu=1}^{N} \sigma_\nu^2 \varphi_\nu^2(t) = \min \quad (9)$$

Следующим шагом в разработке практической методики разложения в дискретный ряд (8) является подбор коэффициентов V_ν. В соответствии с (7) эти коэффициенты коррелированы с исходной функцией $X(t)$. Если учесть, что V_ν имеют нулевые математические ожидания, то простейшими случайными функциями, коррелированными с $X(t)$, являются линейные комбинации значений центрированной случайной функции $\overset{0}{X}(t)$, соответствующих различным значениям аргумента t. Задав произвольную последовательность дискретных отсчетов t_i, определим V_ν в виде ряда

$$V_\nu = \sum_i a_{\nu i} \overset{0}{X}(t_i) \quad (10)$$

с произвольными коэффициентами $a_{\nu i}$. Подставив выражение (10) в условие некоррелированности (2) коэффициентов V_ν, получим, что

$$M[V_\nu, V_\mu] = \sum_i \sum_j a_{\nu i} a_{\mu j} M\left[\overset{0}{X}(t_i)\overset{0}{X}(t_j)\right] = \\ = \sum_i \sum_j a_{\nu i} a_{\mu j} R_x(t_i, t_j) = 0 \quad (11)$$

Число уравнений (11) всегда меньше числа коэффициентов $a_{\nu i}$ (недоопределенная система уравнений), поэтому множество способов выбора коэффициентов $a_{\nu i}$ теоретически бесконечно. Такой произвол в подборе решения весьма удобен с практической точки зрения: можно минимизировать вычислительные затраты, например, благодаря нормировке по коэффициенту с произвольным номером. Наиболее логичным выбором является нормировка по первому коэффициенту

$$\begin{cases} a_{11} = 1 \\ a_{1\nu} = 0, \ \nu \neq 1 \end{cases} \Rightarrow \quad V_1 = \overset{0}{X}(t_1). \quad (12)$$

Тогда

$$\sigma_1^2 = \sigma_x^2(t_1), \\ \varphi_1(t) = \frac{1}{\sigma_1^2} M\left[V_1 \overset{0}{X}(t)\right] = \\ = \frac{1}{\sigma_1^2} M\left[\overset{0}{X}(t)\overset{0}{X}(t)\right] = \\ = \frac{1}{\sigma_1^2} R_x(t_1, t), \ t \geq t_1 \quad (13)$$

Очевидно, что $\begin{cases} \varphi_1(t_1) = 1, \ t = t_1, \\ \varphi_1(t) = 0, \ t < t_1. \end{cases}$

Выражение (10) теперь можно записать в виде

$$V_\nu = \sum_i a_{\nu i} \overset{0}{X}(t_i) - V_1 \varphi_1(t_i), \ \nu = 2, 3, \ldots,$$

где $\varphi_1(t_i) = a_{1i}$, т. к. $V_1 = \overset{0}{X}(t_1)$.

По аналогии с (12)

$$\begin{cases} a_{\nu 1} = 1, \nu = 2, \\ a_{\nu i} = 0, \ i > \nu, \end{cases} \Rightarrow V_2 = \overset{0}{X}(t_2) - V_1 \varphi_1(t_2). \quad (14)$$

Используя выражения (12, 14), легко показать, что коэффициент V_2 некоррелирован с V_1, имеет нулевое математическое ожидание и дисперсию $\sigma_2^2 = \sigma_x^2(t_2) - \sigma_1^2 \varphi_1^2(t_2)$. Параметры второй координатной функции определяются аналогично первой (см. (13)):

$$\varphi_2(t) = \frac{1}{\sigma_2^2} M\left[V_2 \overset{0}{X}(t)\right] = \\ = \frac{1}{\sigma_2^2}\left\{R_x(t_2, t) - \sigma_1^2 \varphi_1(t)\right\}, \ t \geq t_2, \quad (15)$$

$$\varphi_2(t_2) = 1, \quad \varphi_2(t) = 0, t < t_2.$$

С учетом (14-15) можно записать выражение для представления случайной функции двумя членами разложения в ряд:

$$X_0(t) = V_1 \varphi_1(t) + V_2 \varphi_2(t).$$

В точках t_1, t_2 ошибка разложения равна нулю, а при $t \neq t_1, t \neq t_2$ среднеквадратическая ошибка минимальна.

Продолжив по индукции вывод уравнений вида (12-15), получим окончательный рекуррентный алгоритм вычисления параметров канонического разложения в следующем виде:

$$V_1 = \overset{0}{X}(t_1);$$
$$V_k = X(t_k) - \sum_{\nu=1}^{k-1} V_\nu \varphi_\nu(t_k), k=2,3,\ldots; \quad (16)$$

$$\sigma_1^2 = \sigma_x^2(t_1);$$
$$\sigma_k^2 = \sigma_x^2(t_k) - \sum_{\nu=1}^{k-1} \sigma_\nu^2 \varphi_\nu^2(t_k), k=2,3,\ldots; \quad (17)$$

$$\varphi_k(t) = \frac{1}{\sigma_k^2}\left[R_x(t_k,t) - \sum_{\nu=1}^{k-1}\sigma_k^2\varphi_\nu(t_k)\varphi_\nu(t)\right], \quad (18)$$
$k=1,2,\ldots;$

$$X(t) = m_x(t) + \sum_{\nu=1}^{k} V_\nu \varphi_\nu(t), \quad (19)$$
$t \ge t_k$, $k=\overline{1,N}$

Выразим дисперсию и корреляционную функцию $X(t)$ также в виде разложения в ряд по координатным функциям:

$$\sigma_x^2(t) = \sum_{\nu=1}^{k} \sigma_\nu^2 \varphi_\nu^2(t), t \ge t_k, \; k=\overline{1,N}; \quad (20)$$

$$R_x(t,\tau) = \sum_{\nu=1}^{k} \sigma_\nu^2 \varphi_\nu(t)\varphi_\nu(\tau), k=\overline{1,N}. \quad (21)$$

Используя (21), выразим i -ю координатную функцию $\varphi_k(t)$ через соответствующие специальным образом нормированные приращения корреляционной функции на $k-$ м шаге дискретизации:

$$\varphi_k(t) = \frac{R_x(t_k,t) - R_x(t_{k-1},t)}{R_x(t_k,t_k) - R_x(t_{k-1},t_{k-1})}, t \ge t_k. \quad (22)$$

Вычисляя коэффициенты V_k по формулам (16) и координатные функции $\varphi_k(t)$ по формулам (18,21), получаем точное описание функции $X(t)$ в точках $t = t_k$. По существу, коэффициентами V_k описывается мера новой информации о функции $\overset{0}{X}(t)$, которая содержится в ее $k-$м сечении $\overset{0}{X}(t_k)$, а координатными функциями $\varphi_k(t)$ – линейная связь новой информации, полученном на шаге k, с ожидаемыми значениями процесса $X_0(t)$ в моменты $t \ge t_k$. В целом $k-$ й элементарной случайной функцией $V_k\varphi_k(t)$, $t \ge t_k$, определяется вклад, вносимый новой информацией, которая получена на шаге k, в оценку будущих значений $\overset{0}{X}(t)$. Ценность этой информации определяется степенью коррелированности процесса $\overset{0}{X}(t)$. Для всех $t \ne t_k$, $k=\overline{1,N}$ функция $X(t)$ представляется разложением в ряд (19) приближенно, причем с максимально достижимой точностью (минимальной среднеквадратической ошибкой σ_N^2).

В работе анализируются случайные последовательности событий, наблюдаемых в дискретные моменты времени. На ограниченном интервале $t_1 \le t \le t_M$, $M \subseteq N$, можно записать выражение (19) для канонического разложения случайной последовательности в следующем виде:

$$X(k) = m_x(k) + \sum_{\nu=1}^{k} V_\nu \varphi_\nu(k), k=\overline{1,M}, \quad (23)$$

а ее дисперсии и корреляционной функции (20-21) – в виде:

$$\sigma_x^2(k) = \sum_{\nu=1}^{k} \sigma_\nu^2 \varphi_\nu^2(k), \quad k=\overline{1,M}; \quad (24)$$

$$R_x(k,l) = \sum_{\nu=1}^{k} \sigma_\nu^2 \varphi_\nu(k)\varphi_\nu(l), \quad (25)$$
$k=\overline{1,M}$, $l=\overline{1,M}$

Случайная последовательность $X(k)$ разлагается в дискретный ряд теоретически с нулевой ошибкой. Практически, естественно, будут иметь место неизбежные погрешности оценки параметров последовательности, в частности, математического ожидания, и преобразования к виду, удобному для обработки на ЭВМ (дискретизации и квантования).

Отметим, что ограничения на число N членов ряда (8) и (19) не накладывались. Можно определить минимально необходимое число членов этого ряда в зависимости от свойств функции $X(t)$ и максимально допустимой ошибки представления σ_N^2 .

Не накладывались также ограничения на характеристики случайного процесса (последовательности), кроме очевидного ограничения $0<\sigma_x^2(t)<\infty$ (для случайной последовательности $0<\sigma_x^2(k)<\infty$), вытекающего из условия неравенства нулю и ограниченности функции (7).

Однако, чтобы обеспечить заданную точность разложения и, в то же время, приемлемый для практических приложений объем вычислений, необходимо, чтобы интервал дискретизации $t_k - t_{k-1}$ был значительно меньше интервала корреляции случайной последовательности. В противном случае координатные функции быстро убывают, и для достижения заданной точности может потребоваться разложение с неприемлемо большим числом членов ряда.

Существуют и другие методы разложения случайной функции, например, разложение в ряд по собственным функциям (разложение Карунена-Лоэва) [5]. Для его реализации не требуется дискретизации исходной случайной функции. Однако для отыскания координатных функций разложения необходимо решать интегральной уравнение Фредгольма второго рода, ядром которого является корреляционная функция. Решение уравнения Фредгольма в замкнутой форме можно получить только для простейших видов ядра. Как правило, для практических задач решение приходится искать численными методами. Если, как это часто бывает на практике, интервал наблюдения меняется, все вычисления необходимо проделывать заново.

Как отмечалось выше, в рассматриваемой задаче случайная функция является дискретной (случайной последовательностью). Проблема дискретизации в данном случае решается естественным образом. Поэтому применение для анализа случайных входных последовательностей метода канонического разложения В. С. Пугачева достаточно логично и обоснованно.

Рекуррентный алгоритм оценивания

В реальной ситуации функционирования системы сбора информации об экстремальных состояниях телекоммуникационной сети априорные данные о статистических характеристиках процесса, как правило, или отсутствуют полностью, или имеются только частично, самого общего характера. Поэтому при построении алгоритма разложения для выборки переменного (нарастающего) объема необходимо одновременно оценивать требуемые характеристики, учитывая вновь получаемые данные. К таким характеристикам относятся математическое ожидание, дисперсия, корреляционная функция процесса, плотности распределения коэффициентов разложения V_ν.

Теоретические основы оценивания параметров случайных величин (процессов) широко освещены в литературе по теории вероятностей и математической статистике (см., например, [3, 4, 6-8]). Практические вопросы измерения и расчета случайных величин также изучены достаточно исчерпывающе [9-14]. Поэтому не будем здесь останавливаться на свойствах оценок (состоятельность, несмещенность, эффективность), полагая, что они обеспечиваются при построении алгоритма. Запишем выражения, необходимые для оценивания математического ожидания, дисперсии, значений центрированной случайной величины и ее корреляционной функции по $(N+1)$-му элементу выборки нарастающего объема и текущего уточнения канонического разложения (23):

$$m_{N+1}(k)=\frac{1}{N+1}\left[N m_N(k)+X_{N+1}\right]; \qquad (26)$$

$$\Delta m(k)=\frac{1}{N+1}\left[X_{N+1}(k)-m_N(k)\right]$$

(приращение математического ожидания на $N+1$-м шаге по отношению к математическому ожиданию на N-м шаге);

$$\sigma^2_{N+1}(k)=\frac{N}{N-1}\sigma^2_N(k)+\Delta m^2(k)+ \\ +\frac{1}{N}\left[\overset{0}{X}_{N+1}(k)\right]^2,\ k=\overline{1,M} \quad ; \tag{27}$$

$$\overset{0}{X}_{n,N+1}(k)=\begin{cases}\overset{0}{X}_{nN}-\Delta m(k),\ n\le N,\\ X_{N+1}(k)-m_{N+1}(k),\ n=N+1,\end{cases} \tag{28}$$
$$k=\overline{1,M},\ M\subseteq N;$$

$$R_{N+1}(k,l)=\frac{N-1}{N}R_N(k,l)+ \\ +\Delta m(k)\Delta m(l)+ \\ +\frac{1}{N}\overset{0}{X}_{N+1}(k)\overset{0}{X}_{N+1}(l) \quad ; \tag{29}$$

$$V_{kl,N+1}=V_{kl,N}-\sum_{\nu=1}^{k-1}V_{\nu l,N}\,w_\nu(k)-\Delta_m(k), \\ k=\overline{2,M},\ l=\overline{1,N}, \tag{30}$$

где

$$w_{\nu,k-1}(k)=\begin{cases}w_{\nu,k-2}(k)-w_{\nu,k-2}(k-1)\varphi_{k-1,N+1}(k),\ \nu<k-1\\ \varphi_{\nu,N+1}(k)-\varphi_{\nu N}(k),\ \nu=k-1\end{cases}$$

– весовой коэффициент, определяемый через оценки координатных функций;

$$\Delta_m(k)=\Delta m(k)-\sum_{\nu=1}^{k-1}\Delta_m(\nu)\varphi_{\nu,N+1}(k),\quad k=\overline{1,N}$$

– поправка, обусловленная различиями в оценках математических ожиданий;

$$\varphi_{k,N+1}(l)=\frac{N-1}{N}\frac{\sigma^2_{kN}\varphi_{kN}(l)+\left[\Delta m(k)-\Delta m(k-1)\right]\Delta m(l)}{\sigma^2_{k,N+1}}+ \\ +\frac{1}{N}\frac{\left[\overset{0}{X}_{N+1}(k)-\overset{0}{X}_{N+1}(k-1)\right]\overset{0}{X}_{N+1}(l)}{\sigma^2_{k,N+1}},\ k=\overline{1,N},\ l=\overline{1,N}. \tag{31}$$

Рассмотрим методику оценки плотностей вероятности коэффициентов V_ν. Как отмечалось ранее, эти коэффициенты являются некоррелированными. Однако, исходя из физического смысла решаемой задачи, невозможно сделать какие-либо предположения о теоретической функции распределения этих коэффициентов. Следовательно, из некоррелированности коэффициентов в общем случае не вытекает их независимость. Тем не менее, наличие зависимости обусловлено лишь существующими в исследуемом процессе стохастическими связями высших порядков (выше второго), которые в большинстве практических приложений относительно малы и быстро убывают при увеличении временного интервала. По данным соображениям такими связями можно пренебречь и сделать допущение о независимости коэффициентов V_ν . Тогда задача сводится к классической задаче непараметрической статистики восстановления плотности вероятности случайной величины по выборке ее реализаций.

С учетом свойства непрерывности случайных величин V_ν , которое вытекает из непрерывности случайной функции $X(t)$, можно применить непараметрическую оценку Парзена [15] вида:

$$F_N(V)=\frac{1}{N d_f}\sum_{k=1}^{N}g(u_k), \tag{32}$$

где d_f – некоторая константа, называемая коэффициентом размытости;

$g(u_k)$ – весовая сглаживающая функция или функция ядра;

$u_k=(V-V_k)/d_f$, V_k – k-я реализация случайной величины V .

В [14] показано, что при условиях

$$g(u)\ge 0;\ \sup_u|g(u)|<\infty; \\ \lim_{u\to\infty}|u\,g(u)|=0;\quad \int_{-\infty}^{\infty}g(u)\,du=1 \tag{33}$$

и выборе d_f в зависимости от числа наблюдений с соблюдением условий

$$d_f>0;\quad \lim_{N\to\infty}d_f(N)=0; \\ \lim_{N\to\infty}d_f(N)N=\infty \tag{34}$$

оценка (32) является несмещенной, состоятельной и эффективной.

Имея набор реализаций V_k, $k=\overline{1,N}$ и задав каким-либо образом d_f и $g(u)$, можно однозначно определить плотность вероятности случайной величины V . При этом нет необходимости строить гистограммы распределений, и можно упростить трудоемкую процедуру разбиения гистограмм на неэквидистантные интервалы для группирования данных в группы одинакового объема по времени суток (рис. 1-3).

Практическая методика выбора сглаживающего ядра и коэффициентов размытости, удовлетворяющих условиям (33-34), предложена в [16]. Показано, что при условии симметричности ядра $g(u)=g(-u)$ его структура вида:

$$g(u)=\begin{cases} a-bu^2, & |u|\le c, \\ 0, & |u|>c, \end{cases} \qquad (34)$$

является оптимальной по критерию минимума интегральной среднеквадратической ошибки (СКО) аппроксимации. Здесь a, b, c – некоторые константы, выбираемые исходя из особенностей решаемой задачи. Известно, что выбор функции сглаживающего ядра, спадающей к краям, практически всегда дает результат, наиболее близкий к оптимальному. В то же время для рассматриваемой задачи отклонение формы ядра от приведенной выше не слишком критично. Например, при использовании простейшей – прямоугольной функции ядра СКО аппроксимации возрастает всего на 6%. При этом значительно снижается трудоемкость вычислений, которые необходимо выполнять в реальном масштабе времени.

Приведем некоторые соображения по выбору коэффициента d_f – параметра, которым определяется интервал ненулевых значений ядра. Если выбрать его слишком большим, оценка будет слишком сглаженной, нечувствительной к быстрым отклонениям случайной величины. Если же значение d_f слишком мало, оценка будет сглажена недостаточно, будет «зашумленной». Оптимальным между этими крайними позициями для прямоугольной функции ядра будет выбор $d_f=0{,}5\sup\limits_k |V_k-V_{k-1}|,\ V_k \ge V_{k-1},\ k=\overline{2,N}$ как половина наибольшего расстояния между двумя соседними членами случайной последовательности. При этом коэффициент d_f , очевидно, зависит от параметров выборки, что гарантируется отсутствие разрывов области определения оценки $F_N(V)$, т. е. зашумление, и минимальное «заглаживание» оценки не в среднем по всему множеству выборок, а для каждой конкретной выборки.

Таким образом, при использовании прямоугольной функции сглаживания и приведенного выше правила вычисления коэффициентов d_f выражение для оценки $F_N(V)$ имеет вид:

$$F_N(V)=\frac{1}{N}\sum_{k=1}^{N} g_u(V), \qquad (35)$$

где

$$g_u(V)=\begin{cases} 1/2d_f, & V_k-d_f\le V\le V_k+d_f, \\ 0, & |V_k-V|>d_f, \end{cases} \quad k=\overline{1,N}. \qquad (36)$$

При практической реализации рассмотренной методики на ЭВМ с ограниченным быстродействием и объемом памяти учет длительности последействия реальных процессов является весьма актуальной задачей для экономии ресурсов вычислителя. Эффективной оценкой длительности последействия является нормированный коэффициент корреляции процесса, выраженный через координатные функции канонического разложения (см. выражение (25)):

$$r_{\nu x}(\nu,k)=\frac{\sigma_\nu \varphi_\nu(k)}{\sigma_x(k)},\ \nu=\overline{1,N},\ k=\overline{1,N}\ . \qquad (37)$$

В [8] показано, что функцию некоторого специального вида от случайной величины (37) можно считать распределенной по гауссовскому закону с соответствующими математическим ожиданием и дисперсией. Там же определена область принятия гипотезы о равенстве коэффициента корреляции нулю. Таким образом, используя проверку данной гипотезы

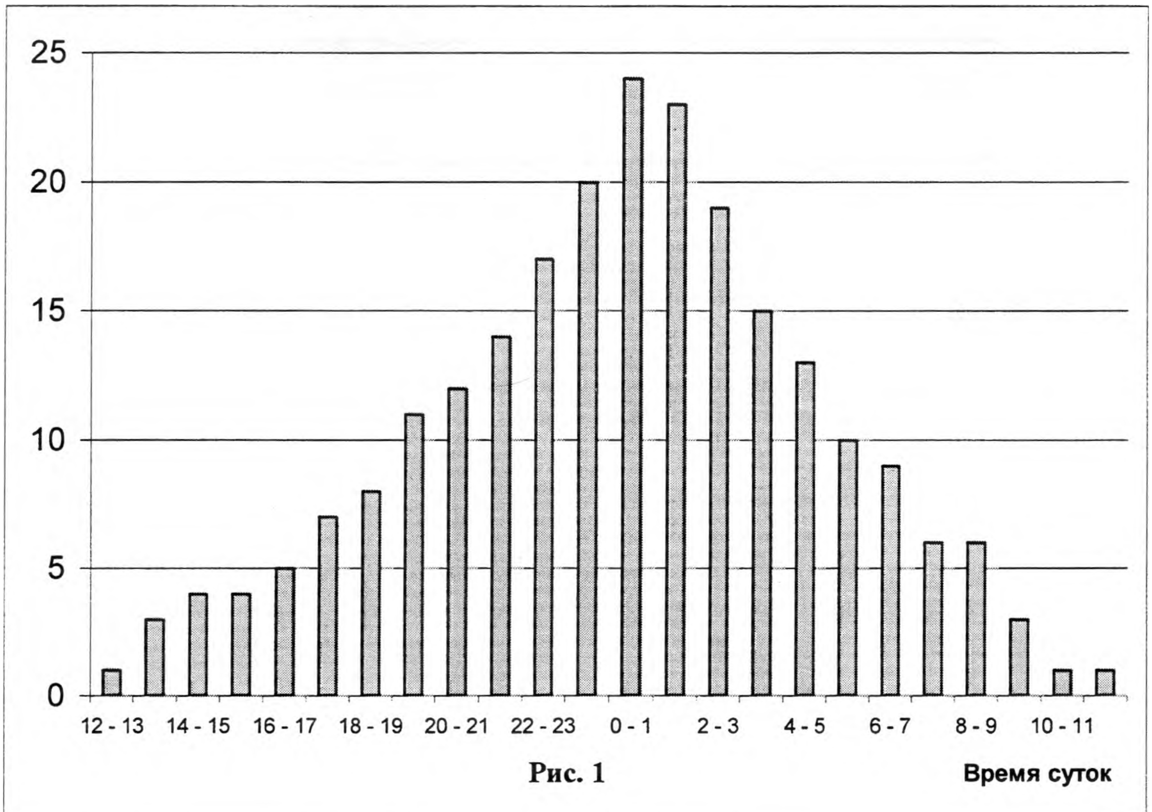
25
20
15
10
5
0
12 - 13
14 - 15
16 - 17
18 - 19
20 - 21
22 - 23
0 - 1
2 - 3
4 - 5
6 - 7
8 - 9
10 - 11
Время суток

Рис. 1

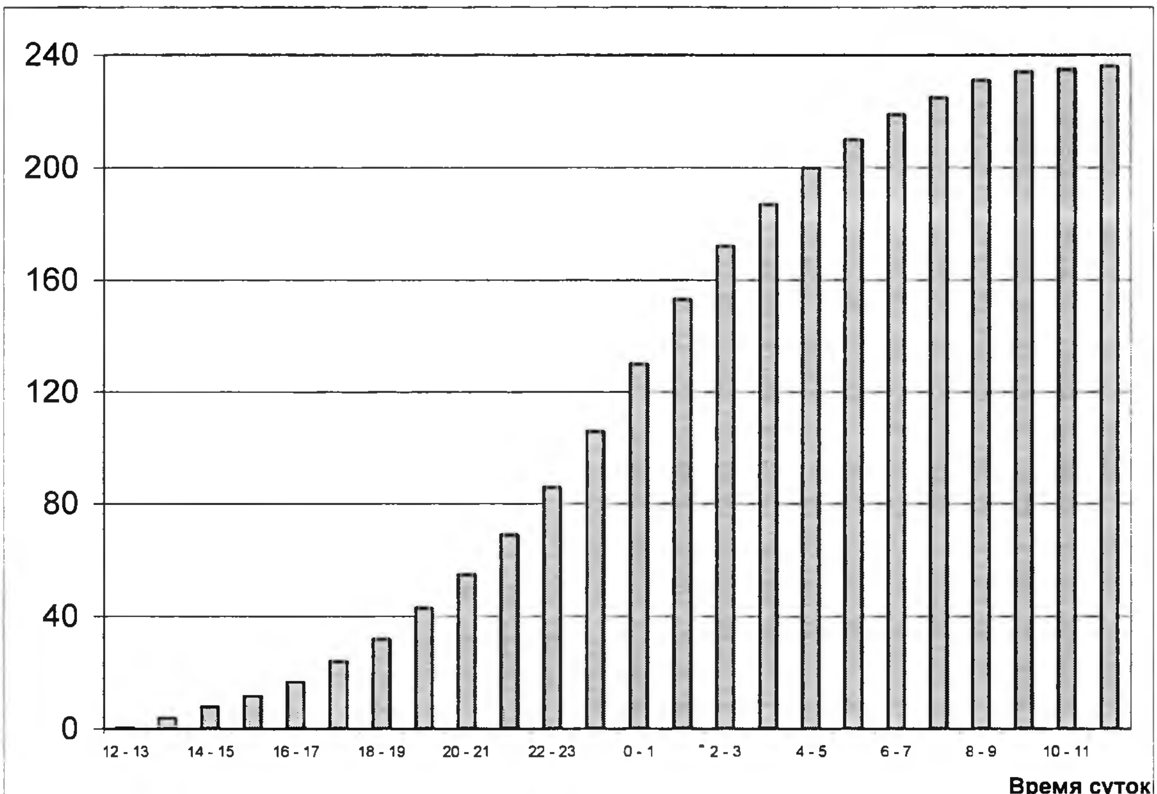
240
200
160
120
80
40
0
12 - 13
14 - 15
16 - 17
18 - 19
20 - 21
22 - 23
0 - 1
2 - 3
4 - 5
6 - 7
8 - 9
10 - 11
Время суток

Рис. 2

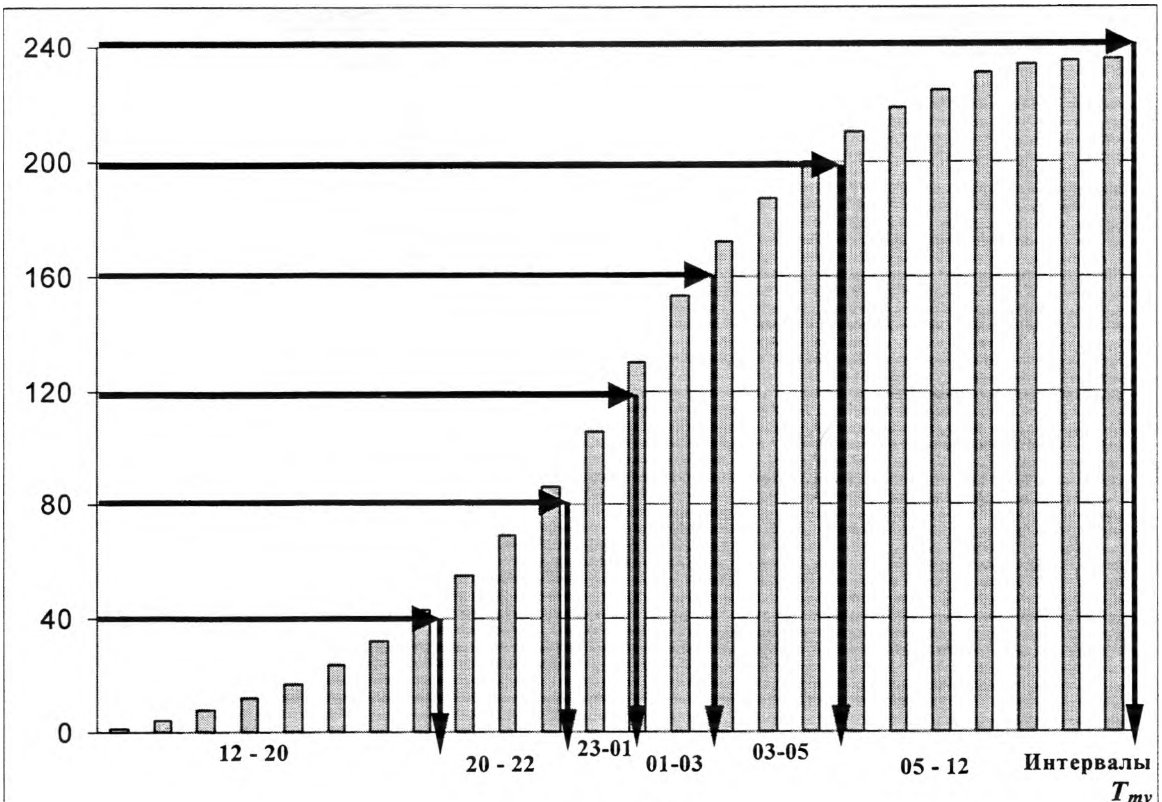

Рис. 3

против простой альтернативы о неравенстве коэффициента корреляции нулю параллельно с общим алгоритмом обработки данных, можно на каждом этапе вычисления параметров разложения определять необходимый объем хранимой информации.

Для практической реализации методики разбиения области значений δr случайной величины $X(k)$ на неэквидистантные интервалы применим метод Н. П. Бусленко [17]. Этот метод сводится к аппроксимации плотности распределения вида (36) ступенчатой кривой, ограничивающей сверху ряд прямоугольников одинаковой площади. Общее число N_r интервалов разбиения области δr определяют практическим путем. Обычно исходят из реально достижимой для данного объема накопленных данных точности аппроксимации. Координаты точек разбиения r_i вычисляют рекуррентно с помощью уравнения

$$\int_{r_{i-1}}^{r_i} F_N(y)dy = \frac{1}{N_r}, \quad i = \overline{1, N_r}. \qquad (38)$$

В уравнении (38) диапазон интегрирования определяется с учетом вида функции $F_N(y)$: $r_0 = \min_k x_k - d_f$, $r_{Nr} = \max_k x_k + d_f$.

Процедура вычисления $F_N(y)$ при произвольном значении y сводится к определению номера интервала r_i, в который попадает результат наблюдения y: $r_{i-1} \le y \le r_i$, после чего искомое значение определяется как

$$F_i(y) = \frac{1}{N_r(r_i - r_{i-1})}. \qquad (39)$$

Поскольку все величины, входящие в уравнения (38-39), известны, процедура определения неэквидистантных интервалов разбиения функции распределения попыток НСД по критерию постоянной интенсивности не зависит от размера выборки, достаточно проста и экономна для реализации на ЭВМ.

Выводы

1. В представленной работе предложена методика компьютерной оценки функций распределения попыток НСД на автономный телекоммуникационный объект при весьма общих предположениях о характере потока. Методика основана на канонических разложениях случайной функции. Для разбиения функции распределения на неэквидистантные интервалы с постоянными интенсивностями потока используется рекуррентный алгоритм, экономичный и легко реализуемый на ЭВМ.

2. Результаты расчетов могут быть использованы для оптимального распределения ресурсов – сил и средств для предотвращения попыток НСД и, в конечном счете, обеспечения требуемого уровня защиты телекоммуникационных сетей.

Литература

1. *Коробко В. В., Скоропаденко А. П., Задоя Г. М., Вовк В. М.* Интегрированная система сбора информации об экстремальных состояниях и защиты телекоммуникационных сетей// Зв'язок. – №1. – 2004.
2. *Вальд А.* Последовательный анализ. – М.: Физматгиз, 1960.
3. *Пугачев В. С.* Теория случайных функций. – М.: Физматгиз, 1962. – 784 с.
4. *Прохоров Ю. В., Розанов Ю. А.* Теория вероятностей. Основные понятия. Предельные теоремы. Случайные процессы. – М.: Наука, 1967. – 496 с.
5. *Лоэв М.* Теория вероятностей. Пер. с англ. Под ред. Ю.В. Прохорова. – М.: Изд-во иностранной литературы, 1962. – 720 с.
6. *Вентцель Е. С.* Теория вероятностей. – М.: Советское радио, 1969. – 576 с.
7. *Пугачев В. С.* Теория вероятностей и математическая статистика. – М.: Наука, 1979. – 496 с.
8. *Крамер Г.* Математические методы статистики.– М.: Мир, 1975. – 648 с.
9. *Дунин-Барковский И. В., Смирнов Н. В.* Теория вероятностей и математическая статистика в технике. – М.: Наука, 1965. – 511 с.
10. *Романенко А. Ф., Сергеев Г. А.* Вопросы прикладного анализа случайных процессов. – М.: Советское радио, 1968. – 256 с.
11. *Мирский Г. Я.* Характеристики стохастической взаимосвязи и их измерения. – М.: Энергоиздат, 1982. – 320 с.
12. *Бендат Дж., Пирсол А.* Измерение и анализ случайных процессов. – М.: Мир, 1974. – 464 с.
13. *Отнес Р., Эноксон Л.* Прикладной анализ временных рядов. Основные методы. – М.: Мир, 1982. – 428 с.
14. *Бендат Дж., Пирсол А.* Применения корреляционного и спектрального анализа: Пер. с англ. – М.: Мир, 1983. – 312 с.
15. *Parzen E.* On Estimation of a Probability Density Function and Mode // Annals of Mathematical Statistics, 1962. – v. 33, – N 3. – pp. 1065-1076.
16. *Епанечников В. А.* Непараметрическая оценка многомерной плотности вероятности // Теория вероятностей и ее применения. – 1969. – № 1. – С. 156-161.
17. *Бусленко Н. П., Шрейдер Ю. А.* Метод статистических испытаний. – М.: Физматгиз, 1961. – 226 с.