УДК 621.391

НЕРАВНОМЕРНОЕ ПО ВХОДУ КОДИРОВАНИЕ СООБЩЕНИЙ, ПОРОЖДЕННЫХ СТАЦИОНАРНЫМ ИСТОЧНИКОМ

В.К. Трофимов, В.И. Агульник, И.И. Резван

Рассмотрено пословное кодирование сообщений, порожденных известным стационарным источником. Предложены эффективные методы кодирования, переводящие слова различной длины в слова одинаковой длины; либо слова различной длины в слова различной длины.

Ключевые слова: энтропия, кодирование, избыточность, стоимость кодирования, источник сообщений.

Основные определения. Постановка задачи

Из-за возрастающего потока и объема информации большое значение имеет проблема сжатия (кодирования) информации [1]. Решение этих проблем значимо и при создании большемасштабных распределенных вычислительных систем [2]. Методы сжатия информации в таких системах, как правило, используют параллельные информационно-вычислительные технологии.

Настоящая работа посвящена кодированию информации, порожденной источником, в её классической форме, предложенной К.Шенноном [1]. Для постановки задачи и формулировки основных утверждений приведем основные определения и обозначения.

Пусть буквы конечного алфавита $A=\{a_1,a_2,...,a_k\}$, $2\le k<\infty$, порождаются источником θ. Мера, заданная на последовательности букв, порождаемой источником, определяет тип источника. Если вероятности порождения букв независимы, то источник называют бернуллиевским. В этом случае $P_\theta(a_j)=\theta_j$, $\theta_1+\theta_2+\ldots+\theta_k=1$. Если же вероятность появления очередной буквы зависит от предыдущей, то $P_\theta(a_i/a_j)=\theta_{ij}$, $\sum_{i=1}^{k}\theta_{ij}=1$, $i=\overline{1,k}$, и в этом случае источник называют марковским. Если вероятность появления очередной буквы зависит от s предшествующих букв, т.е. $P_\theta(a_j/\nu)=\theta_{\nu j}$, где $\nu\in A^s$, то источник θ называют марковским с памятью s. Следует отметить, что для любого слова $\nu\in A^s$, $0\le s<\infty$, выполняется равенство $\sum_{j=1}^{k}\theta_{\nu j}=1$. Множество всех марковских источников с памятью s обозначим Ω_s. Дискретный стационарный источник θ задаётся всеми условными распределениями вероятностей $P_\theta(a_j/\nu)=\theta_{\nu j}$ порождения источником букв a_j, $j=\overline{1,k}$, при заданных ν предшествующих, $\nu\in A^s$, s любое целое неотрицательное число. Здесь, как и выше, при любом заданном ν, $\nu\in A^s$ выполняется равенство:

$$\sum_{j=1}^{k}\theta_{\nu j}=1;\quad s=0,1,2,\ldots$$

Множество всех стационарных источников обозначим Ω_∞. Если u – произвольное слово в алфавите A, то через $P_\theta(u)$ обозначим вероятность слова u, порожденного источником θ. Число $|u|$ букв в слове u назовем его длиной. Энтропию источника θ обозначим $H(\theta)$. Как известно [3-5], если θ – стационарный источник, то

$$H(\theta)=-\lim_{n\to\infty}\frac{1}{n}\sum_{u\in A^n}P_\theta(u)\cdot\log P_\theta(u). \quad (1)$$

Здесь и в дальнейшем $\log x=\log_2 x$, $0\log 0=0$.

Для бернуллиевского источника θ его энтропия $H_0(\theta)$ определяется равенством

$$H_0(\theta)=-\sum_{i=1}^{k}\theta_i\log\theta_i\ . \quad (2)$$

Если θ марковский источник с памятью s, то его энтропия $H_s(\theta)$ находится по формуле

$$H_s(\theta) = -\sum_{\nu \in A^s} \theta_{0\nu} \sum_{i=1}^{k} \theta_{\nu i} \log \theta_{\nu i}, \qquad (3)$$

где $\theta_{0\nu}$ - начальные стационарные вероятности слов ν, $v \in A^s$. При $s = 0$ из (3) получаем соотношение (2). Если θ – произвольный стационарный дискретный источник и $H(\theta)$ его энтропия, то справедливо равенство [3-5]

$$H(\theta) = \lim_{s \to \infty} H_s(\theta). \qquad (4)$$

Рассмотрим T - конечное полное множество слов во входном алфавите. Множество T – полное, если оно префиксное и при любом непустом слове u (в алфавите A) множество слов $T \cup u$ уже не префиксное. Такое множество T назовем кодовым. Примером кодового множества может служить множество всех слов длины *n* взятых в алфавите A, т.е. A^n; множество $A^n \setminus \underbrace{a_1, \ldots, a_k}_{n}$, не является кодовым, потому что оно не полное.

Пусть θ – произвольный источник из Ω_s, T – произвольное кодовое множество. Обозначим через $\theta(T)$ марковскую цепь, состояниями которой являются слова из T, а переходные вероятности $P_{\theta(T)}(u/\nu)$, $u, \nu \in T$, индуцируются источником θ. Будем рассматривать только марковские источники с памятью s, переходные вероятности которых строго положительны. Тогда для любых $u, \nu \in T$ выполняются неравенства $P_{\theta(T)}(u/\nu) > 0$, поэтому для марковской цепи $\theta(T)$ существует стационарное распределение $P^0_{\theta(T)}(u) > 0$, $u \in T$. Средняя длина слова $d_s(T, \theta)$ для множества T, как доказано в [6], равна

$$d_s(T, \theta) = \sum_{u \in T} P^0_{\theta(T)}(u) \cdot |u|. \qquad (5)$$

В этой же работе доказаны тождества Вальда, которые имеют вид

$$\sum_{u \in T} P^0_{\theta(T)}(u) \cdot r_\nu(u) = (d_s(T, \theta) - \hat{s} + 1)\theta_{0\nu}, \quad (6)$$

$$\sum_{u \in T} P^0_{\theta(T)}(u) \cdot r_{\nu i}(u) = (d_s(T, \theta) - s)\theta_{0\nu}\theta_{\nu i} \quad (7)$$

где $r_\nu(u)$, $r_{\nu i}(u)$ - число вхождений блоков v, va_i, $v \in A^s$, в слово u, соответственно, $\hat{s} = \max(s, 1)$.

Полубесконечная последовательность букв, порождаемая источником θ, однозначно разбивается на последовательность слов из фиксированного кодового множества *T*. Полученная последовательность слов из *T* с помощью отображения φ переводятся в слова выходного алфавита *B*, который не уменьшая общности можно считать двоичным. Из неравенства Мак-Милана-Крафта [3-5] следует, что самое общее из всех возможных дешифрируемых кодирований φ такое, что множество слов в выходном алфавите $\varphi(T) = \{\varphi(u),\ u \in T\}$ является префиксным.

Если длины всех слов некоторого множества *D* равны между собой, то говорят, что *D* состоит из блоков; в противном случае из слов переменной длины. В зависимости от видов множества *T* и $\varphi(T)$ логически возможны следующие виды кодирований:

1) кодирование, отображающее блоки в слова переменной длины (обозначается BV);

2) кодирование, отображающее слова переменной длины в блоки (VB);

3) кодирование, отображающее слова переменной длины в слова переменной длины (VV);

4) кодирование, отображающее блоки в слова переменной длины (BB).

Итак, всякое кодирование φ однозначно определяется тройкой $(T, \varphi, \varphi(T))$. Среднее число букв выходного алфавита при кодировании типа σ, $\sigma = BV, VB, VV$, приходящихся на одну букву входного, назовем стоимостью кодирования и обозначим через $C_\sigma(T, \theta, \varphi)$. В [6] доказано, что стоимость кодирования типа σ, $\sigma = BV, VB, VV$, для произвольного кодового множества T и любого источника θ, $\theta \in \Omega_s$, $0 \le s < \infty$, определяется равенством:

$$C_\sigma(T, \theta, \varphi) = \frac{1}{d_s(T, \theta) - \hat{s} + 1} \sum_{u \in T} P^0_{\theta(T)}(u) \cdot |\varphi(u)| \quad (8)$$

Эффективность кодирования φ как обычно [1,4,5] будем оценивать разностью между стоимостью кодирования $C_\sigma(T, \theta, \varphi)$

и энтропией источника $H(\theta)$. Эта разность в дальнейшем называется избыточностью кодирования и обозначается $r_\sigma(T, \theta, \varphi)$, т.е.

$$r_\sigma(T, \theta, \varphi) = C_\sigma(T, \theta, \varphi) - H(\theta). \quad (9)$$

Избыточностью кодирования типа σ для источника θ с заданной сложностью N, назовем величину $R_\sigma(N, \theta)$:

$$R_\sigma(N, \theta) = \inf_\varphi r_\sigma(T, \theta, \varphi). \quad (10)$$

Здесь нижняя грань берется по всем кодированиям φ, для которых кодовое множество *T* имеет не более чем k^N слов. Построение хорошего кодирования при заданной сложности – основной вопрос при изучении передачи сообщений по каналу без шума. Решение поставленной задачи позволяет ответить на вопрос: «какой избыточности можно достигнуть при заданной сложности кодирования?»

Кодирование информации, порожденной известным источником, подробно изучено для различных типов кодирования, например, в работах [1, 3, 6 – 14]. Универсальное кодирование марковских источников различных типов также хорошо изучено. Подробную библиографию по этому вопросу можно найти в [15 – 19].

Равномерное по выходу кодирование информации, порожденной известным марковским источником

В этом параграфе будет предложен метод кодирования известных марковских источников с памятью s, получена оценка избыточности предложенного метода. При доказательстве основного утверждения параграфа нам потребуются следующие понятия и обозначения. Марковский источник θ связанности s задается начальным распределением вероятностей $\theta_{0\nu}$ появления блока ν за первые s шагов работы источника и вероятностями $\theta_{\nu i}$ появления буквы a_i, после блока ν, $a_i \in A$, $\nu \in A^s$. θ. Отсюда следует, что вероятность $P_\theta(u)$ порождения слова u, $|u| > s$, начинающегося блоком ν, $\nu \in A^s$, источником θ определяется равенством

$$P_\theta(u) = \theta_{0\nu} \prod_{\nu \in A^s} \prod_{i=1}^{k} \theta_{\nu i}^{r_{\nu i}(u)}. \quad (11)$$

В работе [12] доказано, что для произвольного марковского источника θ, с памятью s, $0 \le s < \infty$, существует последовательность кодовых множеств $\{T_N^\theta\}$ такая, что избыточность равномерного по выходу кодирования φ_N информации, порожденной источником θ с областью определения T_N^θ, при *N*, стремящемся к бесконечности, стремится к нулю. Более точно – доказана справедливость неравенств

$$0 \le r_{VB}\left(T_N^\theta, \theta, \varphi_N\right) < \frac{c}{d_S\left(T_N, \theta\right)} \quad (12)$$

Неравномерное по входу и выходу кодирование информации, порожденной марковским источником.

Рассмотрим кодирование информации, неравномерное по входу и выходу, порожденное марковским источником с памятью *S*. Основной результат этого параграфа можно сформулировать следующим образом.

<u>Теорема 1</u>. Для любого марковского источника с памятью s, $0 \le s < \infty$, для любой последовательности кодовых множеств в $\{T_N^\theta\}$, $N = 1, 2, \cdots$ такой, что $\min_{u \in T_N} |u|$ стремится к бесконечности с ростом N, существует эффективное кодирование, переводящее слова множества T_N, $N = 1, 2, \cdots$, в слова переменной длины.

<u>Доказательство</u>. Пусть $\theta \in \Omega_s$ произвольный марковский источник с памятью *s*, $0 \le s < \infty$, и $\{T_N^\theta\}$, $N = 1, 2, \cdots$ последовательность кодовых множеств, для которой $\lim_{N \to \infty} \min_{u \in T_N} |u| = \infty$. Зафиксируем *N* и выберем кодовое множество T_N, тогда для любого слова *u* из T_N его вероятность вычисляется по формуле (11). Рассмотрим кодирование φ_N, которое каждому слову $u \in T_N$ ставит в соответствие слово $\varphi_N(u)$ длины

$$|\varphi_N(u)| =]-\log P_\theta(u)[, \; u \in T_N \quad (13)$$

Так как выполнено равенство $\sum_{u \in T_N} P_\theta(u) = 1$, то для длин слов $|\varphi_N(u)|, u \in T_N$ выполнено неравенство

Крафта и, следовательно, дешифруемое кодирование с длинами кодовых слов, определенных равенствами (13) существует. Оценим избыточность такого кодирования. Из равенств (9), (8) и (3) имеем:

$$r_{VV}\left(T_N,\theta,\varphi_N\right)=\frac{1}{d_s\left(T_N,\theta\right)-\hat{s}+1}\cdot \quad (14)$$

$$\cdot\sum_{u\in T_N}P^0_{\theta(T)}(u)\left|\varphi_N(u)\right|+\sum_{\nu\in A^s}\theta_{0\nu}\sum_{i=1}^{k}\theta_{\nu i}\log\theta_{\nu i},$$

Из (14), учитывая (11) и (13), получаем

$$r_{VV}\left(T_N,\theta,\varphi_N\right)=\frac{1}{d_s\left(T_N,\theta\right)-\hat{s}+1}\cdot$$

$$\cdot\sum_{u\in T_N}\sum_{\nu\in A^s}\sum_{i=1}^{k}P^0_{\theta(T)}(u)\cdot\left(-\log\theta_{0\nu}-r_{\alpha i}\log\theta_{\nu i}\right)+ \quad (15)$$

$$+\sum_{\nu\in A^s}\theta_{0\nu}\sum_{i=1}^{k}\theta_{\nu i}\log\theta_{\nu i}.$$

Используя тождества Вальда и применяя для функции $-x\log x$ неравенство Иенсена, из (15) получаем

$$r_{VV}\left(T_N,\theta,\varphi_N\right)\le\frac{1}{d_s\left(T_N,\theta\right)-\hat{s}+1}\left(-\sum_{\nu\in A^s}\theta_{0\nu}\log\theta_{0\nu}+1\right).$$

Так как $d_s\left(T_N,\theta\right)\ge\min_{u\in T_N}\left|u\right|$, то из последнего неравенства получаем

$$r_{VV}\left(T_N,\theta,\varphi_N\right)\le\frac{c}{\min_{u\in T_N}\left|u\right|} \quad (16)$$

Согласно условию $\lim_{N\to\infty}\min_{u\in T_N}\left|u\right|=\infty$, из (16) следует, что $r_{VV}\left(T_N,\theta,\varphi_N\right)$ стремится к нулю с ростом *N*, а это и означает, что стоимость кодирования φ_N с ростом *N* стремится к энтропии источника. Тем самым эффективность кодирования φ_N доказана.

Кодирование VB и VV для известных стационарных источников

В этом параграфе доказаны основные утверждения работы. Имеет место утверждение.

Теорема 2. Для произвольного стационарного источника θ существует эффективное кодирование слов переменной длины блоками.

Доказательство. Каждый стационарный источник θ, $\theta\in\Omega_\infty$, задается условными вероятностными распределениями $\theta_s\left(a_i\middle|\nu\right)$, $a_i\in A$, $\nu\in A^s$, $s=0,1,2,\ldots$ появления буквы a_i после блока ν. Таким образом, каждый стационарный источник θ определяет последовательность марковских источников θ_s, $s=0,1,2,\ldots$, при s, стремящемся к бесконечности, энтропия $H\left(\theta_s\right)$ источника θ_s, не возрастая, сходится к энтропии $H\left(\theta\right)$ источника θ, т.е. верны соотношения:

$$H\left(\theta_0\right)\ge H\left(\theta_1\right)\ge\ldots\ge H\left(\theta_s\right)\ge H\left(\theta_{s+1}\right)\ge\ldots \quad (17)$$

и выполнено равенство (4).

Для любого фиксированного s, $0\le s<\infty$, определена стоимость кодирования $C_{VB}\left(T,\theta,\varphi\right)$ (см.(8)). Покажем, что стоимость $C\left(T_N^s,\theta,\varphi^s\right)$ для кодирования типа VB, предложенного в [11], при N и s, стремящимися к бесконечности, существует и равна энтропии источника $H\left(\theta\right)$. Для этого нам нужно установить, что избыточность кодирования $r_{VB}\left(T_N^s,\theta,\varphi_N^s\right)$ для стационарного источника $\theta,\ \theta\in\Omega_\infty$ стремится к нулю с ростом N и s.

Согласно определениям (8,9) имеем

$$r_{VB}\left(T_N^s,\theta,\varphi_N^s\right)=\frac{\left\lceil\log\left\|T_N^s\right\|\right\rceil}{d_s\left(T_N,\theta\right)-\hat{s}+1}-H\left(\theta\right), \quad (18)$$

или

$$r_{VB}\left(T_N^s,\theta,\varphi_N^s\right)=\left[\frac{\left\lceil\log\left\|T_N^s\right\|\right\rceil}{d_s\left(T_N,\theta\right)-\hat{s}+1}-H\left(\theta_s\right)\right]+ \quad (19)$$

$$+\left[H\left(\theta_s\right)-H\left(\theta\right)\right].$$

В равенстве (19) первое слагаемое в правой части, согласно (12), ограничено асимптотически сверху величиной $c/d_s\left(T_N,\theta\right)$ и поэтому с ростом *N* стремится к нулю. Из соотношения (4) следует, что с ростом s второе слагаемое также стремится к нулю. Таким образом, первое и второе слагаемые правой части равенства (19) стремятся к нулю одновременно с ростом *N*. Тогда из (18) вытекает, что $\lim_{\substack{N\to\infty\\ s\to\infty}} r_{VB}\left(T_N^s,\theta,\varphi_N^s\right)=0$, т.е.

$$\lim_{\substack{N\to\infty\\ s\to\infty}} C\left(T_N^s,\theta,\varphi_N^s\right)=H\left(\theta\right).$$

Теорема доказана.

Перейдём к исследованию неравномерного по входу и по выходу кодирования. Справедливо утверждение.

Теорема 3. Для произвольного стационарного источника θ и для любой последовательности кодовых множеств $\left\{T_N^{\theta}\right\}$, $N=1,2,\cdots$, *для которой величина* $\min\limits_{u\in T_N}|u|$ *с ростом N стремится к бесконечности, существует эффективное кодирование слов переменной длины в слова переменной длины.*

Доказательство. Отметим как и при доказательстве предыдущей теоремы, что для любого фиксированного s, $0\le s<\infty$, определена стоимость кодирования по формуле (8). Покажем, что стоимость кодирования $C\left(T_N^s,\theta,\varphi_N^s\right)$ для кодирования предложенного при доказательстве теоремы 1 при N и s стремящихся к бесконечности, равна энтропии источника $H(\theta)$. В самом деле, по определению (8,9) имеем

$$r_{VV}\left(T_N^s,\theta,\varphi_N^s\right)=$$

$$=\frac{1}{d\left(T_N^s,\theta_s\right)}\sum_{u\in T_N}P^0_{\theta(T_N^s)}(u)\left|\varphi^s(u)\right|-H(\theta)$$

или

$$r_{VV}\left(T_N^s,\theta,\varphi_N^s\right)=\left(r_{VV}\left(T_N^s,\theta,\varphi_N^s\right)-H_s(\theta)\right)+ \quad (20)$$
$$+\left(H_s(\theta)-H(\theta)\right)$$

В соотношении (20) первое слагаемое в правой части по теореме 1 стремится к нулю с ростом N, а второе слагаемое по соотношению (4,17) стремится к нулю при s стремящемся к бесконечности. Первоначально находим s_0, начиная с которого второе слагаемое меньше, чем $\varepsilon/2$, (где ε – сколь угодно малое число), а затем при фиксированном s_0, устремляя N к бесконечности, можно, согласно теореме 1, сделать и первое слагаемое меньше $\varepsilon/2$. Отсюда

$\lim\limits_{\substack{N\to\infty\\ s\to\infty}} r_{VV}\left(T_N^s,\theta,\varphi_N^s\right)=0$, т.е.

$\lim\limits_{\substack{N\to\infty\\ s\to\infty}} C_{VV}\left(T_N^s,\theta,\varphi_N^s\right)=H(\theta)$.

Теорема доказана.

СПИСОК ЛИТЕРАТУРЫ

1. Шеннон, К. Математическая теория связи. Работы по теории информации и кибернетике/ К.Шеннон – М.: 1969. С.243–332.
2. Хорошевский, В.Г. Архитектура вычислительных систем./ В.Г.Хорошевский - М.: МГТУ им.Н.Э.Баумана, 2008. – 520 с.
3. Тарасенко, Ф.П. Введение в курс теории информации./ Ф.П.Тарасенко – Томск, 1963. – 240 с.
4. Фано, Р. Передача информации. Статистическая теория связи./ Р.Фано – М.: 1965.
5. Галлагер, Р. Теория информации и надёжная связь./ Р.Галагер – М.: 1974. – 720 с.
6. Могульский, А.А. Тождество Вальда и стоимость кодирования для цепей Маркова. // А.А.Могульский, В.К.Трофимов. VII Всесоюзная конференция по теории кодирования и передачи информации. Доклады // Теория информации. – Москва-Вильнюс. 1978., ч.I. С.112–116.
7. Кричевский, Р.Е. Длина блока, необходимая для получения заданной избыточности.// Р.Е.Кричевский. Докл. АН СССР. 1966. Т.171, №1.
8. Гильберт, Э.Н. Двоичные кодовые системы переменной длины. // Э.Н.Гильберт, Э.Ф.Мур. Кибернетический сборник. – М.: 1961, № 3, С.103–141.
9. Ходак, Г.Л. Оценки избыточности при пословном кодировании сообщений, порождаемых бернуллиевским источником. // Г.Л.Ходак. Пробл. передачи информ. – 1972. Т.8. № 2. С.21–32.
10. Khodak, G.L. Coding of Markov Sources With Low Redundancy // G.L.Khodak. Proc. of 2nd International Symp. On Inform. Theory Tsahkadzor, Armenia. USSR, 1971, Akademiai Kiado. Budapest. 1973. P.201–204.
11. Jelinek, F. On Variable-Length to Block Coding // F.Jelinek, K.Shneider. IEEE Trans. Inform. Theory. –1972. V.18, №.6. P.756–774.
12. Трофимов, В.К. Эффективное кодирование блоками слов различной длины, порождённых известным марковским источником // В.К.Трофимов. Обработка информации в системах связи. – Л.: 1985. С.9–15.
13. Ziv, J. Variable-to-Fixed Length Codes are Better than Fixed-to-Variable Length Codes for Marcov Sources // J.Ziv. IEEE Trans. Inform. Theory. 1990. V.36. №.4. P.861–863.
14. Кричевский, Р.Е. Связь между избыточностью кодирования и достоверностью сведений об источнике // Р.Е.Кричевский. Пробл. передачи информ. 1968. Т.4. №3. С.48–57.
15. Krichevskii, R.E. The Performace of Universal Encoding // R.E.Krichevskii, V.K.Trofimov. IEEE Trans. on Inform. Theory. 1981. V.IT-27. №2. P.199–207.
16. Shtarkov, Yu.M. Combinatorial Encoding for Discrete Stationary Sources // Yu.M.Shtarkov,.

V.F.Babkin. Proc. of 2nd International Symp. On Inform. Theory Tsahkadzor, Armenia. USSR, 1971, Akademiai Kiado. Budapest. 1973., P.249–256.
17. Трофимов, В.К. Равномерное по выходу кодирование марковских источников при неизвестной статистике // В.К.Трофимов. V международный симпозиум по теории информации. Доклады. Москва – Тбилиси, 1979. ч.II. С.172–175.
18. Krichevsky, R. Universal Compression and Retrieval.// R.Krichevskii. Dordrecht/Boston/London: 1994. P.219.
19. Sergio Verdu. Fifty Years of Shannon Theory// Verdu Sergio. IEEE Trans.on Inform.Theory. 1998. V.IT 44. №6. P.2057-2077.

*Д.т.н., профессор, декан факультета информатики и вычислительной техники **В.К. Трофимов**, тел. (383) 269-82-70, e-mail: trofimov@sibsutis.ru; и.о.доцента **В.И. Агульник**, тел. (383) 269-82-71, e-mail: agulnik@sibsutis.ru; к.т.н., доцент **И.И. Резван**, e-mail: rez-van@rambler.ru; Сибирский государственный университет телекоммуникаций и информатики (г. Новосибирск).*

УДК 681.3.067

АЛГОРИТМ ПРОГНОЗИРОВАНИЯ ОЦЕНОК УРОВНЯ ЗАЩИЩЕННОСТИ ОБЪЕКТОВ ИНФОРМАТИЗАЦИИ НА ОСНОВЕ НЕЧЕТКИХ ВРЕМЕННЫХ РЯДОВ

Е.Н. Пивкин

В статье рассматривают задачу прогнозирования оценок уровня защищенности объектов информатизации на основе нечетких временных рядов. Предложен алгоритм процедуры прогнозирования и проведена апробация на тестовом примере.

Ключевые слова: нечеткий временной ряд, прогнозирование, оценка защищенности.

Определение оценок уровня защищенности объектов информатизации (ОИ), будучи неотъемлемым элементом модели информационной безопасности (ИБ) организации, играет важную роль в формировании комплекса практических мер по реализации ИБ.

Однако, принятие решений на основе классических моделей и методов [1] зачастую не дает желаемого эффекта в связи с неполнотой рассматриваемых данных, упрощениями и погрешностями, возникающими при их обработке. Основными причинами этого, как правило, считают:

1. Невозможность учета всех элементов, определяющих защищенность, влияющих на конечный результат.

2. Отсутствие полной непротиворечивой априорной информации.

3. Влияние на защищенность ОИ различных неконтролируемых воздействий (как внешних, так и внутренних).

Поэтому эмпирические данные (при отсутствии систематизированных статистических материалов) являются единственным исходным источником информации.

Следовательно, оценку уровня защищенности ОИ необходимо осуществлять с учетом того, что информация, лежащая в основе этой деятельности (модели, процедуры), – неполная, нечеткая. Применение в данном случае теории нечетких множеств (ТНМ) – в виде модели нечетких временных рядов [2-3] – можно считать логичным и естественным шагом. Так как погрешность оценки, по сравнению с другими подходами, минимальна [4-5].

Основой исследования являются эмпирические данные – временной ряд $\{\widetilde{Y}(t)\}$ наблюдений (оценки уровня защищенности ОИ за различные периоды). «Погружение» этого ряда в нечеткую среду позволит получить нечеткую функцию $\widetilde{Y}(t)$ аргумента t в универсальном множестве $X_{\text{уз}}$ со значениями в виде нечетких интервалов X^t с функцией принадлежности (ФП) $\mu_{X^t}(x_{\text{узi}})$, т.е. $X^t = \{\mu_{X^t}(x_{\text{узi}})/x_{\text{узi}}\}$, $u_i \in X_{\text{уз}}$, $\mu_{X^t}(x_{\text{узi}}) \in [0, 1]$. Последнее означает, что любая точка из интервала $x_{\text{узi}}$ будет принадлежать множеству X^t со степенью принадлежности $\mu_{X^t}(x_{\text{узi}}) = \mu_i(t)$ с заданными числами $\mu_i(t), i = \overline{1, m}$ при каждом фиксированном $t = 1, 2, \ldots$ Здесь $X_{\text{уз}} = (x_{\text{уз1}}, x_{\text{уз2}}, \ldots, x_{\text{узm}})$ – полное множество (в нашем случае интервал