Министерство образования Республики Беларусь
Учреждение образования
«Белорусский государственный университет
информатики и радиоэлектроники»

И. П. Кобяк

ЗАЩИТА ИНФОРМАЦИИ В ВЫЧИСЛИТЕЛЬНЫХ СЕТЯХ

*Рекомендовано УМО вузов Республики Беларусь
по образованию в области информатики и радиоэлектроники
в качестве учебно-методического пособия для студентов
учреждений, обеспечивающих получение высшего образования
по специальности «Вычислительные машины, системы и сети»*

Минск БГУИР 2011

УДК 004.056.55:004.7(076)
ББК 32.973.26-018.2я73
К55

Р е ц е н з е н т ы:

заведующий кафедрой «Программное обеспечение вычислительной техники и автоматизированных систем» учреждения образования «Белорусский национальный технический университет», кандидат технических наук Н. Н. Гурский;

ведущий научный сотрудник Объединенного института проблем информатики Национальной академии наук Беларуси, кандидат технических наук А. А. Несенчук

Кобяк, И. П.

К55 Защита информации в вычислительных сетях : учеб.-метод. пособие / И. П. Кобяк. – Минск : БГУИР, 2011. – 86 с. : ил.

ISBN 978-985-488-796-8.

Пособие предназначено для изучения программных и аппаратных методов шифрования важной и секретной информации. Данный учебный материал позволяет получить представление о системах шумоподобной связи, алгоритмах криптографии и математических основах наблюдения случайных процессов.

УДК 004.056.55:004.7(076)
ББК 32.973.26-018.2я73

ISBN 978-985-488-796-8

Содержание

ВВЕДЕНИЕ

До настоящего времени вопросы защиты информации в системах хранения и передачи данных рассмотрены недостаточно полно. Данное обстоятельство обусловлено, с одной стороны, высокой востребованностью методов криптографии компьютерами специального назначения, с другой стороны, стремлением шифровальщиков скрыть используемые алгоритмы и аппаратные средства от хакеров и потенциальных противников. Соответственно, и представляемая в печать информация не является пределом математической теории или теории кодирования и носит выраженный учебный характер. Предлагаемые учебные материалы по своей сути являются базовыми с точки зрения прикладных теорий и определяют систему взглядов студентов в задачах защиты, а в некоторых приложениях и сокрытия информации.

Теоретические направления, определяющие перспективу проектирования современных средств защиты важных или секретных данных, включают в себя следующие направления: 1) алгоритмы и аппаратуру формирования шумоподобных каналов связи; 2) математические методы криптографии над конечными полями; 3) методы инъективного отображения знаков открытого сообщения в меньшее подмножество знаков шифротекста; 4) алгоритмы наблюдения случайных процессов, содержащих элементы детерминизма. Все эти направления тесно связаны между собой и, как правило, используются комплексно. Однако и каждое из указанных направлений в криптографии имеет право на самостоятельное развитие.

Методы построения защищенных каналов с шумоподобным представлением данных базируются на достижениях современной теории вероятностей, математической статистики, а также на результатах разделов математики, связанных теорией отображения множеств. В соответствии с этим основные базовые характеристики и подходы к формированию пересылаемой шифрованной информации рассмотрены в первом разделе. Здесь же даются ознакомительные материалы по формированию чисел со случайной и псевдослучайной природой, рассмотрены вопросы регулирования вероятностей и управления вероятностными характеристиками каналов связи в процессе передачи информации.

Во втором разделе пособия приведены алгоритм решения сравнений первой степени с одним неизвестным, метод шифрования данных без передачи ключей, алгоритм шифрования RSA, алгоритмы формирования криптографической подписи и т.д. Рассмотренный в данном разделе основополагающий принцип построения криптосистем на базе вычислительных методов над конечными полями базируется на теореме Ферма. При этом процесс проектирования шифросистемы является достаточно простым, однако, заметим, и декодирование таких данных с помощью компьютеров не представляет особых затруднений. Во втором разделе также рассмотрены математическая модель шифра замены, принципы отображения множеств, некоторые аппаратно-программные средства, ориентированные на реализацию симметричных шифров.

Одним из важнейших направлений в задачах практической криптографии является задача наблюдения каналов связи вероятного противника и, соответственно, обнаружения и декодирования его зашумленной информации. Данному вопросу посвящен третий раздел учебно-методического пособия, где изучаются принципы применения АКФ для анализа лингвистических конструкций различных языков. Показано, что степень зависимости между заданными в условиях задачи подмножествами дискретных аргументов имеет существенное отклонение от асимптотического параметра, если шифрованное сообщение содержит элементы искусственного детерминизма.

Четвертый раздел пособия определен как отдельное направление задачи математического синтеза кодов инъективного отображения двоичных последовательностей. В данном разделе приводятся преобразования, позволяющие установить соответствие между матрицей переходов системы, формируемой на основании конституент таблицы истинности, и структурой цифрового устройства, реализующего обратное инъективное отображение. Выполненный анализ векторно-матричных преобразований позволил сформулировать условия линеаризации формируемого устройства, что эквивалентно сокращению затрат аппаратуры на развёртку состояний цепи Маркова. Если сформулированные условия выполняются, то процесс линеаризации считается возможным. В случае невыполнения хотя бы одного из условий рассматривается возможность перехода к алгоритму формирования нового шифротекста путем выравнивания вероятностей с помощью линейных рекуррентных последовательностей, формируемых с помощью полином над $GF(2)$.

Ограничение методических материалов по объему не позволило, однако, рассмотреть ряд вопросов, связанных с передачей важных или секретных данных в информационных сетях. В частности, важными моментами в системах криптографии являются вопросы синтеза хеш-функций, методы скремблирования зашумленных последовательностей, алгоритмы управления ключами и другие технические и организационные подходы, направленные на укрепление каналов сети. Следует также знать, что современные системы передачи информации могут быть построены с использованием идей многоканальной связи, принципов лазерного генерирования, а также квантово-криптографического протоколирования интерфейсов в оптических системах с заданными параметрами и квантов. Однако данные направления соответствуют другим специальностям университета и требуют усиленной специальной подготовки по общей и квантовой физике, информатике, а также наличия навыков решения задач с исходными данными, представленными комплексными числами.

Вместе с тем материалы предлагаемого учебно-методического пособия определяют концепции подготовки специалистов в области общей математики, теории чисел и теории вероятностей (по специальности 1-40 02 01), что само по себе является положительным фактором в обучении будущих специалистов. Основополагающие знания в области вычислительной техники при этом подразумеваются и, очевидно, должны соответствовать вузовской подготовке.

1. ОСНОВЫ ФОРМИРОВАНИЯ ШУМОПОДОБНЫХ СИГНАЛОВ

1.1. Понятие случайных сигналов

Основной принцип формирования сигналов со случайной природой заключается в регистрации отклонений параметров некоторого константного процесса под воздействием внешних факторов. Физическая сущность этих отклонений (или флуктуаций) может базироваться на дискретном характере зарядов, создающих электрический ток, так называемый дробовой эффект, в тепловом движении этих зарядов, которое получило название теплового шума, и в изменениях проводимости приборов под воздействием случайных факторов, или так называемый модуляционный шум.

Природу дробового шума определяет дискретность и «орбитальность» заряда электрона, что порождает случайное число носителей в системе при поглощении электронами энергии, эквивалентной работе выхода. Таким образом, ток, образованный в реальных условиях, не является строго постоянным, а флуктуирует относительно некоторого среднего значения. Следовательно, полное значение тока через электронный прибор может быть представлено в виде суммы:

$$i(t) = I_{const} \pm i_{\text{var}}. \qquad (1.1)$$

Удаляя аппаратными способами константную составляющую из соотношения (1.1) и усиливая переменную, можно получить шумоподобный сигнал с некоторыми базовыми свойствами, которые в дальнейшем преобразуются к требуемым величинам. Принцип удаления I_{const} реализуется с помощью чувствительных пороговых элементов, настроенных на заданное значение постоянной составляющей в (1.1).

Второй вид шумов – это тепловой шум, который возникает вследствие хаотического движения электронов в некоторой нагретой среде. Поэтому электрический ток может регистрироваться в любом электронном приборе даже без подключения последнего к внешним источникам питания. Однако уровень этого тока близок к нулю, а частота флуктуаций весьма велика. Использование таких источников затруднено необходимостью применения прецизионных усилителей большой мощности, имеющих в общем случае высокую стоимость.

Наиболее перспективными для применения в системах защиты информации являются полупроводниковые приборы, шумовой сигнал в которых определяется комплексом факторов. Основной источник шума в таком устройстве это шумы в области пробоя *p*-*n*-перехода на обратной ветви вольт-амперной характеристики. Их появление объясняется возникновением неоднородностей проводимостей локальных участков в электронном приборе из-за погрешностей в технологических процессах изготовления.

Все рассмотренные методы позволяют сформировать шумоподобные сигналы, подпадающие под определение случайных процессов с непрерывным

временем. Применение их в реальных системах затруднено из-за малой помехозащищенности и сложности построения систем передачи данных.

Для создания сигналов, пригодных для защищенной цифровой связи по информационным сетям, используется принцип дискретизации процессов с непрерывным временем. Базовым элементом при этом является пороговый усилитель, позволяющий сформировать двоичные непозиционные коды, преобразуемые далее в двоичные векторы упакованного формата (8421).

1.2. Характеристики случайного процесса

В большинстве приложений для описания случайного процесса достаточно знать два основных момента – это математическое ожидание и автокорреляционная функция. Эти характеристики являются неслучайными функциями и представляют собой результат вероятностного усреднения множества наблюдений реализаций случайного процесса.

Если зафиксировать значения аргументов в момент времени t_j, то отсчеты реализаций случайного процесса $\xi\left(t_j\right)$ будут представлять собой случайную величину ξ и математическое ожидание для n выборок может быть определено по формуле

$$M\left[\xi\left(t_j\right)\right]=\frac{1}{n}\sum_{\lambda=1}^{n}\xi_\lambda\left(t_j\right),\quad j=1,2,3,\ldots,\infty,\quad n\to\infty, \tag{1.2}$$

где λ – номер реализации.

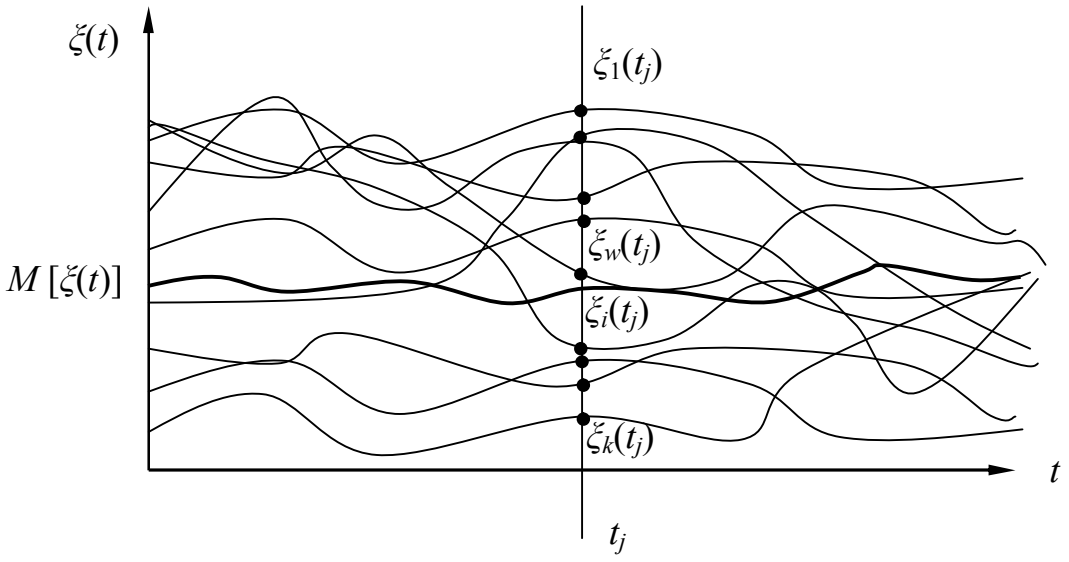

Рис. 1.1

Как правило, для истинно случайного процесса (рис. 1.1) кривая матожидания (1.2) с увеличением числа выборок стремится к прямой линии.

При сравнении характеристик двух процессов с одинаковыми средними

значениями часто используют функцию дисперсии. Эта функция характеризует рассеяние отсчетов случайной величины вблизи математического ожидания и, например, для одной реализации процесса с номером λ определяется формулой

$$D_\lambda\left[\xi(t)\right]=\frac{1}{j_{\max}}\sum_{j=0}^{j_{\max}}\left[\xi_\lambda\left(t_j\right)-M_\xi\right]^2, \tag{1.3}$$

где $M\left[\xi\left(t_j\right)\right]=M_\xi, \quad j_{\max}\to\infty.$

Для всех λ дисперсия (1.3) будет равна

$$D\left[\xi(t)\right]=D_\xi=\frac{1}{n}\frac{1}{j_{\max}}\sum_{\lambda=1}^{n}\sum_{j=0}^{j_{\max}}\left[\xi_\lambda\left(t_j\right)-M_\xi\right]^2.$$

Величина

$$\sigma\left[\xi(t)\right]=\sqrt{D\left[\xi(t)\right]}, \quad \sigma_\xi=\sqrt{D_\xi}, \quad n,j\to\infty$$

называется среднеквадратичным отклонением (с.к.о.) и так же, как и дисперсия, характеризует рассеяние отсчетов случайного процесса вблизи значения математического ожидания.

Однако одномерные характеристики M_ξ и D_ξ не являются достаточными для оценки протекания случайного процесса во времени. Очень часто в системах передачи информации важно знать характер и силу связи между значениями процесса в различные моменты времени. Для этого используется функция математического ожидания произведений значений центрированной случайной величины, взятой при двух моментах времени: t_1 и t_2. Эту функцию называют корреляционной или автокорреляционной функцией (АКФ) процесса $\xi(t)$ и для некоторой λ-реализации процесса рассчитывают по формуле

$$K_{\lambda,\xi}\left(t_1,t_2\right)=M\left[\overset{\circ}{\xi}_\lambda\left(t_1\right)\cdot\overset{\circ}{\xi}_\lambda\left(t_2\right)\right]=\lim_{j_{\max}\to\infty}\frac{1}{j_{\max}}\sum_{j=0}^{j_{\max}}\left[\xi_\lambda\left(t_j\right)-M_\xi\right]\left[\xi_\lambda\left(t_{j+\tau}\right)-M_\xi\right]. \tag{1.4}$$

При условии, что $\tau=0$, из (1.4) имеем

$$K_{\lambda,\xi}\left(t_1,t_2\right)=M\left[\xi_\lambda\left(t_j\right)-M_\xi\right]^2=D_\lambda\left[\xi(t)\right],$$

т. е. автокорреляционная функция одного и того же сечения равна дисперсии случайного процесса.

Соответственно, для всех значений λ и неограниченно растущего числа реализаций случайного процесса АКФ усредняется по правилу

$$K_\xi\left(t_1,t_2\right)=M\left[\overset{\circ}{\xi}\left(t_1\right)\cdot\overset{\circ}{\xi}\left(t_2\right)\right]=\lim_{n\to\infty}\frac{1}{n}\sum_{\lambda=1}^{n}K_{\lambda,\xi}\left(t_1,t_2\right).$$

Для определения статистической зависимости между отсчетами различных случайных процессов используется понятие взаимной корреляционной функции (ВКФ):

$$K_{\xi,\eta}\left(t_1,t_2\right)=M\left[\overset{\circ}{\xi}\left(t_1\right)\cdot\overset{\circ}{\eta}\left(t_2\right)\right]. \tag{1.5}$$

При этом процессы называются коррелированными, если их ВКФ (1.5) не равна нулю, и независимыми – в противном случае.

Для характеристики связи между функциями $\eta(t),\xi(t)$ часто переходят от функции $K_{\xi,\eta}(t_1,t_2)$ к безразмерной характеристике:

$$r_{\xi,\eta}\left(t_1,t_2\right)=\frac{K_{\xi,\eta}\left(t_1,t_2\right)}{\sigma_\xi\,\sigma_\eta}.$$

Данный параметр получил название нормированной ВКФ, или коэффициента корреляции. Удобство использования параметра состоит в том, что теоретически его значения лежат в пределах $-1\le r_{\xi,\eta}\left(t_1,t_2\right)\le 1$.

1.3. Понятие стационарных случайных процессов и цепей Маркова

Случайные процессы, статистические характеристики n выборок которых не зависят от начального момента времени t, называют стационарными. Примером стационарного процесса может служить процесс, приведенный на рис. 1.1. На рис. 1.2 приведен пример нестационарного случайного процесса. При этом характерной особенностью отсчетов элементарных событий является увеличение их абсолютной величины с увеличением времени t.

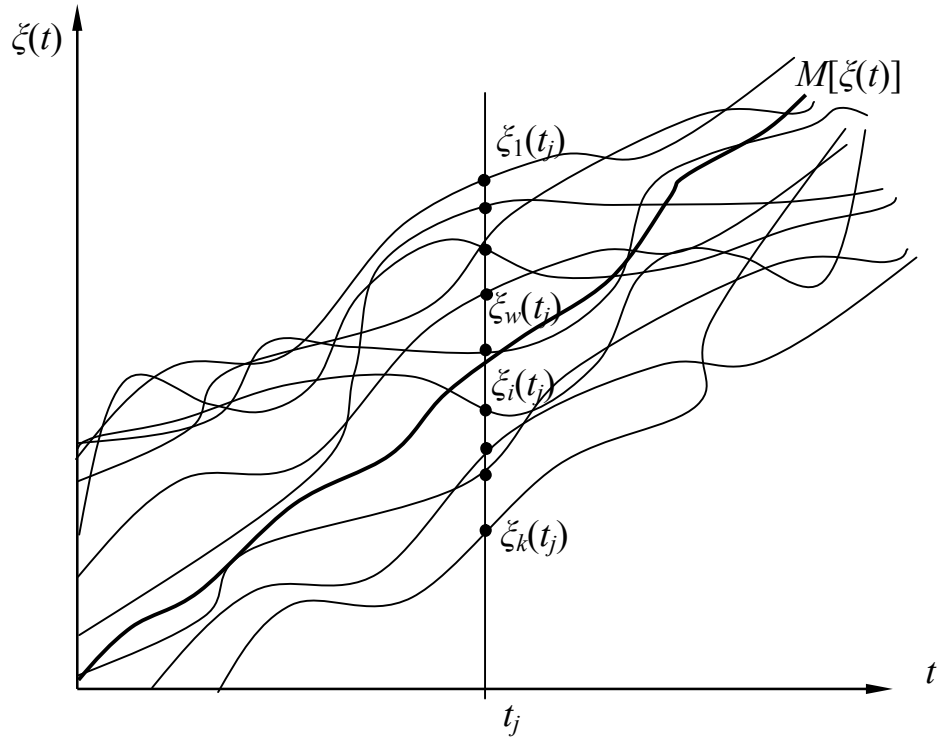

Рис. 1.2

Различают стационарные случайные процессы в узком и широком смысле слова. Так, под *стационарным процессом в узком смысле* слова обычно понимают гармонический процесс со случайно меняющейся частотой и амплитудой. При этом для любого начального момента t распределения вероятностей случайных событий должны совпадать:

$$P\left(\xi_{1,t},\xi_{2,t},\ldots,\xi_{m,t}\right)=P\left(\xi_{1,t+\tau},\xi_{2,t+\tau},\ldots,\xi_{m,t+\tau}\right), \qquad (1.6)$$

иными словами, для процессов, имеющих истинно случайную природу, функция (1.6) имеет равномерное распределение.

Замечание. Плотность (насколько часто) распределения вероятностей случайных событий считается заданной, если известны все значения, принимаемые случайной величиной, и все вероятности, соответствующие элементарным событиям.

Под стационарным случайным процессом *в широком смысле* слова понимают процесс, математическое ожидание которого постоянно во времени, а корреляционная функция $K_\xi(t_1,t_2)$ зависит только от разности $\tau=t_2-t_1$ и обозначается $K_\xi(\tau)$. Дисперсия такого процесса должна удовлетворять равенству

$$D_\xi=K_\xi(\tau=0)=const.$$

Марковскими цепями называют случайные процессы, которые имеют *конечное число возможных состояний*, причем переход из одного состояния в другое зависит только от предшествующего. Марковские процессы свойственны многим устройствам генерирования последовательностей, в связи с чем имеют широкий спектр теоретического и практического применения.

Пусть $p_i(t)$ есть вероятность состояния процесса $A_i(t)$. Тогда совокупность вероятностей для всех i может быть представлена графом с числом вершин, равным числу состояний системы. При этом

$$0\le p_i(t)\le 1, \qquad \sum_i p_i(t)=1.$$

Так как переход из состояния $A_i(t)$ в состояние $A_j(t+1)$ зависит только от уровня вероятностной связи этих двух событий, то каждой такой паре процесса можно поставить в соответствие условную вероятность $p_{i,j}$. Данный параметр указывает, с какой вероятностью система перейдет в состояние A_j в момент времени $t+1$ при условии, что в момент времени t она находилась в состоянии A_i. В данном случае совокупность вероятностей $p_{i,j}$ образует квадратную матрицу, сумма элементов каждой строки которой равна единице. Эта матрица получила название стохастической, а вероятности $p_{i,j}$ – вероятностей перехода. Соответственно, умножение строки на столбцы дает произведение

$$p_j(t+1) = \sum_i p_i(t)\, p_{i,j}, \quad j = 0,1,2,\ldots,$$

или, в матричной форме

$$\mathbf{p}(t+1) = \mathbf{p}(t) \cdot \left\| \mathbf{p}_{i,j} \right\|.$$

Таким образом, марковская цепь полностью определяется стохастической матрицей $\left\| \mathbf{p}_{i,j} \right\|$ и совокупностью вероятностей начальных состояний $p_i(0)$.

Пример. Пусть задан автомат с матрицей переходов и табл.1.1 вероятностей начальных состояний.

$$\left\| \mathbf{p}_{i,j} \right\| = \begin{array}{c|ccc|l} & A_1 & A_2 & A_3 & \leftarrow i \\ A_1 & 0,75 & 0 & 0,25 & \\ A_2 & 0,25 & 0,25 & 0,5 & j\downarrow \\ A_3 & 0,5 & 0,5 & 0 & \end{array}.$$

Таблица 1.1

A_j	$A_1(0)$	$A_2(0)$	$A_3(0)$
$p_j(0)$	0	1	0

Граф переходов такого автомата имеет следующий вид (рис. 1.3)

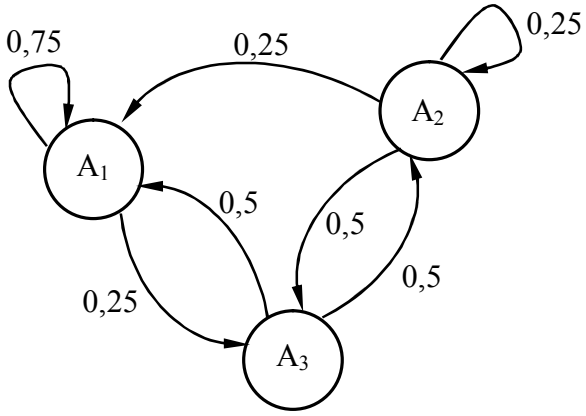

Рис. 1.3

1.4. Формирование последовательностей со случайной природой

Если время t меняется дискретно, то говорят о случайных процессах с дискретным временем или о последовательностях случайных событий. Одним из примеров случайного процесса является последовательность испытаний Бернулли.

Определение. Говорят, что повторные независимые испытания образуют схему Бернулли, если в каждом из них имеется только два возможных исхода (например 0 или 1), а вероятности этих исходов $p(1) = p$ и $p(0) = q$ остаются

неизменными для всего процесса из n испытаний. При этом сумма вероятностей $p+q=1$, а вероятность конкретной реализации процесса равна произведению вероятностей, полученных путем замены в данной выборке 0 и 1 на соответствующие *p*-, *q*-параметры.

Пример. 1 0 1 0 0 1 1 1 0 0 1= p^6q^5.

В большинстве практических приложений существенный интерес представляет суммарное число единиц k, выпавших в последовательности из n испытаний независимо от порядка их следования. Если значение k рассмотреть как некоторую случайную величину, то ее распределение может быть задано функцией

$$P_{k,n}=C_n^k p^k q^{n-k}, \qquad (1.7)$$

где C_n^k – мощность подмножества реализаций с k единицами из n, $p^k q^{n-k}$ – нормирующий множитель, приводящий параметр (1.7) к интервалу «0–1». Приведенное соотношение получило название функции биномиального распределения и может быть представлено в виде треугольника Паскаля. Если процесс испытаний более сложный, то есть число элементарных событий больше двух, то график биномиального распределения одной переменной имеет следующий вид (рис. 1.4).

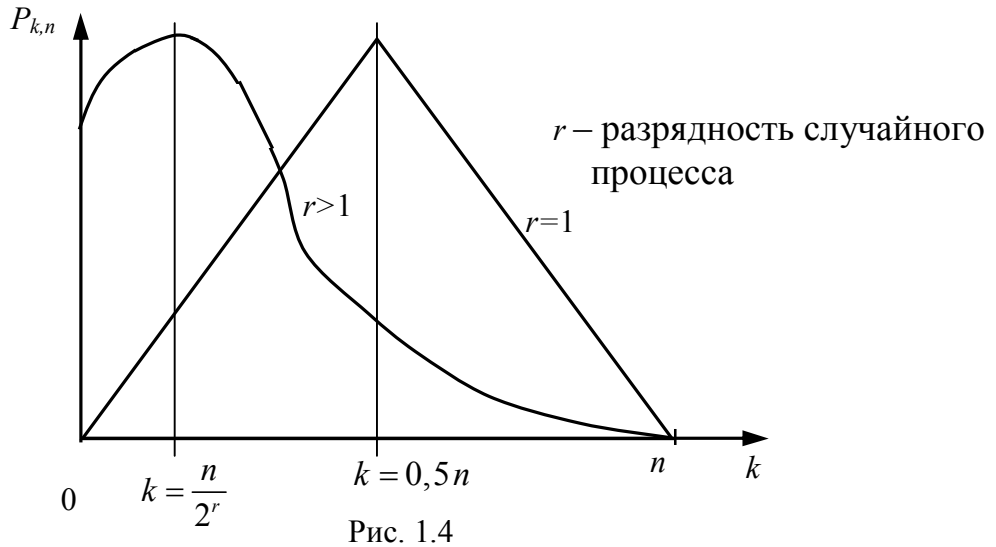

Рис. 1.4

Для кодирования информации с помощью шумоподобных сигналов в системах защиты используют первичные датчики или генераторы случайных чисел (ГСЧ), а также генераторы псевдослучайных чисел (ГПСЧ).

Получение случайных двоичных чисел в ГСЧ с физической основой связано с преобразованием непрерывного «диодного» шума в последовательность случайных импульсов с вероятностью «единицы», равной $p(1)=0,5$. Функцию дискретизации при этом выполняет пороговый элемент, а отсчет чисел осуществляется в заданные моменты времени t. Схема простейшего ГСЧ имеет вид, показанный на рис. 1.5.

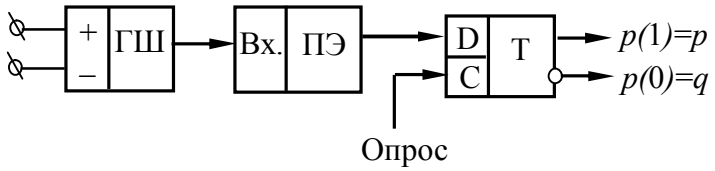

Рис. 1.5

Как правило, применение ГСЧ связано с решением ряда задач, требующих поддержания на высоком уровне вероятностных характеристик формируемых последовательностей. Последовательности же случайных символов, получаемые от реальных датчиков, имеют свойства, далекие от идеальных. Поэтому для выравнивания вероятностей нуля и единицы используют итерационные методы, реализуемые схемами с элементами задержки.

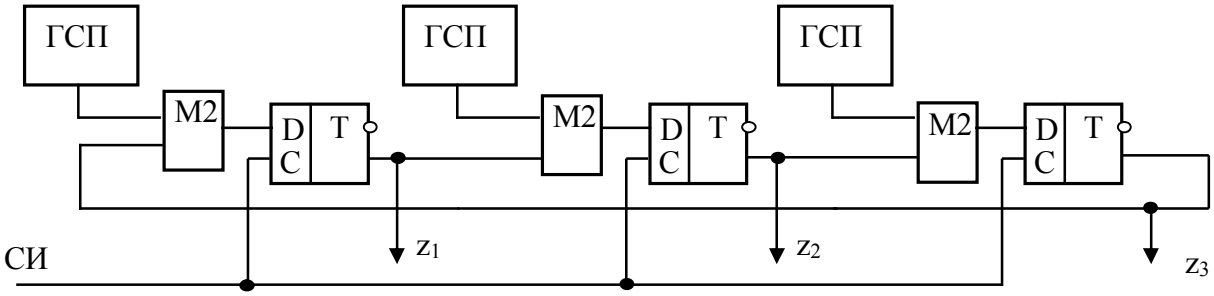

Рис. 1.6

Более простыми в изготовлении и надежными с точки зрения вероятностных характеристик являются генераторы псевдослучайных чисел, построенные полностью из элементов двоичной логики. В таких устройствах числа формируются с использованием регистра сдвига с обратной связью, коэффициенты которой определяются специальными таблицами.

Пример. $\varphi(x)=1+\alpha_1 x^1+\alpha_2 x^2+\ldots+\alpha_l x^l$, где $l-$ разрядность регистра сдвига. В частном случае многочлен $\varphi(x)=1+x^1+x^2+x^4+x^5$ порождает схему, показанную на рис. 1.7.

Формируемые последовательности получили название псевдослучайных, потому что, несмотря на детерминированную структуру, они обладают всеми основными признаками реализаций стационарных случайных процессов.

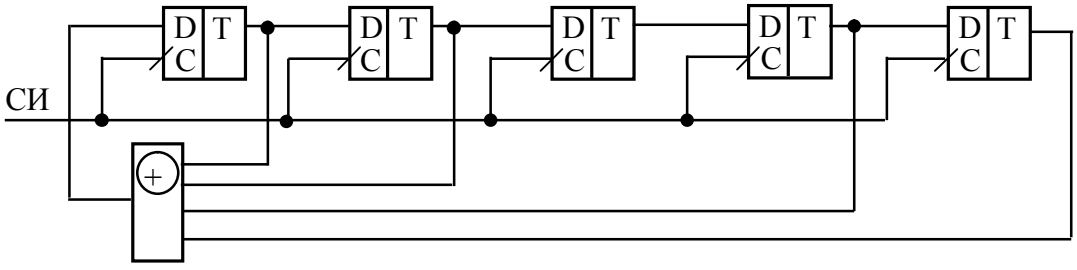

Рис. 1.7

Рассмотрим свойства последовательностей генераторов псевдослучайных чисел. Пусть l – разрядность генератора, тогда $M = 2^l - 1$ это период формируемых чисел, причем полный период двоичной *M*-последовательности обладает следующими свойствами.

1. Число единичных символов в *M*-последовательности всегда на единицу больше, чем нулевых.

2. Серии следующих друг за другом одинаковых символов 0 или 1, 00 или 11, 000 или 111 и т.д. формируются с такой же частотой, что и в случайной последовательности.

3. Любой двоичный набор из $i \le l - 2$ смежных символов (нулей или единиц) встречается в *M*-последовательности ровно 2^{l-2-i} раз. Наборы из $l-1$ и l элементарных событий – по одному разу.

4. Нормированная АКФ *M*-последовательности

$$r_\xi\left[t_2 - t_1 = \tau\right] = \begin{cases} 1 \ при \ \tau = 0, \\ -\dfrac{1}{M} \ при \ \tau \neq 0. \end{cases}$$

Этот показатель эквивалентен АКФ «белого шума» (или истинно случайной последовательности), у которого при больших значениях выборки n для всех $\tau \neq kn$ имеет место равенство $r_\xi[\tau] = 0$, то есть АКФ у *M*-последовательности практически идеальна, т. к. величина $\frac{1}{M} \approx 0$.

Для получения в разрядах ГПСЧ последовательности с максимально возможным периодом, необходимо, чтобы характеристический полином $\varphi(x)$, определяющий структуру цепи обратной связи, был примитивен. Как правило, требуемый многочлен выбирается из специальных таблиц, причем из множества многочленов одинаковой степени выбирают такие, которые позволяют получить наиболее простую структуру. Этому условию удовлетворяют полиномы вида $\varphi(x) = 1 + x^j + x^l$, то есть трехчлены, предполагающие суммирование сигналов только с двух выходов регистра.

Для описания процесса функционирования ГПСЧ часто используют векторно-матричную запись:

$$\begin{vmatrix} x_1^{t+1} \\ x_2^{t+1} \\ x_3^{t+1} \\ \dots \\ x_l^{t+1} \end{vmatrix} = \begin{vmatrix} x_1^{t} \\ x_2^{t} \\ x_3^{t} \\ \dots \\ x_l^{t} \end{vmatrix} \cdot \begin{vmatrix} \alpha_1 & \alpha_2 & \dots & \alpha_{l-1} & \alpha_l \\ 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ . & . & . & . & . \\ 0 & 0 & \dots & 1 & 0 \end{vmatrix}.$$

Очевидно, что $\boldsymbol{X}^{t+1} = \boldsymbol{X}^t \left|\boldsymbol{\alpha}\right|$, кроме того, в данном случае выполняется сочетательный закон и $\boldsymbol{X}^{t+1} = \boldsymbol{X}^1 \left|\boldsymbol{\alpha}\right|^t$.

1.5. Методы регулирования вероятностей. Логические элементы и их свойства как регуляторов вероятности

В большинстве случаев методы синтеза последовательностей со случайной природой предполагают формирование элементарных событий с равномерным законом распределения. Однако даже эталонные генераторы для различных длин выборки n не обеспечивают требуемых свойств, что не позволяет в общем случае сформировать последовательности с произвольно заданной вероятностью событий. В связи с этим будем рассматривать принципы преобразования вероятностных последовательностей с помощью логических элементов с целью определения формальных подходов к задаче управления вероятностью и декорреляции последовательностей.

1.5.1. Функция инверсии

Инверсия $z = \overline{x}$ функционально сводится к замене событий x дополнительными событиями $\overline{x}$, математическое ожидание которых определяется равенством

$$M_z = M_{\overline{x}} = 1 - p\left[x(t)\right] = 1 - M_x.$$

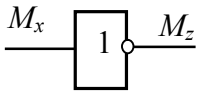

Рис. 1.8

По определению автокорреляционная функция выходных сигналов представляет собой равенство

$$K_z(\tau) = M\left[z(t) - M_z\right]\left[z(t+\tau) - M_z\right],$$

где $t, \tau = 1, 2, 3, \ldots$ – дискретное время. Тогда, учитывая, что $M_x = M_z = 0{,}5$, имеем

$$K_z(\tau) = M\left[z(t)z(t+\tau) - z(t)M_z - z(t+\tau)M_z + M_z^2\right] =$$
$$= M\left[z(t)z(t+\tau) - 0{,}5\left[z(t) + z(t+\tau)\right] + 0{,}25\right].$$

Рассмотрим составляющую в функции $K_z(\tau)$, равную $-0{,}5\left[z(t) + z(t+\tau)\right]$:

$$M\left[-0{,}5\left[z(t) + z(t+\tau)\right]\right] = -0{,}5(0{,}5 + 0{,}5) = -0{,}5.$$

Тогда

$$K_z(\tau) = M\left[z(t)z(t+\tau)\right] - 0{,}5 + M\left[0{,}25\right] = M\left[z(t)z(t+\tau)\right] - 0{,}25.$$

Матожидание и вероятность двоичных переменных численно равны. Поэтому

$$K_z(\tau) = p\left[z(t)z(t+\tau)\right] - 0{,}25 = p\left[\overline{x(t)}\ \overline{x(t+\tau)}\right] - 0{,}25 =$$
$$= p\left[\overline{x(t) \vee x(t+\tau)}\right] - 0{,}25 = 1 - p\left[x(t) \vee x(t+\tau)\right] - 0{,}25. \qquad (1.8)$$

Для статистически связанных событий $x(t)$ и $x(t+\tau)$ имеет место равенство

$$\begin{aligned} p\left[x(t)\vee x(t+\tau)\right] &= p\left[x(t)\right]+p\left[x(t+\tau)\right]-p\left[x(t)x(t+\tau)\right]= \\ &= p\left[x(t)\right]+p\left[x(t+\tau)\right]-p\left[x(t)\right]p\left[x(t+\tau)\right]-K_x(\tau). \end{aligned} \tag{1.9}$$

Учитывая, что $p\left[x(t)\right]=p\left[x(t+\tau)\right]=M_x$, а также равенство (1.8), имеем

$$p\left[x(t)\vee x(t+\tau)\right]=M_x+M_x-M_xM_x-K_x(\tau)=0,75-K_x(\tau).$$

Отсюда на основании (1.9) следует

$$K_x(\tau)=1-0,75+K_x(\tau)-0,25=K_x(\tau).$$

Таким образом, при инвертировании последовательности происходит вычитание вероятности двоичной переменной из единицы, а корреляционная функция не изменяется:

$$\begin{cases} M_z=1-M_x, \\ K_z(\tau)=K_x(\tau), \ \ z=\overline{x}. \end{cases}$$

1.5.2. Функция конъюнкции

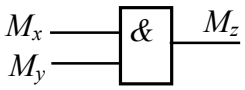

Рис. 1.9

Математическое ожидание выходной последовательности для данной функции будет определяться формулой

$$M_z=p\left[x(t)y(t)\right]=p\left[x(t)\right]p\left[y(t)\right]+K_z(\tau),$$

где $K_z(\tau)$ – взаимная корреляционная функция выходной последовательности при $\tau=0$.

Запишем АКФ $K_z(\tau)$ в стандартной форме:

$$\begin{aligned} K_z(\tau) &= M\left[x(t)-M_x\right]\left[y(t)-M_y\right]=M\left[x(t)y(t)-M_xy(t)-M_yx(t)+M_xM_y\right]= \\ &= M\left[x(t)y(t)\right]-M\left[M_xy(t)\right]-M\left[M_yx(t)\right]+M_xM_y. \end{aligned}$$

В данном соотношении:

$$M\left[M_xy(t)\right]=M_xM\left[y(t)\right]=M_xM_y.$$

Следовательно:

$$K_z(\tau)=M\left[x(t)y(t)\right]-M_xM_y-M_xM_y+M_xM_y=M\left[x(t)y(t)\right]-M_xM_y.$$

Пусть $M_x=M_y=0,5$, тогда

$$K_z(\tau)=M\left[x(t)y(t)\right]-0,25. \tag{1.10}$$

Значение

$$M\left[x(t)y(t)\right]=p\left[x(t)y(t)\right]=p\left[x(t)\right]\cdot p\left[y(t)\right]+K_{x,y}=M\left[x(t)\right]M\left[y(t)\right]+K_{x,y}(\tau).$$

Тогда из (1.10) следует

$$K_z(\tau)=0,25+K_{x,y}-0,25=K_{x,y}(\tau).$$

Таким образом,

$$\begin{cases}M_z=M_xM_y+K_z(\tau),\\ K_z(\tau)=K_{x,y}(\tau).\end{cases}$$

Иными словами, схема конъюнкции не обладает декоррелирующими свойствами, а лишь преобразует ВКФ последовательностей $x(t)$ и $y(t)$ в АКФ выходной последовательности $z(t)$.

1.5.3. Функция дизъюнкции

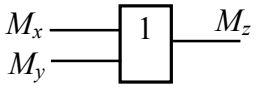

Рис. 1.10

Используя формулу объединения вероятностей событий, можно записать

$$M_z=p\left[x(t)\vee y(t)\right]=p\left[x(t)\right]+p\left[y(t)\right]-p\left[x(t)y(t)\right]=$$
$$=M_x+M_y-M_xM_y-K_z=0,75-K_z(\tau).$$

Для вычисления значения $K_z(\tau)$ выполним преобразование функции дизъюнкции в конъюнкцию, используя правило де Моргана:

$$z(t)=x(t)\vee y(t)=\overline{\overline{x(t)}\,\overline{y(t)}}.$$

Учитывая формулу (1.10) и тот факт, что процесс инвертирования не изменяет значения АКФ, запишем равенство

$$K_z(\tau)=M\left[\overline{\overline{x(t)}\,\overline{y(t)}}\right]-0,25=M\left[\overline{x(t)}\,\overline{y(t)}\right]-0,25=\left(1-M_x\right)\left(1-M_y\right)+K_{\bar{x},\bar{y}}(\tau)-0,25.$$

$$K_{\bar{x},\bar{y}}(\tau)=K_{x,y}(\tau).$$

$$K_z(\tau)=\left(1-M_x-M_y+M_xM_y\right)+K_{x,y}(\tau)-0,25=M_xM_y+K_{x,y}(\tau)-0,25=K_{x,y}(\tau).$$

Итак, операция дизъюнкции также не изменяет уровня корреляции, а преобразует функцию ВКФ_{xy} в АКФ_z выходной последовательности логического элемента:

$$\begin{cases}M_z=M_x+M_y-M_xM_y-K_z(\tau),\\ K_z(\tau)=K_{x,y}(\tau).\end{cases}$$

1.5.4. Операция суммирования по модулю два

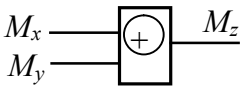

Рис. 1.11

Функцию суммирования по модулю два представим в следующем виде:

$$z(t) = x(t) \oplus y(t) = x(t)\overline{y(t)} \vee \overline{x(t)}y(t).$$

В данном случае логическая функция предполагает выбор одного из двух несовместимых исходов. Матожидание выходной последовательности при этом будет определяться следующей зависимостью:

$$M_z = p\left[x(t)\overline{y(t)} \vee \overline{x(t)}y(t)\right] = p\left[x(t)\overline{y(t)}\right] + p\left[\overline{x(t)}y(t)\right]. \qquad (1.11)$$

Преобразуем соотношение (1.11) с использованием формулы полной вероятности:

$$p\left[x(t)\right] = p\left[x(t)y(t)\right] + p\left[x(t)\overline{y(t)}\right], \quad p\left[x(t)\overline{y(t)}\right] = p\left[x(t)\right] - p\left[x(t)y(t)\right],$$
$$p\left[y(t)\right] = p\left[x(t)y(t)\right] + p\left[\overline{x(t)}y(t)\right], \quad p\left[\overline{x(t)}y(t)\right] = p\left[y(t)\right] - p\left[x(t)y(t)\right].$$

Тогда

$$M_z = p\left[x(t)\right] + p\left[y(t)\right] - 2p\left[x(t)y(t)\right] =$$
$$= p\left[x(t)\right] + p\left[y(t)\right] - 2p\left[x(t)\right]p\left[y(t)\right] - 2K_z(\tau).$$

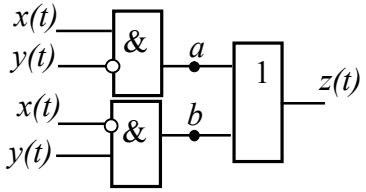

Рис. 1.12

Автокорреляционную функцию выходной последовательности получим, выполнив следующее преобразование схемы суммирования mod2:

$$M_a = M_x M_y^- + K_{x,y}^-(\tau), \quad M_b = M_x^- M_y + K_{x,y}^-(\tau),$$
$$K_{x,y}^-(\tau) = K_{x,y}(\tau), \quad K_{x,y}^-(\tau) = K_{x,y}(\tau). \qquad (1.12)$$
$$M_z = M_a + M_b - M_a M_b - K_{a,b}(\tau).$$

Подставим в M_z (1.12) соотношения для M_a и M_b:

$$M_z = M_x M_y^- + K_{x,y}(\tau) + M_x^- M_y + K_{x,y}(\tau) -$$
$$-\left[M_x M_y^- + K_{x,y}(\tau)\right]\left[M_x^- M_y + K_{x,y}(\tau)\right] - K_{a,b}(\tau).$$

Учитывая, что $M_{x^-} = 1 - M_x$, имеем

$$M_z = M_x\left(1 - M_y\right) + K_{x,y}(\tau) + \left(1 - M_x\right)M_y + K_{x,y}(\tau) -$$
$$-\left[M_x\left(1 - M_y\right) + K_{x,y}(\tau)\right]\left[\left(1 - M_x\right)M_y + K_{x,y}(\tau)\right] - K_{a,b}(\tau) = \qquad (1.13)$$

$$= M_x + M_y - 2M_xM_y + 2K_{x,y}(\tau) -$$

$$-(1-M_x)(1-M_y)M_xM_y - \left[M_x(1-M_y) + (1-M_x)M_y\right]K_{x,y}(\tau) - K^2_{x,y}(\tau) - K_{a,b}(\tau).$$

Рассмотрим значение АКФ K_{ab} с учетом стандартной формы:

$$K_{a,b}(\tau) = M\left[(a-M_a)(b-M_b)\right] = M(ab) - M_bM(a) - M_aM(b) + M_aM_b. \quad (1.14)$$

События a и b несовместимы, поэтому $M(ab) = 0$, а из (1.14) имеем

$$K_{a,b}(\tau) = -M_bM_a - M_aM_b + M_aM_b = -M_aM_b. \quad (1.15)$$

Подставим в равенство (1.15) значения из соотношения (1.12), тогда получаем

$$K_{ab} = -\left[M_xM_{\bar{y}} + K_{x,y}(\tau)\right]\left[M_{\bar{x}}M_y + K_{x,y}(\tau)\right] = -\left[M_x(1-M_y) + K_{x,y}(\tau)\right]\left[(1-M_x)M_y + K_{x,y}(\tau)\right] = \quad (1.16)$$

$$= -(1-M_x)(1-M_y)M_xM_y - \left[M_x(1-M_y) + (1-M_x)M_y\right]K_{x,y}(\tau) - K^2_{x,y}(\tau).$$

Подставим теперь в соотношение (1.16) равенство (1.13) для математического ожидания M_z:

$$M_z = M_x + M_y - 2M_xM_y + 2K_{x,y}(\tau) -$$

$$-\cancel{(1-M_x)(1-M_y)M_xM_y} - \cancel{\left[M_x(1-M_y) + (1-M_x)M_y\right]K_{x,y}(\tau)} - \cancel{K^2_{x,y}(\tau)} +$$

$$+\cancel{(1-M_x)(1-M_y)M_xM_y} + \cancel{\left[M_x(1-M_y) + (1-M_x)M_y\right]K_{x,y}(\tau)} + \cancel{K^2_{x,y}(\tau)} =$$

$$= M + M - 2MM + 2K(\tau).$$

Таким образом, при суммировании по модулю два функция ВКФ $K_{x,y}(\tau)$ преобразуется в АКФ $-K_z(\tau)$ и

$$\begin{cases} M_z = M_x + M_y - 2M_xM_y - 2K_z(\tau), \\ K_z(\tau) = -K_{x,y}(\tau). \end{cases} \quad (1.17)$$

Докажем выравнивающие свойства сумматора по модулю два для вероятности выходных событий элемента. Пусть $M_x = 0{,}5 \pm \omega$ и $M_y = 0{,}5 \pm \gamma$, тогда на основании (1.17)

$$M_z = 0{,}5 \pm \omega + 0{,}5 \pm \gamma - 2\left[(0{,}5 \pm \gamma)(0{,}5 \pm \omega)\right] - 2K_z(\tau).$$

Для простоты определим значение $K_z = 0$. При этом

$$M_z = 1 \pm \omega \pm \gamma - 2\left[(0{,}25 \pm 0{,}5\omega \pm 0{,}5\gamma \pm \gamma\omega)\right] = 1 \pm \omega \pm \gamma - 0{,}5 \mp \omega \mp \gamma \mp 2\gamma\omega =$$

$$= 0{,}5 \mp 2\gamma\omega.$$

Значения ω и γ лежат в пределах $0 \le |\omega, \gamma| \le 1$, поэтому произведение

$$2\gamma\omega << \gamma, \omega.$$

Итак, сумматор по модулю два, так же как и другие элементы, не уменьшает корреляционных связей в выходной последовательности, однако обладает выравнивающими свойствами для вероятности, приближая данный параметр к уровню 0,5.

1.6. Мгновенная относительная частота и первый критерий равномерного распределения элементарных событий

Как правило, априорно проектированию систем защиты информации разработчики цифровых ГПСЧ считают, что любая выборка событий достаточно большой длины n обладает свойствами случайных последовательностей.

Однако реальное применение устройств генерирования *M*-последовательностей указывает на необходимость учета множества объективных факторов как частного, так и общего характера. В первом случае качество эталонных чисел во многом определяется длиной базового регистра сдвига, достигающей на практике нескольких сотен бит. При этом в конкретных задачах используется не вся генеральная совокупность двоичных отсчетов, а только ограниченный набор или подпоследовательность, свойства которой зависят от начальной загрузки и полинома обратной связи. Что же касается процесса суммирования «сигнал+шум», то принцип «случайного детерминизма» формируемых полезных сигналов может приводить к снижению или потере динамики последовательности, не всегда устранимой схемами $\mathrm{mod}2$. Данные обстоятельства требуют внимательного отношения разработчиков аппаратуры к автокорреляционным и другим свойствам ПСП, что в общем случае сводится к выбору степени порождающего полинома и анализу АКФ при ограниченной длине случайной выборки.

Таким образом, исследование механизмов перехода статистических оценок в теоретические параметры могут быть полезны при решении задач выбора вероятностных последовательностей, обладающих эталонными характеристиками на малых длинах выборки, то есть подпадающих под определение термина миниасимптотический случайный процесс. Иными словами, известный интерес представляет собой выбор стартовой точки генератора, обеспечивающей минимальную длительность перехода вида «оценка – теоретический параметр» с учетом заданной границы для параметра.

Итак, пусть эмпирическая вероятность $\hat{p}(t)$ регистрации элементарного события в последовательности Бернулли после j наблюдений составляет величину

$$\hat{p}(t = j) = \frac{1}{j} k(j), \quad j = \overline{1,n}, \tag{1.18}$$

где переменная $k(t = j)$ в равенстве (1.18) соответствует статистике выбранного события в j отсчетах выборки.

На очередном шаге формирования (0,1)-чисел получим равенство

$$\hat{p}(j+1)=\frac{1}{j+1}k(j+1). \tag{1.19}$$

Если теперь вычислить разность между значениями (1.19) и (1.18), то можно записать соотношение для приращения эмпирической вероятности вида

$$\hat{p}(j+1)-\hat{p}(j)=\frac{1}{j+1}\left[x(j+1)-\frac{k(j)}{j}\right],$$

где $x(j+1)$ – очередное (0,1)- элементарное событие.

Из приведенного соотношения следует

$$\hat{p}(j+1)=\hat{p}(j)+\frac{1}{j+1}x(j+1)-\frac{1}{j+1}\frac{k(j)}{j}=\left(1-\frac{1}{j+1}\right)\hat{p}(j)+\frac{1}{j+1}x(j+1).$$

С другой стороны, в *j*-й момент времени, соответствующий достаточно длительному периоду наблюдения элементов выборки, соотношение для $\hat{p}(j+1)$ может быть записано с учетом текущего отклонения от равновероятности $\delta(j)$:

$$\frac{1}{j}k(j)=\frac{1}{2}\pm\delta(j).$$

Тогда для $j+1$-го момента времени будут справедливы выражения

$$\frac{1}{j+1}k(j+1)=\frac{1}{2}\pm\delta(j+1),$$

Разность выборочных значений вероятности дает результат

$$\frac{1}{j+1}k(j+1)-\frac{1}{j}k(j)=\left[\frac{1}{2}\pm\delta(j+1)\right]-\left[\frac{1}{2}\pm\delta(j)\right]=\pm\delta(j+1)\mp\delta(j),$$

$$\hat{p}(j+1)=\hat{p}(j)\pm\delta(j+1)\mp\delta(j).$$

Данное соотношение показывает, что в истинно случайной последовательности в каждом такте формирования (0,1)-событий отклонение от равновероятности на $j+1$ шаге принципиально уменьшается относительно того же параметра на j-м шаге.

Величину допустимой границы отклонения выборочной вероятности от теоретического значения определим соотношением (для ГПСЧ):

$$\delta_0=\frac{s}{2^l-1},\quad s=0,1,2,.... \tag{1.20}$$

Апостериорно, то есть по завершении заданного числа испытаний, в некоторый момент времени, соответствующий j_0-му наблюдению, будет выполнено неравенство

$$\left|\frac{1}{j}k(j)-p(t)\right|<\delta_0,\quad \forall j\ge j_0. \tag{1.21}$$

Очевидно, что значение j_0 будет характеризовать момент преобразования относительной частоты в теоретический параметр с учетом границы δ_0. При этом следует помнить, что известные на сегодняшний день свойства *M*- последовательностей определяются как идеальные только на полной длине выборки, то есть при $n=2^l-1$. В этом случае для всех начальных состояний граница $\delta_0=2^{-l}$, если $p=0,5$.

Сформулируем первое из требований, предъявляемых к структуре случайной последовательности, выполнение которого характеризовало бы статистические свойства (0,1)-выборки как идеальные в процессе движения оценки к вероятностному параметру. С этой целью из неравенства (1.21) выразим сумму событий $k(j)$, используя в качестве границы величину (1.20). Полагая теперь, что для событий Бернулли величина теоретической вероятности $p=0,5$, с учетом (1.21) имеем

$$\frac{1}{j}k(j)=\frac{1}{2}\pm\delta(j).$$

$$k(j)=\frac{j}{2}\pm s,\quad j\ge j_0.$$

Таким образом, последовательности со случайной природой, идеально удовлетворяющие критерию равновероятности (при выполнении требований критерия серий), могут считаться миниасимптотически случайными, если динамика формирования их такова, что в каждом такте генерирования выполняется условие

$$\frac{1}{j}k(j)\to\frac{1}{2},\ \forall j\le n \tag{1.22}$$

или $2k(j)\to j$, что в асимптотике дает равновероятность $k(n)=0,5n$.

Сформулированное условие исключает возможность группирования однотипных элементарных событий в определенной части выборки. Однако вероятность скученности различных серий из нулей и единиц при этом остается. В связи с этим далее исследуем выборочную автокорреляционную функцию и получим аналогичные соотношения, определяющие в миниасимптотике степень связности элементарных событий через τ тактов наблюдения.

1.7. Мгновенная эмпирическая АКФ и второй критерий равномерности для элементарных событий

Известные классические соотношения для эмпирической функции автокорреляции не дают четкого представления об изменении меры зависимости элементарных событий с течением времени t. Поэтому известное выражение

для АКФ запишем относительно момента времени j при $j \to n, n \to \infty$:

$$\hat{K}_x(t \le j, \tau) = \frac{1}{j}\sum_{t=1}^{j}\left[x(t) - M_x\right]\left[x(t+\tau) - M_x\right] =$$
$$= \frac{1}{j}\sum_{t=1}^{j}\left[x(t)x(t+\tau) - M_x\left[x(t) + x(t+\tau)\right] + M_x^2\right], \qquad (1.23)$$
$$x(t) \in \{0,1\}.$$

В момент времени j+1 получим равенство

$$\hat{K}_x(j+1,\tau) = \frac{1}{j+1}\sum_{t=1}^{j+1}\left[x(t) - M_x\right]\left[x(t+\tau) - M_x\right]. \qquad (1.24)$$

Преобразуем соотношение (1.24) и получим зависимость (j+1)-го отсчета автокорреляционной функции от j-го:

$$\hat{K}_x(j+1,\tau) = \frac{1}{j+1}\left[\sum_{t=1}^{j}x(t)x(t+\tau) + x(j+1)x(j+\tau+1) - \sum_{t=1}^{j}M_x\left[x(t) + x(t+\tau)\right] - \right.$$
$$\left. - M_x\left[x(j+1) + x(j+\tau+1)\right] + \sum_{t=1}^{j}M_x^2 + M_x^2\right] =$$
$$= \frac{1}{j+1}\left[\sum_{t=1}^{j}x(t)x(t+\tau) - \sum_{t=1}^{j}M_x\left[x(t) + x(t+\tau)\right] + \sum_{t=1}^{j}M_x^2\right] +$$
$$+ \frac{1}{j+1}\left[x(j+1)x(j+\tau+1) - M_x\left[x(j+1) + x(j+\tau+1)\right] + M_x^2\right] =$$
$$= \frac{j}{j+1}\hat{K}(j,\tau) + \frac{1}{j+1}\left[x(j+1)x(j+\tau+1) - M_x\left[x(j+1) + x(j+\tau+1)\right] + M_x^2\right].$$
$$\hat{K}_x(j,\tau) - \hat{K}_x(j+1,\tau) = \left(1 - \frac{j}{j+1}\right)\hat{K}_x(j,\tau) - \frac{1}{j+1}\left[x(j+1) - M_x\right]\left[x(j+\tau+1) - M_x\right].$$

Данные зависимости показывают, что каждое очередное испытание корректирует предшествующее значение автокорреляционной функции не только на величину $+\frac{1}{j+1}\left[x(j+1) - M_x\right]\left[x(j+\tau+1) - M_x\right]$, но и на значение $-\frac{1}{j+1}\hat{K}_x(j,\tau)$.

С учетом равенства (1.23) сформулируем второе, основное требование к структуре случайного сообщения (то есть к расположению элементарных событий в выборке) при статистическом переходе мгновенной выборочной АКФ в теоретический параметр. С этой целью указанное соотношение преобразуем в неравенство, используя границу для теоретической вероятности δ_0. Полагая далее, что сумма произведений $x(t)x(t+\tau)$ в точке преобразования «оценка – теоретический параметр» может быть характеризована произведением относительных частот (для независимых событий), имеем

$$\frac{k_j \cap k_{j,\tau} = k_\eta}{j} - M_x\left(\frac{k_j}{j} + \frac{k_{j,\tau}}{j}\right) + M_x^2 \approx \left(\frac{1}{2} \pm \frac{s}{2^l - 1}\right)\left(\frac{1}{2} \pm \frac{s}{2^l - 1}\right) - \frac{1}{4},$$

где $k_{j,\tau}$ – статистика элементарных событий, наблюдаемых через интервал времени τ относительно отсчета t. Раскрывая скобки в левой и правой частях данного соотношения, а также упрощая полученное выражение, приходим к равенству

$$\frac{k_\eta}{j}-\frac{1}{2j}\left(k_j+k_{j,\tau}\right)+\frac{1}{4}\approx\left(\frac{1}{4}\pm\frac{s}{2^l-1}+\frac{s^2}{\left(2^l-1\right)^2}\right)-\frac{1}{4}.$$

Отсюда следует

$$k_\eta-\frac{1}{2}\left(k_j+k_{j,\tau}\right)\approx j\left(\frac{1}{4}\pm\frac{s}{2^l-1}+\frac{s^2}{\left(2^l-1\right)^2}\right)-\frac{j}{2}. \qquad (1.25)$$

Из (1.25) можно получить два соотношения, характеризующих структурные свойства (взаимное расположение нулей и единиц) случайного сообщения. Во-первых,

$$k_j\cap k_{j,\tau}\approx j\left(\frac{1}{4}\pm\frac{s}{2^l-1}+\frac{s^2}{\left(2^l-1\right)^2}\right)=j\left(\frac{1}{2}\pm\frac{s}{2^l-1}\right)\left(\frac{1}{2}\pm\frac{s}{2^l-1}\right).$$

Откуда следует

$$\frac{k_j\cap k_{j,\tau}}{j}\approx\left(\frac{1}{2}\pm\frac{s}{2^l-1}\right)\left(\frac{1}{2}\pm\frac{s}{2^l-1}\right)\quad\text{и}\quad\frac{k_j\cap k_{j,\tau}}{j}\to\frac{1}{4}\quad\text{для}\ \forall j. \qquad (1.26)$$

Во-вторых,

$$\frac{1}{2}\left(k_j+k_{j,\tau}\right)\approx\frac{j}{2}. \qquad (1.27)$$

Очевидно, что соотношение (1.27) согласуется с результатом (1.22), а равенство (1.26) исключает возможность группирования наборов элементарных событий в любой части псевдослучайной или случайной последовательности.

Таким образом, равенство (1.26) представляет собой второй критерий равномерности элементарных событий и предполагает выполнение условия (1.22) для пар наблюдаемых объектов. Кроме того, учитывая, что на параметр τ ограничения не налагаются, можно сделать вывод о возможности применения критерия (1.26) ко всем временным интервалам $\tau=\overline{1,n-1}$.

Рассмотрим далее частный случай формирования АКФ для временного интервала $\tau=0$.

1.8. Мгновенная эмпирическая дисперсия и доверительный интервал для вероятности наблюдения (0,1)-событий

Требуемое соотношение для дисперсии может быть получено из классического равенства для АКФ при $\tau=0$:

$$\hat{K}_x(j,\tau)=\frac{1}{j}\sum_{t=1}^{j}\left[x(t)-M_x\right]^2,\qquad x(t)\in\{0,1\}.$$

При $n \to \infty$ из данного равенства следует

$$\hat{K}_x(j,\tau) = \frac{1}{j}\sum_{t=1}^{j} x(t)x(t) - M_x \frac{1}{j}\sum_{t=1}^{j}\left[x(t)+x(t)\right] + M_x^2 = M_x^2.$$

Таким образом, из формулы (1.26) для $n \to \infty$ могут быть получены равенства вида

$$\frac{k_j \cap k_j}{j} = \frac{k_j}{j} \to \frac{1}{2} \quad \text{для} \quad \forall j, \qquad \text{и} \quad \frac{1}{2}\left(2k_j\right) \approx \frac{j}{2}. \tag{1.28}$$

Очевидно, что равенство (1.28) полностью повторяет предыдущие результаты, а, следовательно, соотношение для дисперсии в задаче синтеза миниасимптотического процесса имеет второстепенное значение.

Определим доверительный интервал для параметра $p(t) = p$, отвечающий доверительной вероятности $1-\varepsilon$, учитывая динамику движения оценки к значению теоретического параметра. С этой целью запишем равенство

$$P\left\{\left|\frac{1}{j}k_j - p\right| < \mu\sqrt{\frac{1}{j^2}k_j\left(j-k_j\right)}\right\} = 1-\varepsilon.$$

Очевидно, что использованное в данном соотношении неравенство дает возможность записать пределы для вероятности в виде

$$\frac{1}{j}k_j - \mu\sqrt{\frac{1}{j^2}k_j\left(j-k_j\right)} < p < \frac{1}{j}k_j + \mu\sqrt{\frac{1}{j^2}k_j\left(j-k_j\right)} \tag{1.29}$$

или

$$k_j - \mu j\sqrt{\hat{p}\hat{q}} < jp < k_j + \mu j\sqrt{\hat{p}\hat{q}}, \qquad j \to n. \tag{1.30}$$

Отсюда можно определить величину границы для формулы (1.20). Действительно, учитывая формулы (1.29) и (1.30) и полагая, что требование (1.21) выполняется, имеем

$$s < \mu\sqrt{k_j\left(j-k_j\right)},$$
$$k_j \to 2^{l-1}, \quad j \to 2^l.$$

Из данного соотношения можно определить значение *s* по формуле

$$s \approx 2^{l-1}\mu,$$
$$j = n = 2^l.$$

Итак, при использовании *M*-последовательности в целях защиты информационных сетей необходимо, чтобы формируемые (0,1)-отсчеты удовлетворяли бы двум дополнительным требованиям:

1) эмпирическая вероятность наблюдения бернуллиевских событий в каждом такте формирования чисел должна стремиться к значению 0,5;

2) эмпирическая вероятность наблюдения пар бернуллиевских событий в каждом такте должна стремиться к произведению вероятностей 0,25 для любых сдвигов τ.

1.9. Вероятностные преобразователи информации

При решении задач, связанных с защитой информационных сетей, часто используют устройства, получившие название линейных преобразователей «код – вероятность». Классической схемой такого устройства является схема, представленная на рис.1.13. Функционирование схемы основывается на сравнении числа $\boldsymbol{A}$ и случайного двоичного вектора $X=\{x_1,x_2,...,x_l\}$. Если случайные числа X имеют равномерное распределение в диапазоне представления чисел $A_0, A_1, A_2, ..., A_{2^l-1}$, то на выходе z формируется двоичная последовательность в соответствии с законом

$$z_i=\begin{cases}1 \;\; при\; A_j > X_i,\\ 0 \;\; при\; A_j \le X_i, \quad j=const,\end{cases}$$

где i – номер случайного числа.

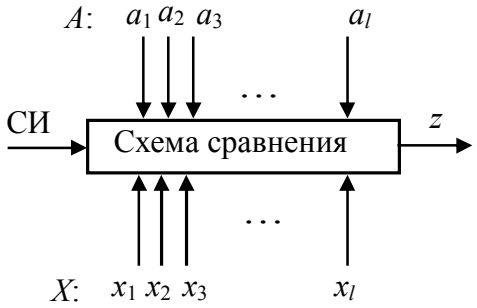

Рис. 1.13

Вероятность единичного значения $p(z)$ на выходе преобразователя в общем случае может быть определена соотношением

$$p(z)=A_j\cdot 2^{-l}.$$

Более сложные устройства строятся с использованием двоичных счетчиков и применяются для интегрирования стохастических последовательностей. Структурная схема стохастического интегратора представлена на рис. 1.14.

Принцип действия таких устройств основывается на следующих математических законах. Если на вход счетчика подаются единичные случайные приращения, то величина интеграла за время t будет равна произведению

$$I=2^{-l}\,p(x)t\,,$$

где l – разрядность счетчика, $p(x)$ – вероятность единичного значения во входной последовательности. При этом

$$p(x)=\frac{k=2^{l-1}}{n=2^{l}},\quad 2^{-l}\,p(x)t=\frac{2^{l-1}}{2^{l}}\cdot\frac{t}{2^{l}}=\frac{0{,}5t}{2^{l}}.$$

Содержимое накопителя <Сm2> в данной схеме определяется соотношением: $<Cm2>=np(x)$,

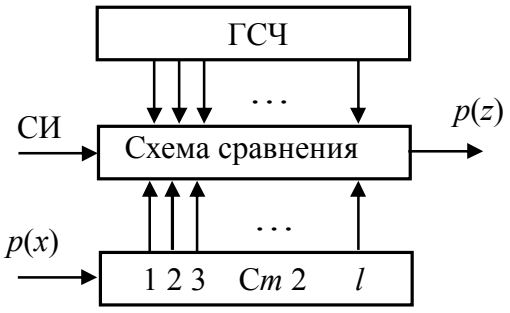

Рис. 1.14

где x – единичное элементарное событие, $n=2^{l}$ – длина последовательности.

Доказательство следует из следующих рассуждений. Очевидно, что при равномерном распределении элементарных событий содержимое <Сm2> будет определяться суммой единичных приращений в каждом такте, умноженных на вероятность этих приращений:

$$M_{Cm2}=x_1p(x_1)+x_2p(x_2)+x_3p(x_3)+\ldots+x_np(x_n)=\sum_{i=1}^{n}x_ip(x_i)=$$

$$=\left|p(x_i)=p(x)\right|=p(x)\sum_{i=1}^{n}x_i=np(x).$$

Теперь, учитывая, что схема преобразователя вырабатывает значения выходной переменной z, используя l – разрядные равномерно распределенные случайные числа (счетчик рассматривается как база для сравнения), для вероятности $p(z)$ будет справедливо соотношение

$$p(z_i)=\frac{1}{2^{l}}<Cm2>_i\xrightarrow{i\to n}=\frac{1}{2^{l}}M_{Cm2}=2^{-l}np(x).$$

Полученная схема может быть модернизирована за счет введения отрицательной обратной связи. Это позволяет получить так называемый следящий режим, а соответствующее устройство получило название следящего интегратора (рис. 1.15).

По цепи обратной связи сигналы выходной последовательности $p(z)$ поступают на вход «–» реверсивного счетчика. В случае совпадения в среднем математических ожиданий последовательностей, действующих на входе и в цепи обратной связи, в схеме устанавливается динамическое равновесие.

Алгоритм генерирования чисел в последовательности $p(z)$ определяется с учетом условия

$$z_i=\begin{cases}1 \;\; при \;\; (Cm2)_i>(ГСЧ)_i,\\ 0 \;\; при \;\; (Cm2)_i\le(ГСЧ)_i.\end{cases}$$

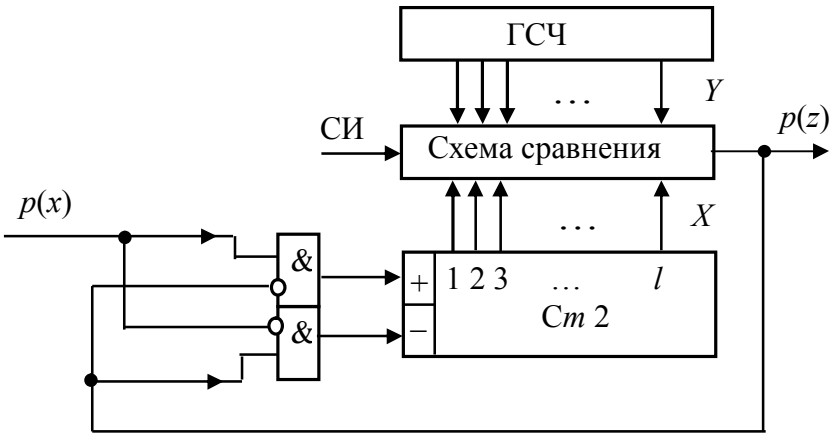

Рис. 1.15

Для математических ожиданий содержимого счетчика можно записать следующее рекуррентное соотношение:

$$M_{Cm2,,i} = M_{Cm2,,i\text{-}1} - \frac{1}{2^l} M_{Cm2,,i\text{-}1} + p(x_i) = \left| 2^l = n,\ p(x_i) = p,\ \forall i \right| = M_{Cm2,,i\text{-}1}\left(1 - \frac{1}{n}\right) + p,$$

где i – номер машинного такта.

Полагая, что в начальные моменты времени $M_{Cm2,0} = 0$, $M_{Cm2,1} = p$, выпишем значения счетчика в последующие такты работы:

$$\begin{aligned} M_{Cm2,2} &= p + p\left(1 - \frac{1}{n}\right), \\ M_{Cm2,3} &= p + p\left(1 - \frac{1}{n}\right) + p\left(1 - \frac{1}{n}\right)^2, \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ M_{Cm2,m} &= p\left(1 + a + a^2 + \ldots + a^{m-2} + a^{m-1}\right),\ a = 1 - \frac{1}{n}, \end{aligned} \tag{1.31}$$

где выражение в скобках есть сумма членов геометрической прогрессии со знаменателем a. При общем числе членов m из соотношений (1.31) получим

$$M_{Cm2,m} = p\frac{1-a^m}{1-a} = p\left(1-a^m\right)\frac{1}{1-1+\frac{1}{n}} = np\left(1-a^m\right).$$

Сомножитель $1-a^m$ преобразуем к виду

$$1 - a^m = 1 - \left[\left(1 - \frac{1}{n}\right)^n\right]^{\frac{m}{n}} \xrightarrow{n>>1} 1 - e^{-\frac{m}{n}}.$$

Отсюда

$$p(z) = \frac{1}{n} M_{Cm2,m} = p(x)\left(1 - e^{-\frac{m}{n}}\right) = \left| m = n_1 > n \right| = p(x)\left(1 - e^{-\frac{n_1}{n}}\right). \tag{1.32}$$

Соотношение (1.32) показывает, что если входная случайная последовательность интегратора является стационарной, то процесс изменения содержимого реверсивного счетчика подчиняется экспоненциальному закону и переходные явления в устройстве затухают при $n_1 > n$.

1.10. Поточная шифросистема Д. Гиффорда

Поточная система шифрования Гиффорда получила распространение начиная с 1984 г. Основной элемент системы это 64-разрядный криптографический генератор с линейной обратной связью *FB* и нелинейной функцией выхода *FF*. Ключом сообщения является код начального заполнения регистра. Принципиально схема реализует шифрование путем наложения псевдослучайной гаммы на текст предварительно кодированного или открытого сообщения.

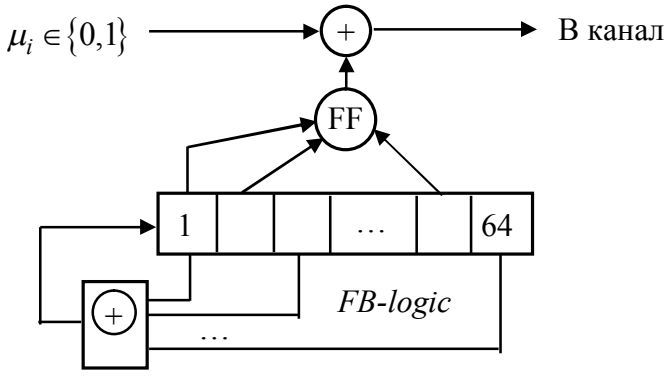

Рис. 1.16

Однако практика применения показала, что данная система оказалась не достаточно стойкой к вскрытию. В частности, в печати имеются сведения о том, что расшифровка сообщений, закрытых с помощью схемы Гиффорда, требует в среднем 4 часа времени при использовании в целях декодирования 8 специальных станций (*Sparc*).

Тем не менее линейные регистры сдвига до настоящего времени применяются для различных компьютерных приложений, так как обладают рядом положительных свойств:

1) в схеме кодирования используются только простейшие операции сложения по модулю два и логического умножения;

2) высокое быстродействие создаваемых алгоритмов кодирования;

3) большое количество теоретических исследований, позволяющих использовать ГПСЧ в различных криптографических приложениях.

Алгоритм установления шифрованной связи между абонентами при использовании ГПСЧ состоит в следующем.

Для согласования работы своих шифрующих устройств абоненты договариваются о выборе числа g, так, чтобы выполнялось условие взаимной простоты $\left(g, 2^{64} - 1\right) = 1$.

После этого независимо друг от друга участники связи выбирают числа α и β из условия

$$1 < \alpha, \beta \le 2^{64} - 1.$$

Далее первый из абонентов вычисляет значение $g^{\alpha} \pmod{2^{64}}$, а второй – $g^{\beta} \pmod{2^{64}}$, после чего они обмениваются результатами.

Зная число $g^{\beta} \pmod{2^{64}}$, первый абонент определяет ключ $g^{\beta\alpha} \pmod{2^{64}}$, а второй, зная $g^{\alpha} \pmod{2^{64}}$, получает значение $g^{\alpha\beta} \pmod{2^{64}}$. Таким образом, оба абонента обладают одним и тем же ключом:

$$g^{\alpha\beta} \pmod{2^{64}} = g^{\beta\alpha} \pmod{2^{64}}.$$

В заранее обусловленное время корреспонденты закладывают ключ в свои шифрующие устройства, что и позволяет достичь требуемого процесса синхронизации.

Сокращенный список примитивных полиномов по модулю два, используемых в целях шифрования, может быть представлен табл. 1.2.

Таблица 1.2

$\varphi(x) = 1 + x^j + x^l$					
l	j или $l–j$	$M = 2^l–1$	l	j или $l–j$	$M = 2^l–1$
4	1	15	18	7	262143
5	2	31	20	3	1048575
6	1	63	21	2	2097151
7	1 или 3	127	22	1	4194303
9	4	511	23	5 или 9	8388607
10	3	1023	25	3 или 7	33554431
11	2	2047	28	3,9 или 13	268435455
15	1,4 или 7	32767	31	3,6,7 или 13	2147483647
17	3	131071	33	13	8589934591

Пример. Для полиномов степени $l = 23$ можно составить следующие трехчлены:

$$\varphi(x) = 1 + x^5 + x^{23},$$
$$\varphi(x) = 1 + x^9 + x^{23}.$$

Примечание. В случае если полином требуемой разрядности в таблице отсутствует, необходимо воспользоваться классическими первоисточниками литературы, например [23].

2. МАТЕМАТИЧЕСКИЕ ОСНОВЫ КОМПЬЮТЕРНОЙ КРИПТОГРАФИИ

2.1. Криптосистема без передачи ключей

Основу данного метода криптографии составляют теоремы Эйлера и Ферма, а также алгоритмы решения сравнений первой степени, базирующиеся, в свою очередь, на алгоритме Евклида.

Сущность алгоритма Евклида состоит в отыскании нод(a,b) и заключается в следующем. Пусть a и b – положительные целые числа и $a > b$. Тогда для указанных значений будут справедливы соотношения

$$\begin{aligned}
&a = bq_1 + r_2, && 0 < r_2 < b,\\
&b = r_2q_2 + r_3, && 0 < r_3 < r_2,\\
&r_2 = r_3q_3 + r_4, && 0 < r_4 < r_3,\\
&\qquad\qquad . \; . \; . \; . \; . \; . \; . \; . \; . \; . \; . \; . \; .\\
&r_{n-2} = r_{n-1}q_{n-1} + r_n, && 0 < r_n < r_{n-1},\\
&r_{n-1} = r_nq_n.
\end{aligned}$$

Алгоритм заканчивается при получении некоторого остатка $r_{n+1} = 0$. Наибольшим общим делителем тогда будет последний не равный нулю остаток r_n.

Алгоритм Евклида можно представить также в виде следующих арифметических действий:

a	b
bq_1	q_1
r_2	

b	r_2
r_2q_2	q_2
r_3	

r_2	r_3
r_3q_3	q_3
r_4	

...

r_{n-2}	r_{n-1}
$r_{n-1}q_{n-1}$	q_{n-1}
r_n	

r_{n-1}	r_n
r_nq_n	q_n
0	

(525, 231)=21.

525	231
462	2
63	

231	63
189	3
42	

63	42
42	1
21	

42	21
42	q_4=2
0	n=4

В теории существует математическая связь данного алгоритма и представления частного от деления двух чисел в виде полиномиальных дробей:

$$\begin{aligned}
&P_s = q_sP_{s-1} + P_{s-2}\ ,\\
&Q_s = q_sQ_{s-1} + Q_{s-2}, \qquad P_0 = 1,\ Q_0 = 0,\ Q_1 = 1.
\end{aligned}$$

Из данных равенств следует

$$\frac{P_1}{Q_1} = q_1, \qquad \frac{P_2}{Q_2} = q_1 + \frac{1}{q_2}, \qquad \frac{P_3}{Q_3} = q_1 + \frac{1}{q_2 + \dfrac{1}{q_3}},$$

$$\frac{P_4}{Q_4} = q_1 + \frac{1}{q_2 + \dfrac{1}{q_3 + \dfrac{1}{q_4}}}, \quad \ldots \quad \frac{P_n}{Q_n} = q_1 + \frac{1}{q_2 + \dfrac{1}{\ldots + \dfrac{1}{q_{n-1} + \dfrac{1}{q_n}}}} = \frac{a}{b}.$$

Обычно эти вычисления представляют в виде табл. 2.1.

Таблица 2.1

q_s		q_1	q_2	…	q_{s-2}	q_{s-1}	q_s	…	q_{n-1}	q_n
P_s	1	q_1	P_2	…	P_{s-2}	P_{s-1}	P_s	…	P_{n-1}	a
Q_s	0	1	Q_2	…	Q_{s-2}	Q_{s-1}	Q_s	…	Q_{n-1}	b

Пример. Пусть $a = 105,\ b = 38,$ тогда в соответствии с алгоритмом Евклида значения q будут равны $q_s = \{2,1,3,4,2\}$, а табл. 2.1 примет вид

Таблица 2.2

	s=0	s=1	s=2	s=3	s=4	s=5
q_s		2	1	3	4	2
P_s	1	2	3	11	47	105
Q_s	0	1	1	4	17	38

Из табл. 2.2 имеем:

$$\frac{P_1}{Q_1} = q_1 = \frac{2}{1}, \qquad \frac{P_2}{Q_2} = q_1 + \frac{1}{q_2} = 2 + \frac{1}{1} = \frac{3}{1}, \qquad \frac{P_3}{Q_3} = q_1 + \frac{1}{q_2 + \dfrac{1}{q_3}} = 2 + \frac{1}{1 + \dfrac{1}{3}} = \frac{11}{4},$$

$$\frac{P_4}{Q_4} = q_1 + \frac{1}{q_2 + \dfrac{1}{q_3 + \dfrac{1}{q_4}}} = 2 + \frac{1}{1 + \dfrac{1}{3 + \dfrac{1}{4}}} = \frac{47}{17}, \qquad \frac{P_5}{Q_5} = q_1 + \frac{1}{q_2 + \dfrac{1}{q_3 + \dfrac{1}{q_4 + \dfrac{1}{q_5}}}} = 2 + \frac{1}{1 + \dfrac{1}{3 + \dfrac{1}{4 + \dfrac{1}{2}}}} = \frac{105}{38}.$$

Дальнейшее математическое обоснование криптосистемы без передачи ключей базируется на решении сравнений первой степени с одним неизвестным.

Для определения понятия «сравнение» будем рассматривать некоторые целые числа, представляющие собой остатки от деления на другое, заданное натуральное число m, которое назовем модулем.

Сравнимость чисел a и b по модулю m записывается в виде

$$a \equiv b \pmod{m}$$

и читается «a сравнимо с b по модулю m».

Сравнения обладают рядом свойств, характерных для обыкновенных равенств:

а) если два числа сравнимы с третьим, то все три числа сравнимы между собой по заданному модулю;

б) сравнения можно почленно складывать:

$$a_1 \equiv b_1 \pmod{m},\ a_2 \equiv b_2 \pmod{m},\quad a_3 \equiv b_3 \pmod{m},\ \ldots\ a_n \equiv b_n \pmod{m},$$

тогда

$$a_1 + a_2 + a_3 + \ldots + a_n \equiv b_1 + b_2 + b_3 + \ldots + b_n \pmod{m};$$

в) слагаемое, стоящее в какой-либо части сравнения, можно переносить в другую часть, переменив знак на противоположный:

$$a + b \equiv c \pmod{m} \longrightarrow a \equiv c - b \pmod{m}.$$

Пример. $15 + 7 \equiv 9 \pmod{13},\ 15 \equiv 9 - 7 \pmod{13}$;

г) к каждой части сравнения можно прибавить любое число, кратное модулю:

$$a \equiv b \pmod{m} \longrightarrow a + nm \equiv b \pmod{m};$$

д) сравнения можно почленно перемножать:

$$\begin{gathered} a_1 \equiv b_1 \pmod{m},\ a_2 \equiv b_2 \pmod{m},\quad a_3 \equiv b_3 \pmod{m},\ \ldots\ a_n \equiv b_n \pmod{m}, \\ a_1 \cdot a_2 \cdot a_3 \cdot \ldots \cdot a_n \equiv b_1 \cdot b_2 \cdot b_3 \cdot \ldots \cdot b_n \pmod{m}; \end{gathered} \qquad (2.1)$$

е) следствие из (2.1)

$$a^k \equiv b^k \pmod{m};$$

ж) если выражения для многочленов A и B сравнимы по модулю m с некоторым числом w, то есть

$$\begin{aligned} A:\ & a_0 x^n + a_1 x^{n-1} + a_2 x^{n-2} + \ldots + a_n x^0 \equiv w \pmod{m}, \\ B:\ & b_0 x^n + b_1 x^{n-1} + b_2 x^{n-2} + \ldots + b_n x^0 \equiv w \pmod{m}, \end{aligned}$$

то сравнимы и сами многочлены A и B по модулю m:

$$\sum_{i=0}^{n} a_i x^{n-i} \equiv \sum_{i=0}^{n} b_i x^{n-i} \pmod{m}$$

(данное свойство является обобщением предыдущих свойств сложения и умножения);

з) обе части сравнения и модуль можно умножить на одно и тоже число:

$$a \equiv b \pmod{m}, \quad ak \equiv bk \pmod{mk}.$$

Рассмотрим теперь две теоремы, являющиеся базой для криптографического метода с закрытой системой ключей.

Теорема Эйлера. Пусть числа a и m взаимно простые, то есть $(a,m)=1$, тогда будет справедливо равенство

$$a^{\varphi(m)} \equiv 1 \pmod{m}, \quad m > 1,$$

где $\varphi(m)$ – это функция Эйлера, которая показывает число чисел взаимно простых с m с учетом 1. Считается, что 1 тоже обладает взаимной простотой с m.

Пример. $\varphi(6)=2$, так как в ряду 1,2,3,4,5 только два числа 1 и 5 считаются взаимно простыми с числом 6. Соответственно:

$$7^{\varphi(6)} \equiv 1 \pmod{6}.$$

Теорема Ферма. Если p – простое число, а число a не делится на p, то есть $a \neq sp$, и тогда имеет место сравнение вида

$$a^{p-1} \equiv 1 \pmod{p},$$

где $\varphi(p) = p-1$. Например $9^{\varphi(7)=6} \equiv 1 \pmod{7}$.

Приведенные две теоремы в задачах криптографии применяются для решения сравнений, корни которых используются для выбора секретных ключей.

Пусть требуется решить сравнение первой степени с одним неизвестным:

$$ax \equiv b \pmod{m}, \quad (a,m)=1. \tag{2.2}$$

Принцип решения сравнения (2.2) основывается на разложении отношения m/a в непрерывную (полиномиальную) дробь. Согласно свойствам непрерывных дробей, можно записать

$$P_s Q_{s-1} - P_{s-1} Q_s = (-1)^s.$$

Тогда, рассматривая две последние подходящие дроби, имеем

$$\frac{P_{n-1}}{Q_{n-1}}, \frac{P_n}{Q_n} = \frac{m}{a}, \qquad mQ_{n-1} - aP_{n-1} = (-1)^n.$$

Отсюда следует, что

$$aP_{n-1} \equiv (-1)^{n-1} \pmod{m},$$

$$a(-1)^{n-1} P_{n-1} b \equiv b \pmod{m}, \quad x \equiv \frac{b}{a} \pmod{m}, \quad x \equiv \frac{a(-1)^{n-1} P_{n-1} b}{a} \pmod{m}, \tag{2.3}$$

$$x \equiv (-1)^{n-1} P_{n-1} b \pmod{m}.$$

Итак, для решения сравнения (2.2) необходимо определить только числитель $n-1$ подходящей дроби, согласно алгоритму Евклида, и подставить его в решение (2.3).

Пример. Решить сравнение вида $111x \equiv 75 \pmod{321}$. Для решения поставленной задачи необходимо указать на то, что

$$(111,321) = 3, \qquad 111 = 37 \cdot 3, \qquad 321 = 107 \cdot 3.$$

Таким образом, начальный вариант сравнения можно сократить на 3, откуда следует

$$37x \equiv 25 \pmod{107}. \tag{2.4}$$

Разложим отношение $\frac{107}{37}$ в непрерывную дробь и составим табл. 2.3 для числителей P_s.

$$\begin{array}{r|l} 107 & 37 \\ \hline 74 & 2 \\ \hline 37 & 33 \\ \hline 33 & 1 \\ \hline 33 & 4 \\ \hline 32 & 8 \\ \hline 4 & 1 \\ \hline 4 & 4 \\ \hline 0 & \end{array}$$

Таблица 2.3

	$s=0$	$s=1$	$s=2$	$s=3$	$s=4$
q_s		2	1	8	4
P_s	1	2	3	26	107

Для данного примера $n=4,\ P_3 = 26,\ b = 25.$ Следовательно, решение сравнения (2.4) может быть представлено в виде

$$x \equiv (-1)^3\, 26 \cdot 25 \equiv 99 \pmod{107}.$$

Решением исходного сравнения будут три корня вида

$$x \equiv 99 + k \cdot 107 \pmod{321}, \quad k = 0,1,2;$$
$$x \equiv 99,\ 206,\ 313 \pmod{321}.$$

Криптографический метод на основе теорем Эйлера и Ферма основывается на следующем утверждении.

Следствие из теоремы Ферма. Если число p –простое, а число $0 < \mu \le p-1$, то существует такое $k \equiv 1 \pmod{p-1}$, при котором

$$\mu \equiv \mu^k \pmod{p}.$$

Итак, пусть абоненты A и B условились организовать между собой секретную от наблюдателя C переписку. С этой целью они выбирают достаточно большое простое число p, такое, что значение $p-1$ хорошо разлагается на большое число сомножителей. Каждый из абонентов независимо один от дру-

гого выбирает числа-ключи a и b и сопутствующие им неслучайные параметры α и β из условия

$$a \cdot \alpha \equiv 1 \pmod{p-1}, \qquad 0 < \alpha < p-1;$$
$$b \cdot \beta \equiv 1 \pmod{p-1}, \qquad 0 < \beta < p-1.$$

Таким образом, абонент A обладает своими секретными ключами a и α, неизвестными абоненту B, а абонент B обладает секретными ключами b и β, неизвестными абоненту A. В данной постановке задачи очевиден факт

$$a \cdot \alpha \cdot b \cdot \beta \equiv 1 \pmod{p-1}.$$

Пусть абонент A решает передать сообщение μ (заметим, что не все μ кодируются данным методом) абоненту B, причем $0 < \mu \le p-1$. Тогда A зашифровывает сообщение своим первым секретным ключом, находит шифротекст

$$\mu_1 \equiv \mu^{\alpha} \pmod{p}, \quad 0 < \mu_1 < p$$

и отправляет его абоненту B.

Приняв шифротекст, абонент B, в свою очередь зашифровывает сообщение μ_1 своим первым секретным ключом

$$\mu_2 \equiv \mu_1^{\beta} \pmod{p}, \quad 0 < \mu_2 < p$$

и пересылает его обратно абоненту A. Абонент A, получив обратно свое теперь уже дважды шифрованное сообщение, шифрует его в третий раз своим вторым ключом

$$\mu_3 \equiv \mu_2^{a} \pmod{p}, \quad 0 < \mu_3 < p$$

и вновь отправляет его абоненту B. Последний расшифровывает это сообщение при помощи своего второго ключа

$$\mu_4 \equiv \mu_3^{b} \pmod{p}, \quad 0 < \mu_4 < p\,.$$

Действительно, последнее равенство может быть записано в виде

$$\mu_4 \equiv \mu^{\alpha\beta ab} \equiv \mu^{k} \pmod{p},$$

а так как $\alpha\beta ab = k \equiv 1 \pmod{p-1}$, то $\mu^{k} \equiv \mu \pmod{p}$ или $\mu_4 \equiv \mu \pmod{p}$.

Пример. Абоненты A и B решили установить между собой защищенную от наблюдателя C связь без передачи ключей с простым модулем $p = 23$.

Абонент A выбирает случайным образом число $a = 5$ и находит к нему неслучайный параметр α, решая сравнение $5\alpha \equiv 1 \pmod{\varphi(23)}$; $\alpha = 9$. Следовательно, числа 5 и 9 – секретные ключи абонента A.

Абонент B выбирает случайным образом число $b = 7$ и находит к нему неслучайный параметр β, решая сравнение $7\beta \equiv 1 \pmod{\varphi(23)}$; $\beta = 19$. Следовательно, числа 7 и 19 – секретные ключи абонента B.

Абонент A решает секретно передать очень важное сообщение $\mu = 17$ абоненту B. Он шифрует сообщение своим первым ключом:

$$\mu_1 \equiv 17^5 \equiv 21 \pmod{23}.$$

В процессе обмена выполняются действия

$$\mu_2 \equiv 21^7 \equiv 10 \pmod{23},$$
$$\mu_3 \equiv 10^9 \equiv 20 \pmod{23},$$
$$\mu_4 \equiv 20^{19} \equiv 17 \pmod{23}.$$

Итак, $\mu = 17,$ так как $0 < 17 < 23.$

В целом, основу данного метода криптографии составляет сравнение

$$a^p \equiv a^{p \equiv 1 (\bmod\ p-1)} \equiv a \pmod{p}, \text{ или } a^{x \equiv 1 (\bmod\ p-1)} \equiv 1 \pmod{p},$$

позволяющее, в принципе, произвольно выбирать число абонентов и число ключей у каждого абонента.

2.2. Криптосистема с открытым ключом (*RSA*)

Данный метод не является самостоятельным, так как базируется на предыдущем алгоритме. *RSA – Rivest*, *Shamir*, *Adleman*.

Пусть абоненты A и B хотят наладить между собой секретный обмен данными с открытым ключом. С этой целью каждый из них независимо друг от друга выбирает два больших простых числа ξ_1, ξ_2 и ψ_1, ψ_2 и находит их произведение, равное $\rho_A = \xi_1 \cdot \xi_2$ и $\rho_B = \psi_1 \cdot \psi_2$ соответственно. После этого определяется функция Эйлера от этих чисел:

$$\varphi(\rho_A) = (\xi_1 - 1)(\xi_2 - 1) \text{ и } \varphi(\rho_B) = (\psi_1 - 1)(\psi_2 - 1).$$

Для шифрования сообщений абоненты выбирают свои случайные числа (открытые ключи) из условия

$$(a, \varphi(\rho_A)) = 1, \quad 0 < a < \varphi(\rho_A) = (\xi_1 - 1)(\xi_2 - 1),$$
$$(b, \varphi(\rho_B)) = 1, \quad 0 < b < \varphi(\rho_B) = (\psi_1 - 1)(\psi_2 - 1).$$

Значения a и b, ρ_A и ρ_B носят открытый характер и размещаются в Интернете на заранее оговоренных сайтах.

Далее каждый из абонентов находит свой секретный ключ из сравнений

$$ax \equiv 1 \left(\bmod \varphi(\rho_A)\right) \text{ и } by \equiv 1 \left(\bmod \varphi(\rho_B)\right), \tag{2.5}$$

где $0 < x = \alpha < \varphi(\rho_A),$ а $0 < y = \beta < \varphi(\rho_B)$ – секретные ключи, известные только самим пользователям.

При необходимости посылки информации от абонента A к абоненту B текст разделяется на блоки длиной $0 < \mu < \rho_B$. После чего каждый блок кодиру-

ется в соответствии с правилом

$$\mu_1 \equiv \mu^b \ (\mathrm{mod}\ \rho_B)$$

и передается в сеть.

Абонент B расшифровывает это сообщение, используя свой секретный ключ β:

$$\mu_2 \equiv \mu_1^{\beta} \ (\mathrm{mod}\ \rho_B).$$

Учитывая, что

$$\mu_2 \equiv \mu^{b\beta} \ (\mathrm{mod}\ \rho_B),$$

имеем равенство $\mu = \mu_2$, так как $b\beta \equiv 1 \ (\mathrm{mod}\ \varphi(\rho_B))$ с учетом формулы (2.5).

Отличительной особенностью данного алгоритма по отношению к криптосистеме без передачи ключей является наличие своих модулей $\varphi(\rho_A)$ и $\varphi(\rho_B)$ для передачи информации у каждого пользователя сети.

Другая особенность метода состоит в использовании больших простых чисел ξ_1, ξ_2 и ψ_1, ψ_2 для формирования произведения $\rho_A = \xi_1 \cdot \xi_2$ и $\rho_B = \psi_1 \cdot \psi_2$, что дает возможность выбирать ключи a и b в широком диапазоне значений и, следовательно, существенно затруднить процесс расшифрования.

2.3. Электронная криптографическая подпись

Криптосистемы с открытым ключом обладают существенным недостатком, а именно: получатель шифрованных данных не имеет информации об отправителе, если абонентов-источников $A_1, A_2, ..., A_n$ несколько. Этого недостатка лишена шифросистема с электронной криптографической подписью. Сущность данного метода шифрования заключается в следующем.

Пусть имеется сеть передачи информации (рис. 2.1).

Для организации секретной связи каждый из абонентов A_i и банк данных независимо друг от друга выбирают по два достаточно больших простых числа. Пусть ξ_1 и ξ_2 – простые числа банка, а $\psi_{1,i}, \psi_{2,i}$ – простые числа абонентов A_i. Для всех значений i определяются числа r_i, $i = \overline{1,n}$, как произведения

$$R = \xi_1 \cdot \xi_2, \qquad r_i = \psi_{1,i} \cdot \psi_{2,i}.$$

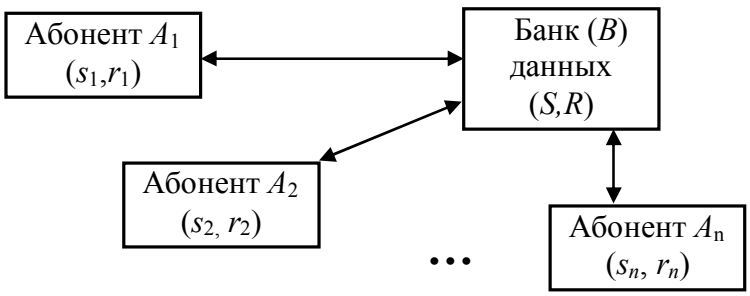

Рис. 2.1

При этом должно выполняться неравенство $R \geq r_i$.

На следующем этапе все участники связи выбирают случайные числа s_i из условия

$$0 < S < \varphi(R), \quad (S, \varphi(R)) = 1,$$
$$0 < s_i < \varphi(r_i), \quad (s_i, \varphi(r_i)) = 1,$$

после чего в Интернете размещается таблица открытых ключей

$$B: \ S, R$$
$$A_1 : s_1, r_1 \quad A_2 : s_2, r_2 \quad ..., \quad A_n : s_n, r_n,$$

которая доступна всем желающим.

Далее банк и каждый из абонентов A_i находят свои секретные ключи T и t_i из условий

$$ST \equiv 1 \pmod{\varphi(R)}, \quad 0 < T < \varphi(R),$$
$$s_i t_i \equiv 1 \pmod{\varphi(r_i)}, \quad 0 < t_i < \varphi(r_i).$$

Теперь пусть абонент A_1 хочет передать информацию μ абоненту B, причем будем считать, что в данном случае выполняется неравенство $R > r_1$, при этом используем значения $\mu < r_1$,где $(\mu, r_1) = 1$. Абонент A_1 шифрует сообщение μ своим первым секретным ключом, а потом открытым ключом банка:

$$\begin{aligned} \mu_1 &\equiv \mu^{t_1} \pmod{r_1}, \quad 0 < \mu_1 < r_1, \\ \mu_2 &\equiv \mu_1^S \pmod{R}, \quad 0 < \mu_2 < R. \end{aligned} \tag{2.6}$$

Абонент B, получив шифрованное сообщение μ_2, расшифровывает его, пользуясь сначала своим секретным ключом T:

$$\mu_3 \equiv \mu_2^T \pmod{R}, \quad 0 < \mu_3 < R, \tag{2.7}$$

а потом открытым ключом s_1 абонента A_1:

$$\mu \equiv \mu_4 \equiv \mu_3^{s_1} \pmod{r_1}, \quad 0 < \mu_4 < r_1.$$

Математическое доказательство тождеств (2.6)–(2.7) следует из ряда соотношений, так как

$$\mu_3 \equiv \mu_2^T \equiv \left(\mu_1^S\right)^T \equiv \mu_1^{ST} \pmod{R}. \tag{2.8}$$

При $(\mu_1, R) = 1$ из (2.8) имеем

$$\mu_1^{ST \equiv 1 \pmod{\varphi(R)}} \equiv \mu_1 \pmod{R}.$$

Следовательно, $\mu_3 \equiv \mu_1 \pmod{R}$. Учитывая же, что $0 < \mu_3 < R$, а также $0 < \mu_3 < r_1$, окончательно имеем $\mu_3 = \mu_1$.

Из сравнений $\mu_4 \equiv \mu_3^{s_1} \equiv \mu_1^{s_1} \equiv \left(\mu^{t_1}\right)^{s_1} \equiv \mu^{s_1 t_1} \pmod{r_1}$ при $(\mu, r_1) = 1$ имеем

$$\mu^{s_1 t_1 \equiv 1 \left(\mathrm{mod}\,\varphi(r_1)\right)} \equiv \mu \pmod{r_1}.$$

Опять же, при $0 < \mu < R$, а также $0 < \mu < r_1$ окончательно имеем $\mu_4 = \mu$. Причем если полученное сообщение принадлежит смысловым конструкциям используемого языка, то считается, что сообщение прислано абонентом A_1.

Пример. Пусть банк данных выбрал свои простые числа 11 и 23, а абонент *A* значения 7 и 13. Таким образом, $r = 7 \cdot 13 = 91$ и $R = 11 \cdot 23 = 253$. Банк и абонент выбирают открытые ключи 31 и 5, а числа 71 и 29 – секретные ключи банка и абонента соответственно. Данный факт доказывается тождествами

$$5 \cdot 29 \equiv 1 \left(\mathrm{mod}\,72\right), \qquad 72 = \varphi\left(7 \cdot 13\right),$$
$$31 \cdot 71 \equiv 1 \left(\mathrm{mod}\,220\right), \qquad 220 = \varphi\left(11 \cdot 23\right).$$

Тогда публикуемая в Интернете таблица будет представлена в виде

$$\textit{Банк данных}: \; S = 31, \quad R = 253.$$
$$\textit{Абонент}: \; s = 5, \; r = 91.$$

Абонент A посылает запрос банку в виде сообщения $\mu = 41$. Замечая, что $R > r$, инициатор запроса кодирует посылаемое слово своим секретным ключом 29, а затем открытым ключом банка 31:

$$\mu_1 \equiv 41^{29} \equiv 6 \pmod{91},$$
$$\mu_2 \equiv 6^{31} \equiv 39 \left(\mathrm{mod}\ 253\right).$$

Банковская система доступа декодирует сообщение, используя сначала свой секретный ключ, а затем открытый ключ абонента:

$$\mu_3 \equiv 39^{71} \equiv 6 \pmod{253},$$
$$\mu_4 \equiv 6^{5} \equiv 41 \left(\mathrm{mod}\ 91\right).$$

Итак, запрос абонента есть число 41.

2.4. Шифросистема Эль-Гамаля

Основу данного способа криптографии составляет малая теорема Ферма, которая используется следующим образом.

Очевидно, что любое сообщение μ может быть представлено в виде

$$\mu \equiv \mu \cdot \alpha^{xy} \left(\alpha^{xy}\right)^{-1} \pmod{p}. \tag{2.9}$$

Введем обозначения в сравнение (2.9) в соответствии с равенствами

$$\alpha^x \equiv \beta \pmod{p}, \; \alpha^y \equiv \gamma \pmod{p},$$
$$(x, p-1) = 1, \qquad (y, p-1) = 1.$$

Тогда $\mu \equiv \mu \cdot \beta^y \left(\gamma^x\right)^{-1} \pmod{p}$.

Теперь, полагая, что $\mu_1 \equiv \mu \cdot \beta^y \pmod{p}$, создаем систему открытых и секретных ключей:

$$\text{открытый ключ} - \ p, \alpha, x,$$
$$\text{секретный ключ} - \ y.$$

В данном случае необходимо, чтобы секретный ключ y был известен как абоненту-источнику, так и приемнику информации. При этом правило обмена данными будет заключаться в выполнении следующих действий.

Абонент-источник выбирает сообщение μ, кодирует его в соответствии с соотношением

$$\mu_1 \equiv \mu \cdot \alpha^{xy} \pmod{p}$$

и предает его в канал связи.

Абонент-приемник принимает слово μ_1 и, используя множитель γ, расшифровывает его в соответствии со сравнением

$$\mu \equiv \mu_1 \left(\gamma^x\right)^{-1} \pmod{p}.$$

Шифросистема Эль-Гамаля обретает криптографическую стойкость, если для каждого слова μ_i выбирается свой секретный ключ $0 < y_i < p-1$, который кодировано передается с шифротекстом $\mu_{1,i}$. Однако длина текста при шифровании в такой системе удваивается, что является существенным недостатком с точки зрения обнаружения самого канала связи.

2.5. Цифровая криптографическая подпись Эль-Гамаля

Цифровая подпись Эль-Гамаля используется для передачи сообщения μ от абонента-источника к абоненту-приемнику, а также установления подлинности источника, пославшего сообщение.

Пусть p – простое число и $0 < \alpha \le p-1$, тогда, выбрав в качестве полусекретного ключа значение $0 < x < p-1$, вычислим

$$\beta \equiv \alpha^x \pmod{p}$$

и разместим в Интернете набор вида (p, α, β).

Подпись для сообщения μ зашифровывается по следующему правилу.

Абонент-источник выбирает случайное целое число $0 < y < p-1$, $(y, p-1) = 1$, являющееся, по сути, переменным секретным ключом, и определяет параметры

$$\gamma \equiv \alpha^y \pmod{p},$$
$$\delta \equiv (\mu - x\gamma) y^{-1} \pmod{p-1}.$$

Параметры (γ, δ) являются пересылаемой информацией, содержащей под-

утверждение или подпись для сообщения μ.

Абонент-приемник, зная полуоткрытый ключ абонента-источника, выполняет следующие действия. Во-первых, получает произведение

$$\beta^{\gamma}\gamma^{\delta} \equiv \beta^{\alpha^y}\gamma^{(\mu-x\gamma)y^{-1}} \equiv \alpha^{x\alpha^y}\alpha^{y\left(\mu-x\alpha^y\right)y^{-1}} \equiv \alpha^{x\alpha^y}\alpha^{\mu-x\alpha^y} \pmod{p}, \qquad (2.10)$$

а из сравнения (2.10) – цифровую подпись вида

$$\beta^{\gamma}\gamma^{\delta} \equiv \alpha^{\mu} \pmod{p}. \qquad (2.11)$$

Во-вторых, зная μ и открытое значение α, легко устанавливает идентичность присланного сообщения и подписи (2.11).

Схема цифровой подписи Эль-Гамаля послужила образцом для построения большого семейства во многом сходных по своим свойствам алгоритмов формирования подписей. В их основе лежит проверка сравнения вида

$$\alpha^{A}\beta^{B} \equiv \gamma^{C} \pmod{p},$$

в котором тройка A,B,C определяется некоторым набором параметров μ,δ,γ при некотором выборе знаков.

Например, исходная схема Эль-Гамаля получается при $A=\mu$, $B=-\gamma$ и $C=\delta$. При этом получаем

$$\alpha^{\mu}\beta^{-\gamma} \equiv \gamma^{\delta} \pmod{p}. \qquad (2.12)$$

Подставляя в (2.12) значения $\beta=\alpha^x$, $\gamma=\alpha^y$, имеем сравнение

$$\alpha^{\mu}\alpha^{-x\cdot\alpha^y} \equiv \alpha^{y\delta} \pmod{p}. \qquad (2.13)$$

При $\delta=\left(\mu-x\gamma\right)y^{-1}$ из (2.13) следует тождество

$$\alpha^{\mu}\alpha^{-x\cdot\alpha^y} \equiv \alpha^{y\left(\mu-x\cdot\alpha^y\right)y^{-1}} \pmod{p},$$

$$\alpha^{\mu-x\cdot\alpha^y} \equiv \alpha^{\mu-x\cdot\alpha^y} \pmod{p}.$$

Таким образом, (2.12) достоверно и $\alpha^{\mu} \equiv \gamma^{\delta}\beta^{\gamma} \pmod{p}$.

На базе схем подписи из этого семейства построены стандарты цифровой криптоподписи США (*DSS – Digital Signature Standard*) и России, что подтверждает эффективность метода на соответствующем классе задач.

Пример. Пусть $\alpha=7$ и $p=29$, таким образом, $0<\alpha\le 28$. Выбираем первый полусекретный ключ $x=14$ и определяем параметр $\beta \equiv \alpha^x \equiv 7^{14} \equiv 1 \pmod{29}$. Значения $p=29, \alpha=7, \beta=1$ помещаются в Интернете.

Далее абонент-источник выбирает второй (полусекретный) ключ $y=23$ таким образом, чтобы выполнялись соотношения

$$0 < y < 28, \quad (y, p-1) = 1 \longrightarrow (23, 28) = 1.$$

После этого определяется первый пересылаемый элемент подписи:

$$\gamma \equiv \alpha^{y} \equiv 7^{23} \equiv 20 \pmod{29}.$$

Выбранное сообщение, например $\mu = 15$, используется для получения второго элемента подписи:

$$\delta \equiv (\mu - x\gamma) y^{-1} \equiv (15 - 14 \cdot 20) 23^{-1} \equiv -\frac{265}{23} \pmod{28}. \qquad (2.14)$$

Умножим обе части (2.14) на 23 и решим сравнение вида

$$23\delta \equiv -265 \equiv 15 \pmod{28}. \qquad (2.15)$$

Разложим отношение $\frac{28}{23}$, используя алгоритм Евклида, и составим таблицу для значений q_s и P_s (табл. 2.4).

Таблица 2.4

	s=0	s=1	s=2	s=3	s=4	s=5
q_s		1	4	1	1	2
P_s	1	1	5	6	11	28

Решением сравнения (2.15) будет соотношение

$$\delta \equiv (-1)^{4} 11 \cdot 15 \equiv 165 \pmod{28}$$

или $\delta \equiv 25 \pmod{28}$.

Итак, парой символов, сопровождающих сообщение $\mu = 15$, будут значения $\gamma = 20,\ \delta = 25$.

Абонент-приемник принимает μ, γ, δ и формирует функции

$$\beta^{\gamma} \gamma^{\delta} \equiv 1 \cdot 20^{25} \equiv 7 \pmod{29},$$
$$\alpha^{\mu} \equiv 7^{15} \equiv 7 \pmod{29}.$$

Так как теоретическое значение α^{μ} и переданное $\beta^{\gamma} \gamma^{\delta}$ совпадают, считаем, что принятое сообщение μ – достоверно.

2.6. Криптографическая подпись Фиат-Шамира

Пусть h – некоторая хеш-функция (или функция формирования сигнатур) преобразует исходное сообщение μ в строку символов длиной l. Для формирования подписи выберем два простых числа ξ_1 и ξ_2 и определим модуль сравнения для составляющих подписи в виде $m = \xi_1 \cdot \xi_2$.

В качестве секретного ключа каждый абонент должен сгенерировать l различных случайных чисел $a_0, a_1, ..., a_{l-1} > 0, \quad (a_i, m) = 1,$ а в качестве открытого ключа значения

$$b_i \equiv \left(a_i^{-1}\right)^2 \pmod{m}, \qquad i = 0,1,...,l-1. \tag{2.16}$$

Тогда алгоритм вычисления криптографической подписи для сообщения μ будет состоять в следующем.

1. Выбираем случайное число α, $0 < \alpha \le m-1$. Например, полагая значения $\xi_1 = 3, \xi_2 = 41,$ имеем $m = 123$ и можем выбрать $\alpha = 33.$

2. Вычисляем значение: $\beta \equiv \alpha^2 \pmod{m}$. Для нашего примера это $\beta = 105.$

3. Вычисляем хеш-функцию для сообщения μ вида $h(\mu, \beta) = s = s_0, s_1, ..., s_{l-1}$. Значение s используем как первый элемент подписи.

Пусть, например, последовательность бит исходного сообщения имеет вид

$$\mu = \{1,1,1,0,0,1,1,1,1,0,0,0,1,0,0,0\},$$

$$\beta = \{1101001\}.$$

Тогда значение сигнатуры определим с помощью функций счета:

$$h(\mu, \beta) = 12 = 1100_2.$$

4. Выбираем ряд случайных чисел $a_0 = 4, a_1 = 8, a_2 = 14, a_3 = 16,$ больших нуля, и находим второй элемент подписи по формуле

$$t \equiv \alpha \prod_{i=0}^{l-1} a_i^{s_i} \pmod{m}.$$

Для нашего примера это число $t \equiv 33 \cdot 4^1 \cdot 8^1 \cdot 14^0 \cdot 16^0 \equiv 72 \pmod{123}$. Открытый ключ (2.16) при заданных a_i будет равен $b_0 = 100, b_1 = 25, b_2 = 91, b_3 = 37$.

Итак, криптографической подписью для сообщения μ следует считать пару чисел $(s,t) = (12,72).$

Алгоритм проверки подписи состоит в выполнении следующих действий.

1. По открытому ключу $b_0, b_1, ..., b_{l-1} \pmod{m}$ и значению t определяется параметр

$$w \equiv t^2 \prod_{i=0}^{l-1} b_i^{s_i} \pmod{m},$$

который при $b_0 = 100, b_1 = 25, b_2 = 91, b_3 = 37$ и $t = 72$ будет равен 105.

2. Определяется значение хеш-функции $h(\mu, w)$:

$$\mu \cup w = \{1,1,1,0,0,1,1,1,1,0,0,0,1,0,0,0,\ \ 1,1,0,1,0,0,1\} \to 12.$$

3 Сравнивается переданное по информационному каналу $h(\mu, \beta)$ и расчетное значение $h(\mu, w)$. Совпадение кодов свидетельствует о достоверности передачи данных.

2.7. Классификация алгоритмов шифрования

Первичным признаком, по которому производится классификация шифров, является тип преобразования, осуществляемого с открытым текстом при шифровании. Например, если фрагменты открытого текста на языке A (отдельные буквы или группы букв) заменяются некоторыми эквивалентами из подмножества символов другого языка B, то соответствующий шифр относится к классу *шифров замены*. Если буквы открытого текста при шифровании лишь меняются местами в рамках собственного алфавита, то мы имеем дело с *шифром перестановки*.

С целью повышения надежности систем передачи информации шифрованный текст, засекреченный некоторым алгоритмом, может быть еще раз зашифрован с использованием другого шифра. Всевозможные такие композиции алгоритмов объединяются в третий класс, который обычно называют композиционными шифрами. При этом сами вновь создаваемые методы могут быть организованы так, что не будут входить ни в первый, ни во второй классы шифров. Таким образом, первый уровень классификации алгоритмов шифрования может быть представлен в следующем виде (рис. 2.2).

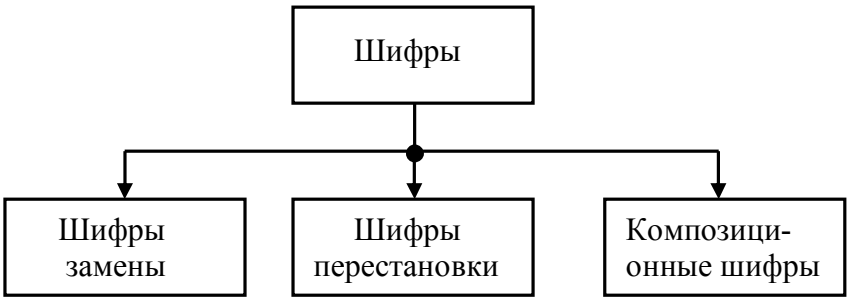

Рис. 2.2

2.8. Математическая модель шифра замены

Пусть $A=\{\mu_1, \mu_2,\ldots,\mu_n\}$ – словосочетания или буквы открытого текста передающей системы, $B=\{\lambda_1,\lambda_2,\ldots,\lambda_m\}$ – словосочетания или буквы алфавита шифрования (шифротекста), X – открытый текст сообщения, Y – шифрованное сообщение. Для шифрования текста X выбирается подмножество символов $B^* \in B$, при этом сам текст X является подмножеством символов $A^* \in A$. Таким образом, $X \in A^*$, $Y \in B^*$.

Перед кодированием открытый текст представляется в виде последовательности подслов, называемых шифровеличинами. При шифровании шифровеличины заменяются некоторыми эквивалентами из шифротекста, которые называют шифрообозначениями. И те и другие символы принадлежат элементам A^* и B^*.

Выберем в качестве конкретной реализации передаваемого по сети сооб-

щения (иными словами множество шифровеличин) последовательность символов $A^* = \{\mu_h, \mu_j, \ldots, \mu_k\}$, а в качестве шифрообозначений $B^* = \{\lambda_r, \lambda_s, \ldots, \lambda_l\}$.

Для определения правила шифрования $E_s(\mu)$ в общем случае вводится ряд обозначений и понятие распределителя, который, по сути, будет осуществлять выбор в каждом такте кодирования замену из алфавита B^*, соответствующей шифровеличине из A^*.

Представим множество B^* в виде объединения непустых подмножеств $B^*_{\alpha,i}$:

$$B^* = \bigcup_{i=1}^{w} B^*_{\alpha,i}, \qquad B^*_{\alpha,i} \neq \varnothing.$$

Далее каждое λ_z будем считать членом семейства $B^*_{\alpha,i} = \bigcup_{z=1}^{r} \lambda_z$, $r < n$, что позволяет процесс шифрования представить в виде биекции множества событий A^* на $B^*_{\alpha,i}$ в виде

$$A^* \xrightarrow{\varphi_\alpha(\mu)} \left\{B^*_{\alpha,1}, B^*_{\alpha,2}, \ldots, B^*_{\alpha,w}\right\},$$

где $\varphi_\alpha(\mu_i) = B^*_{\alpha,i}$ – распределитель шифра замены. При этом $B^*_{\alpha,i}$ могут быть пересекающимися подмножествами.

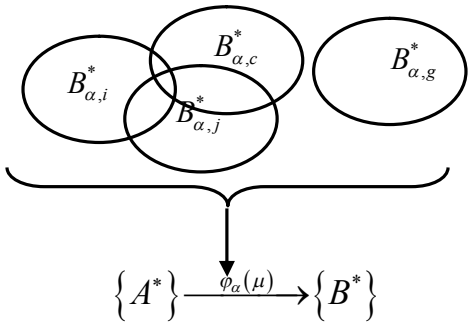

Рис. 2.3

При этом под биективным отображением понимают такое преобразование, при котором элементы множества $\{X\}$ преобразуются в элементы большего множества $\{Y\}$, а обратная функция однозначно приводит к элементам начально-исходного множества $\{X\}$. Функция преобразования при этом имеет вид

$$\{X\} \xrightarrow{\varphi} \{Y\}, \quad \{Y\} \xrightarrow{\varphi^{-1}} \{X\}.$$

Пример. $\{X\} = \{4, 9, 36, 64, 81\}$, $\varphi = squirt$. $\{Y\} = \{\pm 2, \pm 3, \pm 6, \pm 8, \pm 9\}$.

Под инъективным отображением понимают такое преобразование, при котором элементы множества $\{X\}$ однозначно преобразуются в элементы множества $\{Y\}$, а обратная функция также однозначно приводит к элементам ис-

ходно-начального множества $\{X\}$. Функция преобразования в данном случае имеет вид

$$\{X\} \xrightarrow{\varphi} \{Y\}, \quad \{Y\} \xrightarrow{\varphi^{-1}} \{X\}.$$

Пример. $\{X\} = \{2,8,4,15,16\}$, $\varphi = \times 3$. $\{Y\} = \{6,24,12,45,48\}$.

И, наконец, под сюрьективным отображением понимают такое преобразование, при котором элементы множества $\{X\}$ преобразуются в элементы большего множества $\{Y\}$, а функция преобразования имеет вид $\{X\} \xrightarrow{\varphi} \{Y\}$.

Пример. $\{X\} = \{4, 9,36,64,81\}$, $\varphi = C_n^x$. $\varphi^{-1} = \left\{\left[C_n^x\right]^{-1}, \left[C_n^{n-x}\right]^{-1}\right\}$.

Следует также различать такое понятие, как инъективное отображение множества $\{X\}$ в множество $\{Y\}$. При этом элементы $\{X\}$ однозначно преобразуются в элементы меньшего множества $\{Y\}$, а обратное преобразование однозначно дает множество $\{X\}$.

Математическая модель шифра замены может быть представлена на основании отображения ψ, такого, что $\mu \to C(n \times m)$, где $C = \{c_{i,j}\}$, $(i \le n,\ j \le m)$ – символы алфавита в плоскости отображения. Тогда последовательность действий $\psi(\mu = a_i \cup b_j)$ будет называться распределителем, отвечающим значениям i и j на плоскости символов замены ***C***. При этом: $\psi(a_i, b_j) = c_{i,j}$, где a_i, b_j образуют координаты плоскости.

В общем случае использование шифров замены может базироваться на нескольких преобразованиях с использованием нескольких алгоритмов и алфавитов. Общая схема замены в этом случае будет образовывать некоторый пространственный объект с межслойной передачей данных, образованных распределителями предыдущих шагов шифрования.

2.9. Классификация шифров замены

Если ключ зашифровывания совпадает с ключом расшифровывания $K_з = K_p$, то такие шифры называют симметричными, при $K_з \ne K_p$ – асимметричными. Соответственно, данный факт и дает следующий шаг классификации.

Правило шифрования в первом случае записывают в виде

$$\mu_1 = E_s(\mu), \quad \mu = E_s(\mu_1),$$

а во втором

$$\mu_1 = E_s(\mu), \quad \mu \ne E_s(\mu_1).$$

Однозначные шифры замены базируются на свойстве однозначной замены элементов множества $\{X\}$ элементами множества $\{Y\}$, то есть на инъективном преобразовании. Для многозначных же шифров правило шифрования носит обычно характер биективного отображения.

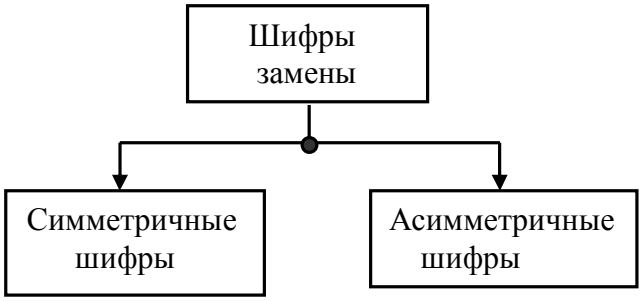

Рис. 2.4

Наибольшее распространение на практике получили шифры однозначной замены и, в частности, шифры гаммирования. Сущность данного метода состоит в образовании нового множества $\{Y\}$ путем «сложения» в той или иной форме элементов множества $\{X\}$ и элементов гаммы, представляющих собой некоторую ключевую последовательность.

К шифрам замены относятся также поточные и блочные алгоритмы, сущность которых заключается в следующем.

Поточные шифры представляют собой двухблочную систему шифрования, в которой первый блок вырабатывает последовательность номеров шифрующих преобразований, а второй – замену знаков открытого текста. Используется, как правило, для кодирования и передачи больших массивов информации.

Блочные шифры отличаются от поточных мощностью используемых алфавитов, а также способом деления информации. В данном случае открытый текст разбивается на блоки и каждый блок шифруется своим алгоритмом.

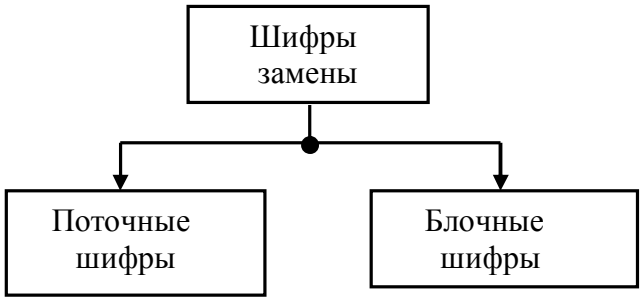

Рис. 2.5

В зависимости от числа алфавитов шифры замены бывают также одноалфавитные – шифры простой замены, и многоалфавитные.

В общем случае на практике трудно провести границу между реально используемыми шифрами. Часто вновь создаваемые криптографические методы обладают различными свойствами частных алгоритмов, подпадающих под признаки той или иной классификации. В целом же приведенная система деления шифров на группы имеет целью познакомить разработчиков аппаратуры и алгоритмов со всем многообразием шифров и методов кодирования, применяемых в криптографии.

2.10. Шифры перестановки

Широкое распространение на практике получили так называемые маршрутные перестановки, основанные на использовании некоторых геометрических фигур. Фрагмент открытого текста записывается в такую фигуру по некоторой траектории, а шифрованным текстом является последовательность, полученная при считывании текста по другой траектории.

Пример. Текст: «пример маршрутной перестановки». Для шифрования используем прямоугольную таблицу размером 4×7 (табл. 2.5). Запись фразы выпоняется по строкам слева направо и справа налево. Считывание выполняется по столбцам начиная с левого верхнего угла вниз по столбцу и далее вверх по соседнему справа столбцу; затем вниз и т. д. В результате получаем фразу кодированного сообщения, с трудом поддающуюся раскодированию без знания таблицы: «пноикйтр иупвоерм ешрнаерр маст».

Таблица 2.5

п	р	и	м	е	р	м
н	т	у	р	ш	р	а
о	й	п	е	р	е	с
и	к	в	о	н	а	т

Другой класс алгоритмов перестановок базируется на использовании различных решеток и таблиц. Примером такого алгоритма следует считать классический алгоритм перестановок Кардано. При этом используется решетка, приведенная на рис. 2.6. Кодирование выполняется путем 4- кратного поворота системы и записи букв в свободные шифроклетки.

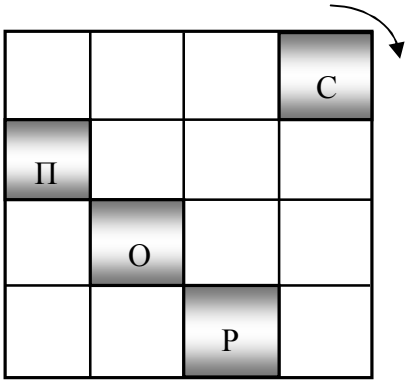

Рис. 2.6

Пример. Пусть имеется текст открытого сообщения: «спортлото это игра».

Кодированное сообщение, полученное с помощью алгоритма Кордано, имеет вид «иотсплэг оортоарт».

В практических целях может использоваться более сложная решетка, имеющая в своем составе 25, 36 … клеток.

2.11. Композиционный шифр в блочной системе шифрования (*DES*)

Одним из методов, получивших широкое распространение в системах блочного кодирования информации, является метод сетей Фейстеля. Он основывается на использовании специального регистра сдвига с элементами меж-

разрядной логики, показанными на рис. 2.7.

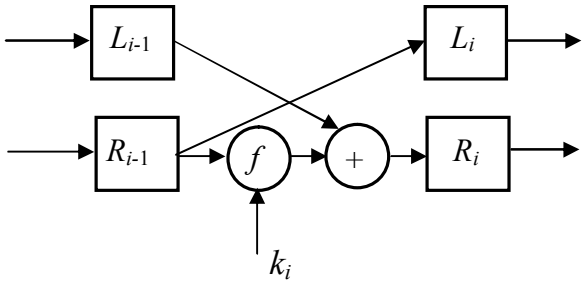

Рис. 2.7

При сдвиге информации слева направо в схеме выполняются следующие преобразования:

$$\begin{aligned} L_i &= R_{i-1}, \\ R_i &= L_{i-1} \oplus f_i\left(k_i, R_{i-1}\right), \end{aligned} \tag{2.17}$$

где $f_i\left(k_i, R_{i-1}\right)$ – произвольная функция от двух аргументов: случайного ключа k_i и значения предыдущего разряда регистра R_i.

Достоинством преобразований данного вида является возможность обратного преобразования (2.17), даже в случае если внутренняя функция $f_i\left(k_i, R_{i-1}\right)$ не является обратимой. Действительно из базового равенства следует, что

$$\begin{aligned} R_{i-1} &= L_i, \\ L_{i-1} &= R_i \oplus f_i\left(k_i, R_{i-1}\right), \end{aligned}$$

то есть выходные символы ячейки Фейстеля могут быть поданы на свои же входы, а на выходах считаны начально-исходные значения. Данный факт следует из рассуждений. Пусть на вход устройства поданы значения $R_{i-1} = a,\ L_{i-1} = b.$ Тогда на выходе схемы получим

$$\begin{aligned} L_i &= a, \\ R_i &= b \oplus f_i\left(k_i, a\right). \end{aligned} \tag{2.18}$$

Подавая на вход данного блока значения (2.18), на выходе получаем результат, представленный на рис. 2.8.

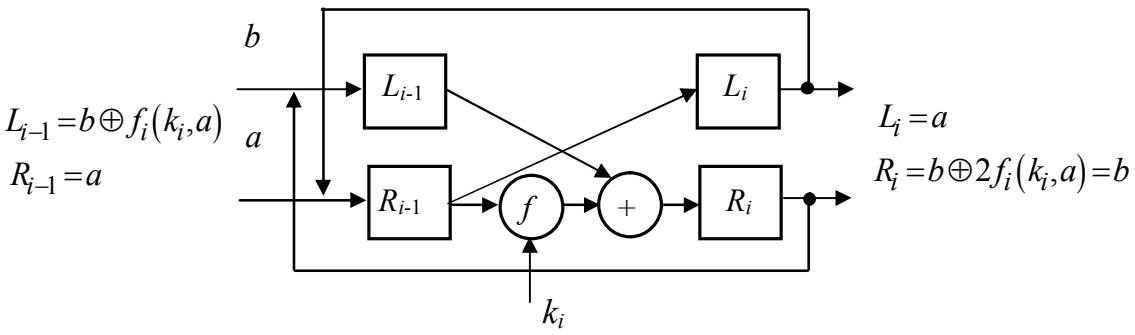

Рис. 2.8

При использовании нескольких ячеек данной схемы процедура декодирования будет заключаться в подаче на вход блока шифрованной последовательности при условии изменения порядка следования ключей на противоположный. Таким образом, шифр является симметричным. Следует учитывать также, что нечетное число преобразований в схеме предполагает перестановку входных переменных относительно выходных, а четное – прямую подачу символов.

Достоинство *DES* заключается в простоте системы, высокой скорости аппаратной и программной реализации, в достаточно высокой криптографической стойкости по отношению к другим методам.

Алгоритм *DES*, используя принцип перестановки аргументов и подстановки последовательности ключей, осуществляет шифрование 64-битных слов с помощью 31-битного ключа k. Процесс шифрования состоит в начальной перестановке бит входного 64-разрядного блока в соответствии со стандартными таблицами в требуемом числе циклов шифрования и конечной перестановке бит выходного слова.

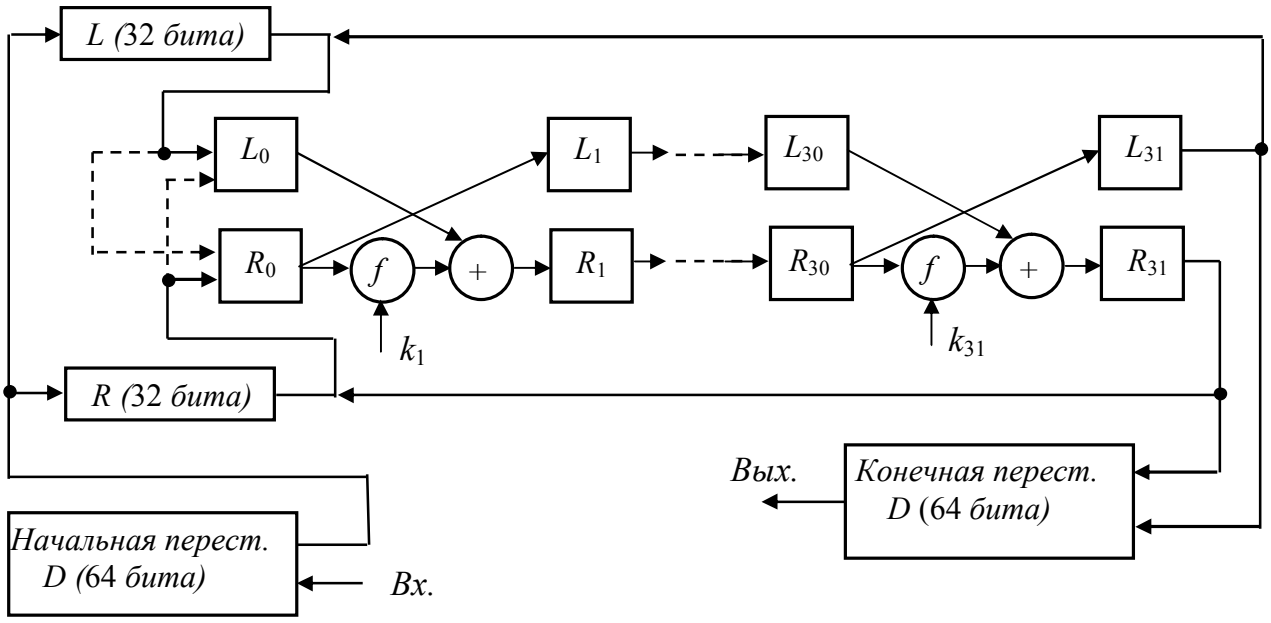

Рис. 2.9

К недостаткам алгоритма *DES* относится небольшое число ключей в стандартной схеме, что дает возможность их полного перебора с помощью компьютера за реальное время.

2.12. Векторно-матричный симметричный шифр замены

Известный из комбинаторной теории принцип включения и исключения для свойств дискретных объектов в известной общей форме позволяет в некотором множестве N_l^0 (принципиально $N_l^0=n$) векторов определить количество

$N_{l-i}(\overline{x}_{u,h},\overline{x}_{u,b},...,\overline{x}_{u,v})$ элементов, не обладающих ни одним из заданных i свойств $x_{u,h},x_{u,b},...,x_{u,v}$, по формуле

$$N_{l-i}(\overline{x}_{u,h},\overline{x}_{u,b},...,\overline{x}_{u,v}) = N_l^0 + \sum_{j=1}^{2^i-1}(-1)^s N_s(x_{u,l-1}^{\sigma_{u,l-1}},x_{u,l-2}^{\sigma_{u,l-2}},...,x_{u,0}^{\sigma_{u,0}}), \quad (2.19)$$

$$j = 2^0\sigma_{u,h} + 2^1\sigma_{u,b} + ... + 2^{i-1}\sigma_{u,v},\ \ \sigma_{u,\lambda} = \text{var}:\ \ \lambda \in \{h,b,...,v\}\ \ ,$$

$$s = \sum_{\lambda}\sigma_{u,\lambda}\ \ ,\ \ \sigma_{u,\lambda} \in \{0,1\},$$

где s – число (0,1)-свойств, характеризующих подмножество N_s; наборы свойств $x_{u,h},...,x_{u,v}$, $x_{u,h}x_{u,b},...,x_{u,b}x_{u,v}$, … $x_{u,h}x_{u,b}...x_{u,v}$, ... в подмножествах N_S определяются на множестве исключаемых элементов, причем $x_{u,l-1},...,x_{u,0}$ – это полный набор (или l-свойств) объектов множества N_l^0. Общий индекс *u* определяет принадлежность векторов форме записи (2.19).

Очевидно, что, в случае, когда число объектов, обладающих s свойствами, $s = \overline{1,i}$, оказывается равным для каждого из значений s, т. е.

$$N_1(x_{u,h}) = N_1(x_{u,b}) = ... = N_1(x_{u,v}),$$
$$N_2(x_{u,h},x_{u,b}) = N_2(x_{u,b},x_{u,v}) = ... = N_2(x_{u,h},x_{u,v}),\ \ldots$$

из соотношения (2.19) следует

$$N_{l-i}(\overline{x}_{u,h},\overline{x}_{u,b},...,\overline{x}_{u,v}) = N_l^0 + \sum_{s=1}^{i}(-1)^s C_i^s N_s\ . \quad (2.20)$$

Разделим равенство (2.20) на величину $N_l^0 \to \infty$. Тогда отношение подмножеств вида

$$p_s = C_i^s\frac{N_s}{N_l^0}, \quad (2.21)$$

может быть интерпретировано как вероятность встретить объект с набором *s* свойств во множестве символов N_l^0, а равенство (2.20) учетом (2.21) даст соотношение

$$p_{l-i}(\overline{x}_{u,h},\overline{x}_{u,b},...,\overline{x}_{u,v}) = 1 + \sum_{s=1}^{i}(-1)^s p_s\ .$$

Известно, что в теории вероятностей данное равенство называют теоремой Пуанкаре. При этом используют запись

$$p_{l-i}(\overline{x}_{u,h},\overline{x}_{u,b},...,\overline{x}_{u,v}) = \prod_{\lambda=0}^{i-1}\left[1 - p_1(x_{u,\lambda})\right]. \quad (2.22)$$

В приведенных выше соотношениях предполагается, что

$$p_S = p_S(x_{u,s-1},...,x_{u,0}) = p_1(x_{u,s-1})\cdot ...\cdot p_1(x_{u,1})p_1(x_{u,0}),$$

где $s \in \{1,2,...,l\}$. То есть все вероятности $p_1\left(x_{u,l-1}\right),\ldots,p_1\left(x_{u,0}\right)$ независимы между собой.

С другой стороны, из теории вероятностных машин известно, что соотношение, эквивалентное (2.22) порождается функциональным элементом ИЛИ-НЕ в процессе сложения i некоррелированных вероятностей. Данный факт следует из правила де Моргана:

$$\prod_{\lambda=0}^{i-1}\left[1-p_1(x_{u,\lambda})\right] \Rightarrow \bigwedge_{\lambda=0}^{i-1} p_1(\overline{x}_{u,\lambda}) = not\left[\bigvee_{\lambda=0}^{i-1} p_1(x_{u,\lambda})\right],$$

где знаки $\prod$ и $\wedge$ означают произведение численных значений вероятностей и логическое произведение вероятностных последовательностей, знак $\vee$ – логическую сумму последовательностей или вероятностей бернуллиевских случайных событий.

Используя форму записи (2.22) для целочисленных объектов, можно записать равенство

$$N_{l-i}(\overline{x}_{u,h},\overline{x}_{u,b},...,\overline{x}_{u,v}) = \bigcap_{\lambda=0}^{i-1}\left[N_l^0 - N_1(x_{u,\lambda})\right], \qquad (2.23)$$

где

$$N_l^0 \cap N_1(x_{u,\lambda}) = N_1(x_{u,\lambda}).$$

Таким образом, раскрывая в соотношении (2.23) скобки и полагая, что все подмножества с s свойствами, имеют одинаковую мощность, от равенства (2.23) легко перейти к равенству (2.20), что подтверждает логику перехода от подмножеств к вероятностям.

Однако переход от подмножеств объектов (2.19) в область теории вероятностей (2.22) и обратно указывает на известную частность рассматриваемого классического принципа. Дело в том, что номенклатура реализуемых функций в вычислительной технике весьма велика, и, как следствие, велико количество алгоритмов, применяемых для синтеза вероятностей. Следовательно, преобразование подмножеств, получившее в литературе название принципа включения и исключения, очевидно, представляет собой только частное преобразование комбинаторной теории и требует дополнительных исследований с точки зрения прикладных вопросов.

Идентичность формы записи равенства Пуанкаре и полиномиального представления булевых функций позволяет предположить, что принцип включения и исключения для подмножеств, равно как и функции двоичной логики, может быть представлен несколькими тождественными формами или соотношениями. Данную гипотезу подтверждает и известная универсальность диаграмм Вена, интерпретирующих булево-алгебраические высказывания, вероятностные соотношения и операции преобразования множеств. Очевидная параллель между различными математическими категориями предполагает существование доказательств тождественности различных форм и для формул включения и исключения. В лучшем

же случае могут быть получены и доказаны соотношения, устанавливающие аналитическую взаимосвязь между различными формами данной методологии.

Для доказательства многовариантности представления рассматриваемого алгоритма запишем равенство, определяющее число векторов N_l^0 как сумму подмножеств N_{l-i}, образуемых из векторов свойств таблицы, аналогичной таблице истинности, умноженных на (0,1)-коэффициенты a_z:

$$N_l^0 = \sum_{z=0}^{2^l-1} a_z N_{l-i}\left[(\overline{x}_{g,l-1} \oplus \sigma_{g,l-1})(\overline{x}_{g,l-2} \oplus \sigma_{g,l-2})...(\overline{x}_{g,0} \oplus \sigma_{g,0})\right], \quad (2.24)$$

где $\sigma_{g,\lambda}, a_z \in \{0,1\}$, $z = 2^{l-1}\sigma_{g,l-1} + 2^{l-2}\sigma_{g,l-2} + ... + 2^0\sigma_{g,0}$, $i = \overline{\sigma}_{g,l-1} + \overline{\sigma}_{g,l-2} + ... + \overline{\sigma}_{g,0}$, а значения переменных

$$\overline{x}_{g,\lambda} \oplus \sigma_{g,\lambda} = \begin{cases} x_{g,\lambda}, \sigma_{g,\lambda} = 1, \\ \overline{x}_{g,\lambda}, \sigma_{g,\lambda} = 0. \end{cases}$$

При этом g указывает на принадлежность векторов из подмножеств N_{l-i} булево-алгебраической форме «включения и исключения», а a_z, как уже говорилось, конституенты таблицы истинности СДНФ, например табл. 2.6, где приведены только подмножества N_{l-i}, соответствующие единичным значениям конституент и, в соответствии с формулой (2.24), дающие сумму, равную $N_l^0 = 68$.

Таблица 2.6

z	$\overline{x}_{g,3} \oplus \sigma_3$	$\overline{x}_{g,2} \oplus \sigma_2$	$\overline{x}_{g,1} \oplus \sigma_1$	$\overline{x}_{g,0} \oplus \sigma_0$	$a_z=1$	N_{l-i}	$\dot{x}_{g,\lambda}$
0	$\overline{x}_{g,3} \oplus 0$	$\overline{x}_{g,2} \oplus 0$	$\overline{x}_{g,1} \oplus 0$	$\overline{x}_{g,0} \oplus 0$	a_0	N_0=10	$\overline{x}_{g,3}\overline{x}_{g,2}\overline{x}_{g,1}\overline{x}_{g,0}$
5	$\overline{x}_{g,3} \oplus 0$	$\overline{x}_{g,2} \oplus 1$	$\overline{x}_{g,1} \oplus 0$	$\overline{x}_{g,0} \oplus 1$	a_5	N_2=7	$\overline{x}_{g,3}x_{g,2}\overline{x}_{g,1}x_{g,0}$
8	$\overline{x}_{g,3} \oplus 1$	$\overline{x}_{g,2} \oplus 0$	$\overline{x}_{g,1} \oplus 0$	$\overline{x}_{g,0} \oplus 0$	a_8	N_1=13	$x_{g,3}\overline{x}_{g,2}\overline{x}_{g,1}\overline{x}_{g,0}$
10	$\overline{x}_{g,3} \oplus 1$	$\overline{x}_{g,2} \oplus 0$	$\overline{x}_{g,1} \oplus 1$	$\overline{x}_{g,0} \oplus 0$	a_{10}	N_2=4	$x_{g,3}\overline{x}_{g,2}x_{g,1}\overline{x}_{g,0}$
12	$\overline{x}_{g,3} \oplus 1$	$\overline{x}_{g,2} \oplus 1$	$\overline{x}_{g,1} \oplus 0$	$\overline{x}_{g,0} \oplus 0$	a_{12}	N_2=10	$x_{g,3}x_{g,2}\overline{x}_{g,1}\overline{x}_{g,0}$
13	$\overline{x}_{g,3} \oplus 1$	$\overline{x}_{g,2} \oplus 1$	$\overline{x}_{g,1} \oplus 0$	$\overline{x}_{g,0} \oplus 1$	a_{13}	N_3=7	$x_{g,3}x_{g,2}\overline{x}_{g,1}x_{g,0}$
14	$\overline{x}_{g,3} \oplus 1$	$\overline{x}_{g,2} \oplus 1$	$\overline{x}_{g,1} \oplus 1$	$\overline{x}_{g,2} \oplus 0$	a_{14}	N_3=13	$x_{g,3}x_{g,2}x_{g,1}\overline{x}_{g,0}$
15	$\overline{x}_{g,3} \oplus 1$	$\overline{x}_{g,2} \oplus 1$	$\overline{x}_{g,1} \oplus 1$	$\overline{x}_{g,0} \oplus 1$	a_{15}	N_4=4	$x_{g,3}x_{g,2}x_{g,1}x_{g,0}$

С другой стороны, используя правую часть соотношения (2.19) и полиномиальное представление булевых функций, можно записать равенство

$$N_l^0 = c_0 N_0 + \sum_{j=1}^{2^l-1} (-1)^{1+s} c_j N_s \left(x_{u,l-1}^{\sigma_{u,l-1}}, x_{u,l-2}^{\sigma_{u,l-2}}, ..., x_{u,1}^{\sigma_{u,1}}, x_{u,0}^{\sigma_{u,0}} \right). \quad (2.25)$$

Здесь $j = 2^{l-1}\sigma_{u,l-1} + 2^{l-2}\sigma_{u,l-2} + ... + 2^1\sigma_{u,1} + 2^0\sigma_{u,0}$, а $s = \sum_{\lambda=0}^{l-1} \sigma_{u,\lambda}$, $\sigma_{u,\lambda} \in \{0,1\}$.

В равенстве (2.25) коэффициенты c_j представляют собой целочисленные значения, $c_j = 0, \pm 1$, и определяют характер включения и исключения векторов каждого из подмножеств N_s в множество N_l^0, индекс u – указывает на принадлежность объектов в N_s полиномиальной форме, N_0 – элементы вида $\overline{x}_0, \overline{x}_1, ..., \overline{x}_l$.

Учитывая, что закон следования индексов j в (2.25) соответствует принципу формирования функций счета, наборы коэффициентов $\sigma_{u,l-1}\sigma_{u,l-2},...,\sigma_{u,0}$ оказываются идентичны состояниям l-разрядного двоичного счетчика в моменты времени $t = j$. Таким образом, число свойств $x_{u,\lambda}$, характеризующих каждое из подмножеств

$$N_s\left(x_{u,l-1}^{\sigma_{u,l-1}}, x_{u,l-2}^{\sigma_{u,l-2}}, ..., x_{u,0}^{\sigma_{u,0}}\right),$$

будет всегда равно сумме единиц в двоичном разложении числа j и не превысит величины l, т. е. $x_{u,\lambda}^{\sigma_{u,\lambda}=1} = x_{u,\lambda}$, а $x_{u,\lambda}^{\sigma_{u,\lambda}=0} = 1$, например, при $j = 5$:

$$N_2\left(x_{u,l-1}^{\sigma_{u,l-1}}, x_{u,l-2}^{\sigma_{u,l-2}}, ..., x_{u,0}^{\sigma_{u,0}}\right) = N_2\left(x_{u,2}, x_{u,0}\right).$$

Очевидно, если в равенстве (2.25) все коэффициенты $c_j = 1$, $j = \overline{0, 2^l - 1}$, то при $0 < i \le l-1$ из данного соотношения следует тождество, аналогичное (2.19):

$$N_l^0 - N_{l\text{-}i}\left(\overline{x}_{u,i\text{-}1}, \overline{x}_{u,i\text{-}2}, ..., \overline{x}_{u,0}\right) = \sum_{j=1}^{2^i\text{-}1} (\text{-}1)^{1+\sum_{\lambda=0}^{i\text{-}1}\sigma_\lambda} N_s\left(x_{u,l\text{-}1}^{\sigma_{u,l\text{-}1}}, x_{u,l\text{-}2}^{\sigma_{u,l\text{-}2}}, ..., x_{u,0}^{\sigma_{u,0}}\right),$$
$$j = 2^{i-1}\sigma_{u,i-1} + 2^{i-2}\sigma_{u,i-2} + ... + 2^1\sigma_{u,1} + 2^0\sigma_{u,0},$$
$$\sigma_{u,\lambda} = 1, \quad \lambda \in \{i, i+1, ..., r\}.$$

Если же в соотношении (2.25) коэффициенты c_j будут представлять собой произвольно взятый набор нулей и единиц, то принцип включения и исключения будет порождать новый закон формирования подмножеств векторов на базе общей формы алгоритма суммирования.

Определим зависимости, связывающие коэффициенты двух рассмотренных форм суммирования. С этой целью исследуем случай формирования подмножеств, когда набор коэффициентов c_j априорно неизвестен, а на входы схемы суммирования, отождествляющей преобразования (2.25), коммутируются подмножества объектов в соответствии с естественным порядком следования индексов j. При этом на выходе схемы суммирования (без реализации функции запоминания, т. е. без регистра памяти) будем получать следующие частичные суммы:

$$N_l^0(j=0)=c_0N_0,$$
$$N_l^0(j=1)=c_0N_0+c_1N_1(x_{u,0}),$$
$$N_l^0(j=2)=c_0N_0+c_2N_1(x_{u,1}),$$
$$N_l^0(j=3)=c_0N_0+c_1N_1(x_{u,0})+c_2N_1(x_{u,1})-c_3N_2(x_{u,1},x_{u,0}),$$
$$N_l^0(j=4)=c_0N_0+c_4N_1(x_{u,2}),$$
$$N_l^0(j=5)=c_0N_0+c_1N_1(x_{u,0})+c_4N_1(x_{u,2})-c_5N_2(x_{u,2},x_{u,0}),$$

В общем случае можно показать, что для произвольного значения j будет иметь место равенство

$$N_l^0(j)=c_0N_0+\sum_{q=1}^{j}(-1)^{s+1}\left[\sum_{i=1}^{q}C_q^i\,a_i(\bmod 2)\right]N_s\left(x_{u,l-1}^{\sigma_{u,l-1}},x_{u,l-2}^{\sigma_{u,l-2}},\ldots,x_{u,0}^{\sigma_{u,0}}\right),\qquad(2.26)$$

где $q=\sum_{\lambda=0}^{l-1}\sigma_{u,\lambda}2^{\lambda}$, $j=\overline{1,M-1}$, $M=2^l$.

Очевидно, что левая часть равенства (2.24) и соотношение (2.26) при $j=2^l-1$ будут численно совпадать и давать сумму N_l^0. Следовательно, для правых частей (2.24) и (2.26) также будет выполняться равенство

$$\sum_{z=0}^{2^l-1}a_zN_{l-i}\left[\left(\bar{x}_{g,l-1}\oplus\sigma_{g,l-1}\right)\left(\bar{x}_{g,l-2}\oplus\sigma_{g,l-2}\right)\ldots\left(\bar{x}_{g,0}\oplus\sigma_{g,0}\right)\right]=\qquad(2.27)$$
$$=c_0N_0+\sum_{q=1}^{M-1}(-1)^{s+1}\left[\sum_{i=1}^{q}C_q^i\,a_i(\bmod 2)\right]N_s\left(x_{u,l-1}^{\sigma_{u,l-1}},x_{u,l-2}^{\sigma_{u,l-2}},\ldots,x_{u,0}^{\sigma_{u,0}}\right).$$

Взаимосвязь коэффициентов c_q и a_z в формуле (2.27) может быть доказана следующим образом. Очевидно, что применительно к исследуемому принципу задача получения формы представления значений N_l^0 в виде (2.25) из подмножеств (2.24) (или наоборот) может быть решена путем перегруппировки подмножеств в исходной функции в новые подмножества формируемого набора векторов. При этом для простоты будем считать, что мощность подмножеств в (2.25) для всех N_{l-i} равна 1, что позволяет искомую перегруппировку свести к преобразованиям над коэффициентами, а соотношение (2.26) переписать в виде

$$F(a_0,a_1,\ldots,a_{M-1})\equiv\sum_{q=1}^{z}(-1)^{s+1}C_z^q\cdot c_q\quad(\bmod 2).\qquad(2.28)$$

Для функции $F(a_0,a_1,\ldots,a_{M-1})=F(a_z)=a_z$ будет справедлива следующая теорема.

***Теорема* 2.1.** Коэффициенты $a_z\in\{0,1\}$ и $c_q\in\{0,\pm1\}$ связаны между собой биномиальным преобразованием над $GF(2)$:

$$a_0\equiv c_0,\ a_z\equiv\sum_{q=1}^{z}(-1)^{s+1}C_z^q\cdot c_q\quad(\bmod 2),\ z=\overline{1,M-1},\qquad(2.29)$$

обратная взаимосвязь определяется взаимно-обратным равенством.

Доказательство (2.28) следует из (2.27) при $N_s = 1$ и $j = \overline{1, M-1}$. Для установления зависимости вида (2.29) нужны более длительные рассуждения.

Известно, что при булево-алгебраических преобразованиях двоичный вектор $C' = \left\{c'_0, c'_1, ..., c'_{2^l-1}\right\}$ некоторого положительно поляризованного многочлена $f(x)$ над $GF(2)$ может быть получен из конституент СДНФ $A' = \left\{a'_0, a'_1, ..., a'_{2^l-1}\right\}$ с использованием векторно-матричного соотношения:

$$C' = S^l A', \tag{2.30}$$

где $S^l - (0,1)-$ матрица вида

$$S^l = \begin{vmatrix} S^{l-1} & 0 \\ S^{l-1} & S^{l-1} \end{vmatrix}, \qquad S^0 = 1.$$

Согласно правилу построения, для каждого элемента данной матрицы будет выполняться равенство

$$\varepsilon_{\eta,\omega} \oplus \varepsilon_{\eta,j} = \varepsilon_{q,j}, \tag{2.31}$$

где значения q, j определяются соотношениями $q = \eta + 2^{l-1},\ j = \omega + 2^{l-1},\ \forall \eta, \omega$.

Для установления общей связи между элементами $\varepsilon_{q,j}$ в матрице S^l докажем следующую теорему.

***Теорема* 2.2.** (0,1)-матрица S^l есть матрица биномиальных коэффициентов C_q^j над полем $GF(2)$ при $0 \le q, j \le 2^l - 1$.

Доказательство. Введем обозначение для матрицы S^0 в форме биномиального коэффициента $S^0 = C_0^0$. Тогда при $l = 1$ матрица S^l будет иметь следующий вид:

$$S^1 = \begin{vmatrix} C_0^0 & 0 \\ C_0^0 & C_0^0 \end{vmatrix}. \tag{2.32}$$

Из комбинаторики известно, что

$$C_{q-1}^{j-1} + C_{q-1}^j = C_q^j.$$

Это соотношение позволяет выразить элементы $\varepsilon_{q,j}$ в правой части соотношения (2.31) через элементы предыдущей строки матрицы S^l. Действительно, используя рекурсию для равенства (2.31), можно показать, что

$$C_q^j = C_{q-2^{l-1}}^j + \sum_{k=1}^{2^{l-1}-1} C_{2^{l-1}}^k C_{q-2^{l-1}}^{j-k} + C_{q-2^{l-1}}^{j-2^{l-1}}.$$

В силу того что биномиальный коэффициент $C_{2^{l-1}}^k$ четен для $k < 2^{l-1}$, при-

веденное соотношение будет соответствовать равенству

$$C_q^j \equiv C^{j}_{q-2^{l-1}} + C^{j-2^{l-1}}_{q-2^{l-1}} \qquad (\mathrm{mod}\ 2). \tag{2.33}$$

Обозначив в полученном соотношении величины $\eta = q - 2^{l-1}$ и $\omega = j - 2^{l-1}$, из (2.33) имеем

$$C_q^j \equiv C_\eta^j + C_\eta^\omega \qquad (\mathrm{mod}\ 2). \tag{2.34}$$

Отсюда следует, что $\varepsilon_{q,j} \equiv C_q^j \ (\mathrm{mod}\ 2)$ и, таким образом, матрица S^l есть матрица, образованная из остатков по модулю два биномиальных коэффициентов C_q^j. Теорема 2.2 доказана.

Однако методика формирования матриц S^l из подматриц S^{l-1} дает возможность заключить, что S^l есть только определенная комбинация коэффициентов C_q^j. Соответствие же вновь образованного массива упорядоченному набору факториальных моментов должно быть установлено особо.

***Теорема* 2.3.** Матрица S^l, формируемая из подматриц S^{l-1}, является упорядоченной матрицей биномиальных коэффициентов C_q^j над $GF(2)$ для всех $0 \le q, j \le 2^l - 1$.

Доказательство. Пусть $l = 1$. Тогда для матрицы S^1 будет справедливо равенство

$$S^1 = \begin{vmatrix} C_0^0 & 0 \\ C_0^0 & C_0^0 \end{vmatrix} = \begin{vmatrix} C_0^0 & 0 \\ C_1^0 & C_1^1 \end{vmatrix}.$$

Используя вариант записи S^1 с коэффициентами C_0^0 (2.32) и правило формирования матриц S^l из S^{l-1}, сформируем определители более высокого порядка. Для элементов этих определителей будут иметь место сравнения по модулю два (2.31) и (2.34). Следовательно, для доказательства теоремы необходимо показать механизм действия закона (2.34) на границах матриц S^{l-1}, составляющих матрицу S^l.

Очевидно, что из основного рекуррентного соотношения и равенства (2.33) следует

$$C^{j}_{2^{l-1}+k-1} + C^{j+1}_{2^{l-1}+k-1} \equiv C^{j+1}_{2^{l-1}+k} \qquad (\mathrm{mod}\ 2), \tag{2.35}$$

где $0 \le k \le 2^{l-1}$, $0 \le j \le 2^{l-1} - 1 + k$. В то же время сохраняется и привязка к верхней половине S^l, например, вида

$$C^{j+1}_{2^{l-1}+k} \equiv C_k^{j+1} + C_k^{j+1-2^{l-1}} \equiv C_k^{j+1} \qquad (\mathrm{mod}\ 2).$$

В результате, при $k = q$ из (2.35) вытекает соответствие всех элементов матрицы S^l биномиальным коэффициентам $C_q^j (\mathrm{mod}\ 2)$ при $0 \le j, q \le 2^l - 1$.

Таким образом, нижняя половина матрицы S^l будет также представлять собой упорядоченный набор биномиальных коэффициентов, но теперь уже относительно верхней половины над полем $GF(2)$. Теорема 2.3 доказана.

Итак, для булевых форм представления функций в (2.30) имеем

$$c'_0 \equiv a'_0, \qquad c'_q \equiv \sum_{z=1}^{q} C_q^z \cdot a'_z \pmod 2, \qquad (2.36)$$
$$\overline{q = 1, 2^l - 1}, \quad c'_q \in \{0,1\}.$$

Обратное преобразование выполняется по формулам обращения:

$$a'_z \equiv \sum_{q=1}^{z} C_z^q \cdot c'_q \pmod 2$$
$$\overline{z = 1, 2^l - 1}, \quad a'_z \in \{0,1\}.$$

В дискретной математике аналогичные соотношения порождает общий случай формальной зависимости для временной производной от значений булевых аргументов. Простейший вариант данного преобразования соответствует сумме некоторых переменных $x(t)$ над $GF(2)$:

$$\frac{\partial x(t)}{\partial t} \equiv x(t) \oplus x(t \mp 1).$$

С учетом же эквивалентности приведенного соотношения и формулы (2.34), а также учитывая эквивалентность закона формирования производных высших степеней и основного рекуррентного соотношения, можно показать, что

$$\frac{\partial^{(q)} x(t)}{\partial t^q} \equiv \sum_{j=0}^{q} C_q^j \cdot x(t+j) \pmod 2.$$

Данное равенство позволяет соотношения вида (2.35) рассматривать не как объекты булевой алгебры, а как объекты дискретной математики. Тогда с учетом знаков преобразования в теореме 2.1 можно представить равенствами

$$c_q \equiv \frac{\partial^{(q)} A}{\partial t^q} \equiv \sum_{z=1}^{q} (-1)^{s+1} C_q^z \cdot a_z \pmod 2, \quad \overline{q = 1, 2^l - 1}, \qquad (2.37)$$
$$z = \sum_{\lambda=0}^{l-1} \sigma_{u,\lambda} 2^\lambda, \quad s = \sum_{\lambda=0}^{l-1} \sigma_{u,\lambda}, \; a_z \in \{0,1\},$$
$$a_z \equiv \frac{\partial^{(z)} C}{\partial t^z} \equiv \sum_{q=1}^{z} (-1)^{s+1} C_z^q \cdot c_q \pmod 2, \quad \overline{z = 1, 2^l - 1},$$
$$q = \sum_{\lambda=0}^{l-1} \sigma_{u,\lambda} 2^\lambda, \quad s = \sum_{\lambda=0}^{l-1} \sigma_{u,\lambda}, \; c_q \in \{0, \pm 1\}.$$

Здесь векторы $C = \left\{c_1, c_2, ..., c_{2^l-1}\right\}$ и $A = \left\{a_1, a_2, ..., a_{2^l-1}\right\}$ принадлежат равенству

(2.27), а знакопеременность c_q строго устанавливается набором коэффициентов a_z.

Таким образом, тождественность принципа включения и исключения и алгоритма суммирования подмножеств векторов с двоичными коэффициентами a_z можно считать доказанной. Как следствие формул (2.37), также доказана возможность перехода от формы представления алгоритма (2.25) с $c_q \in \{0, \pm 1\}$ к сумме векторов вида (2.24).

Пример. Для подмножеств, приведенных в табл. 2.6, в соответствии с формулой (2.37) получить набор коэффициентов c_q для равенства (2.27).

Очевидно, что на основании (2.36) имеем $c_0 \equiv a_0 \equiv 1 \pmod 2$. Далее, для $q = 1$, имеем

$$c_1 \equiv (-1)^2 C_1^1 \cdot a_1 \equiv |a_1 = 0| \equiv 0 \pmod 2.$$

Остальные коэффициенты полиномиальной формы определяются следующими равенствами:

$$c_2 \equiv (-1)^2 C_2^1 \cdot a_1 + (-1)^2 C_2^2 \cdot a_2 \equiv |a_1 = a_2 = 0| \equiv 0 \pmod 2;$$

$$c_3 \equiv (-1)^2 C_3^1 \cdot a_1 + (-1)^2 C_3^2 \cdot a_2 + (-1)^3 C_3^3 \cdot a_3 \equiv 0 \pmod 2;$$

$$c_4 \equiv (-1)^2 C_4^1 \cdot a_1 + (-1)^2 C_4^2 \cdot a_2 + (-1)^3 C_4^3 \cdot a_3 + (-1)^2 C_4^4 \cdot a_4 \equiv 0 \pmod 2;$$

$$c_5 \equiv (-1)^3 C_5^5 \cdot a_5 \equiv -1 \pmod 2;$$

$$c_6 \equiv (-1)^3 C_6^5 \cdot a_5 \equiv -6a_5 \equiv 0 \pmod 2;$$

$$c_7 \equiv (-1)^3 C_7^5 \cdot a_5 \equiv -21a_5 \equiv -1 \pmod 2;$$

$$c_8 \equiv (-1)^3 C_8^5 \cdot a_5 + (-1)^2 C_8^8 \cdot a_8 \equiv -56a_5 + a_8 \equiv 1 \pmod 2;$$

$$c_9 \equiv (-1)^3 C_9^5 \cdot a_5 + (-1)^2 C_9^8 \cdot a_8 \equiv -126a_5 + 9a_8 \equiv 1 \pmod 2;$$

$$c_{10} \equiv (-1)^3 C_{10}^5 \cdot a_5 + (-1)^2 C_{10}^8 \cdot a_8 + (-1)^3 C_{10}^{10} \cdot a_{10} \equiv -252a_5 + 45a_8 - a_{10} \equiv 0 \pmod 2;$$

$$c_{11} \equiv -462a_5 + 165a_8 - 11a_{10} \equiv 0 \pmod 2;$$

$$c_{12} \equiv -792a_5 + 495a_8 - 66a_{10} - a_{12} \equiv 0 \pmod 2;$$

$$c_{13} \equiv -1287a_5 + 1287a_8 - 286a_{10} - 13a_{12} + a_{13} \equiv 0 \pmod 2;$$

$$c_{14} \equiv -2002a_5 + 3003a_8 - 1001a_{10} - 91a_{12} + 14a_{13} + a_{14} \equiv 0 \pmod 2;$$

$$c_{15} \equiv -3003a_5 + 6435a_8 - 3003a_{10} - 455a_{12} + 105a_{13} + 15a_{14} - a_{15} \equiv -1 \pmod 2.$$

Таким образом, $C = \{1, 0, 0, 0, 0, -1, 0, -1, 1, 1, 0, 0, 0, 0, 0, -1\}$. Подстановка данных значений в правую часть равенства (2.27) дает результат, совпадающий с суммой в табл. 2.6:

$$N_l^0 = 1\cdot N_0 + (-1)\cdot(-1)N_2\left(x_{u,2}, x_{u,0}\right) + (-1)\cdot N_3\left(x_{u,2}, x_{u,1}, x_{u,0}\right) + 1\cdot N_1\left(x_{u,3}\right) +$$
$$+(-1)\cdot N_2\left(x_{u,3}, x_{u,0}\right) + (-1)\cdot(-1)N_4\left(x_{u,3}, x_{u,2}, x_{u,1}, x_{u,0}\right) = 10+(7+7+4)-4+$$
$$+(13+4+10+7+13+4)-(7+4)+4 = 10+18-4+51-11+4 = 68.$$

Выполним проверку правильности преобразования с помощью формул (2.37) путем расчета коэффициентов a_z на основании теперь уже известного вектора:

$$C = \{c_0 = 1,\ c_5 = -1,\ c_7 = -1,\ c_8 = c_9 = 1,\ c_{15} = -1\}.$$

Очевидно, что в соответствии с обратным равенством получим значения:

$a_0 \equiv c_0 \equiv 1 \pmod 2$;

$a_1 \equiv (-1)^2 C_1^1 \cdot c_1 \equiv 0 \pmod 2$;

$a_2 \equiv (-1)^2 C_2^1 \cdot c_1 + (-1)^2 C_2^2 \cdot c_2 \equiv |c_1 = c_2 = 0| \equiv 0 \pmod 2$;

$a_3 \equiv (-1)^2 C_3^1 \cdot c_1 + (-1)^2 C_3^2 \cdot c_2 + (-1)^3 C_3^3 \cdot c_3 \equiv 0 \pmod 2$;

$a_4 \equiv (-1)^2 C_4^1 \cdot c_1 + (-1)^2 C_4^2 \cdot c_2 + (-1)^3 C_4^3 \cdot c_3 + (-1)^2 C_4^4 \cdot c_4 \equiv 0 \pmod 2$;

$a_5 \equiv (-1)^3 C_5^5 \cdot c_5 \equiv (-1)^4 \equiv 1 \pmod 2$;

$a_6 \equiv (-1)^3 C_6^5 \cdot c_5 \equiv -6c_5 \equiv 0 \pmod 2$;

$a_7 \equiv (-1)^3 C_7^5 \cdot c_5 + (-1)^4 C_7^7 c_7 \equiv -21c_5 + c_7 \equiv 0 \pmod 2$;

$a_8 \equiv (-1)^3 C_8^5 \cdot c_5 + (-1)^4 C_8^7 \cdot c_7 + (-1)^2 C_8^8 \cdot c_8 \equiv -56c_5 + 8c_7 + c_8 \equiv 1 \pmod 2$;

$a_9 \equiv (-1)^3 C_9^5 \cdot c_5 + (-1)^4 C_9^7 \cdot c_7 + (-1)^2 C_9^8 \cdot c_8 + (-1)^3 C_9^9 \cdot c_9 \equiv 0 \pmod 2$;

$a_{10} \equiv (-1)^3 C_{10}^5 \cdot c_5 + (-1)^4 C_{10}^7 \cdot c_7 + (-1)^2 C_{10}^8 \cdot c_8 + (-1)^3 C_{10}^9 \cdot c_9 \equiv 1 \pmod 2$;

$a_{11} \equiv -462c_5 + 330c_7 + 165c_8 - 55c_9 \equiv 0 \pmod 2$;

$a_{12} \equiv -792c_5 + 792c_7 + 495c_8 - 220c_9 \equiv 1 \pmod 2$;

$a_{13} \equiv -1287c_5 + 1716c_7 + 1287c_8 - 715c_9 \equiv 1 \pmod 2$;

$a_{14} \equiv -2002c_5 + 3432c_7 + 3003c_8 - 2002c_9 \equiv 1 \pmod 2$;

$a_{15} \equiv -3003c_5 + 6435c_7 + 6435c_8 - 5005c_9 - c_{15} \equiv$

$\equiv |3003 - 5005 = -2002 \equiv 0 \bmod 2| - (c_{15} = -1) \equiv 1 \pmod 2$.

Итак, полученный вектор $A = \{1, 0, 0, 0, 0, 1, 0, 0, 1, 0, 1, 0, 1, 1, 1, 1\}$ полностью совпадает с заданными значениями из табл. 2.6.

В процессе рассмотрения данного примера следует учитывать, что образование множеств $N_s\left(x_{u,l\text{-}1}^{\sigma_{u,l\text{-}1}}, x_{u,l\text{-}2}^{\sigma_{u,l\text{-}2}}, \ldots, x_{u,0}^{\sigma_{u,0}}\right)$ в правой части (2.27) из элементов табл. 2.6 соответствует нахождению покрытия векторов вида (2.24) векторами сокращенной ДНФ булевой функции.

Таким образом, если в качестве открытого текста из табл. 2.6 выбраны значения $\{10, 0, 0, 0, 0, 7, 0, 0, 13, 0, 4, 0, 10, 7, 13, 4\}$, то в результате шифрования получается набор чисел $\{10, 0, 0, 0, 0, -18, 0, -4, 51, 11, 0, 0, 0, 0, 0, -4\}$.

Недостаток метода состоит в обязательном совпадении первого символа сформированного открытого текста и соответствующего элемента в зашифрованном сообщении. Кроме того, обратное преобразование шифрованного текста является неоднозначным, что определяет преобразование подмножеств как сюръективное. Эффективность метода проявляется в двоичной системе, где открытый текст соответствует элементам таблицы истинности, а шифрованное сообщение отождествляется с коэффициентами положительно поляризованного полинома по модулю два:

$$\varphi(x) = c_0 \oplus c_1 x_1 \oplus c_2 x_2 \oplus c_3 x_1 x_2 \oplus c_4 x_3 \oplus c_5 x_1 x_3 \oplus \ldots \oplus c_{2^l - 1} x_1 x_2 \ldots x_l.$$

В этом случае преобразование является инъективным.

2.13. Инъективное преобразование множества $\{X\}$ в элементы меньшего множества $\{Y\}$

Алфавит языка открытого сообщения представляет собой совокупность символов, с помощью которых записываются все текстовые модули, подлежащие шифрованию. Наиболее часто для кодирования символов компьютерных текстов используются 8-разрядные двоичные коды упакованного формата. Длина сообщения в этом случае определяется формулой: $n' = 8n$, где n – число символов текста.

Однако исследования в области языковых множеств показали, что относительные частоты появления различных букв в различных языках при $n \to \infty$ не подчиняются равномерному закону распределения. Более того, частоты появления таких букв, как «о» и «щ», существенно разнятся и составляют величины 0,09 и 0,003 соответственно (то есть в 30 раз). Данный факт указывает на возможность применения системы кодирования чисел, основывающейся на неравномерной длине кодов для символов, имеющих разную частоту появления.

Сущность неравномерного кодирования заключается в следующем. На первом этапе выбирается группа символов-разделителей букв, указывающих на начало каждой буквы. Например, пара цифр 01 может служить элементами разделения символов сообщения. На втором этапе формируются кодовые наборы различной длины, не содержащие кодов выбранных разделяющих пар. С этой целью удобно использовать счетчик Джексона, формирующий первую часть всех состояний схемы (например, табл. 2.7).

На третьем этапе составляется сводная таблица частот символов языка и каждому символу назначается свой код, длина которого увеличивается с уменьшением частоты регистрации. В нашем примере для разделителей букв вида 01 набор кодов для русского алфавита может быть представлен табл. 2.8.

Таблица 2.7

l	0000 1110 1000 1111 1100	*l*	00000 11110 10000 11111 11000 11100
4		*5*	

Таблица 2.8

о 1	е 0	а 10	и 11	н 00	т 100	с 110
р 111	в 000	л 0000	к 1000	м 1100	д 1110	п 1111
у 11111	я 11110	ы 11100	з 11000	ь 10000	ъ 00000	г 000000
ч 100000	б 110000	х 111000	ж 111100	ш 111110	ю 111111	э 1000000
ф 1111111	щ 1111110	ц 1111100				

В реальных системах большие информационные массивы могут иметь свои частоты регистрации символов. Следовательно, целесообразным представляется измерение статистических характеристик в каждом тексте открытого сообщения, после чего буквы кодируются в соответствии с табл. 2.8.

Пример. «Основа государства это самоорганизация и развитие» – текст, состоящий из 44 букв, который при кодировании байтами имеет длину 352 бита. Использование кодов из табл. 2.8 неравномерной длины приводит к образованию последовательности, состоящей, с учетом разделителей, из 216 бит.

011 01110 0100 011 01000 0110 01000000 011 01110 0111111 011110 0110 (57)

01111 01110 01100 01000 0110 011000000 01100 011 01110 0110 011100 (56)

011 011 01111 01000000 0110 0100 0111 0111000 0110 011111100 0111 01 (57)

11110 0111 01111 0110 0111000 01000 0111 01100 0111 010 (46)

Это на 38,63 % меньше, чем начально-исходный массив. В целом, процесс кодирования на этом шаге может быть закончен. Однако, зная, что число букв русского языка составляет 32 символа, полученный текст может быть разбит на «пятерки» бит (табл. 2.9), а алгоритм кодирования в соответствии с таблицей неравномерных кодов выполнен еще раз (неполная «пятерка» доопределяется нулевыми символами). Заметим, однако, что практические испытания метода показали, что повторное применение алгоритма зачастую не дает положительного результата и даже приводит к увеличению длины сообщения. Данный факт следует из необходимости хранения при каждом преобразовании таблиц кодирования данных и другой служебной информации.

01101 11001 00011 01000 01100 10000 00011 01110 01111 11011 11001 10011
13 25 3 8 12 16 3 14 15 27 25 19
11011 10011 00010 00011 00110 00000 01100 01101 11001 10011 10001 10110
27 19 2 3 6 0 12 13 25 19 17 22
11110 10000 00011 00100 01110 11100 00110 01111 11000 11101 11110 01110
30 16 3 4 14 28 6 15 24 29 30 14
11110 11001 11000 01000 01110 11000 11101 0(0000)
30 25 24 8 14 24 29 0

Таблица 2.9

а 1	б 2	в 3	г 4	д 5	е 6	ж 7
з 8	и 9	к 10	л 11	м 12	н 13	о 14
п 15	р 16	с 17	т 18	у 19	ф 20	х 21
ц 22	ч 23	ш 24	щ 25	ы 26	ь 27	ъ 28
э 29	ю 30	я 31		* 0		

Символ 00000 отождествляется, например, как * и перекодированию не подлежит.

Полученный шифротекст имеет следующий вид и размер:

«нщвзмрвопьщу ьубве*мнщусц юрвгоъепшэюо ющшзошэ*» 44×5=220.

Повторное применение алгоритма шифрования в данном случае оказывается неэффективным, поэтому полученный набор символов считаем окончательным вариантом шифрованного сообщения.

Аналитический подход к предложенному методу кодирования, а именно определение средней длины кода по формуле

$$\overline{l} = \sum_{i=1}^{32} \hat{p}_i l_i = 4,233 \quad \text{бита},$$

где $\hat{p}_i$ – частоты символов языка, l_i – длины кодов из табл. 2.8, указывает на среднестатистическую эффективность (53 %) предложенного алгоритма.

Замечание. Символы, повторяющиеся подряд много раз, например пробелы, шифруются однократно своим кодом с простановкой рядом числа-определителя повторов кода.

3. ПРИКЛАДНЫЕ АЛГОРИТМЫ ПОИСКА ДЕТЕРМИНИЗМА И ШИФРОВАНИЯ В КАНАЛАХ СВЯЗИ

3.1. Включение аргументов времени в АКФ (ВКФ) с помощью набора (0,1)-коэффициентов

Практика применения статистических методологий в системах наблюдения случайных процессов в существенной мере ограничивается отсутствием законченных теоретических исследований, определяющих целесообразность применения оценок АКФ, сформированных тем или иным способом. В частности, до настоящего времени нет строгого математического обоснования для оптимальной размерности данных выборочных функций, а также не исследованы свойства и границы вероятностных аргументов, в рамках которых АКФ имеет преимущества перед другими методами формирования свертки.

Для наблюдения детерминизма в последовательностях со случайной природой воспользуемся классической формулой, определяющей взаимное влияние двух процессов $\boldsymbol{X}$ и $\boldsymbol{Y}$ в соответствии с соотношением

$$\hat{K}_{x,y}(\tau) = \overline{\operatorname{cov}}(X,Y) = \frac{1}{n}\sum_{t=0}^{n-1}\left[x(t)-\overline{x}\right]\left[y(t+\tau)-\overline{y}\right] ,$$

где $x(t)$, $y(t+\tau)$ – элементарные события, наблюдаемые в моменты времени t и $t+\tau$, $\overline{x} \to M_x$ и $\overline{y} \to M_y$ при $n \to \infty$.

Если в качестве отсчетов процесса $\boldsymbol{X}$ выбирать наборы элементарных событий длиной i:

$$x(t+0), x(t+1), ..., x(t+i-1), \qquad t = 0,1,2,3,...,$$

а в качестве отсчётов процесса $\boldsymbol{Y}$ наборы событий длиной j, также принадлежащие процессу $\boldsymbol{X}$:

$$x(t+\tau+0), x(t+\tau+1), ..., x(t+\tau+j-1),$$

то можно записать соотношение, характеризующее автокорреляционную связь двух подмножеств аргументов мощностью i и j вида

$$K_x(\tau) = M\left(\left[\sum_{\omega=0}^{i-1}\sigma_\omega x(t+\omega) - M_x\sum_{\omega=0}^{i-1}\sigma_\omega\right]\left[\sum_{\eta=0}^{j-1}\sigma_\eta x(t+\tau+\eta) - M_x\sum_{\eta=0}^{j-1}\sigma_\eta\right]\right), \quad (3.1)$$

где $\sigma_\omega, \sigma_\eta \in \{0,1\}$ – постоянные коэффициенты, определяющие закон включения или исключения аргументов в АКФ (предполагается, что аргументы не зависят друг от друга) при $n \to \infty$ или в выборочную АКФ, если имеет место реальный эксперимент.

Перемножим аргументы в функции (3.1) и получим соотношение, приводящее к требуемой форме включения аргументов в функцию

$$K_x(\tau)=M\left[\sum_{\omega=0}^{i-1}\sigma_\omega x(t+\omega)\sum_{\eta=0}^{j-1}\sigma_\eta x(t+\tau+\eta)-M_x\sum_{\omega=0}^{i-1}\sigma_\omega x(t+\omega)\sum_{\eta=0}^{j-1}\sigma_\eta-\right. \quad (3.2)$$
$$\left.-M_x\sum_{\omega=0}^{i-1}\sigma_\omega\sum_{\eta=0}^{j-1}\sigma_\eta x(t+\tau+\eta)+M_x^2\sum_{\omega=0}^{i-1}\sigma_\omega\sum_{\eta=0}^{j-1}\sigma_\eta\right].$$

В соотношении (3.2) выполним замену индексов вида $i=i_0$, $j=i_1$, $\omega,\eta=z$, при этом будем считать, что индекс индекса соответствует номеру слова в функции автокорреляции (3.2). Соответственно, коэффициент $\sigma_{z,k}$ будет принадлежать z-й позиции k-го слова. Тогда для двух слов $(k=0,1)$ имеем равенство

$$K_x(\tau)=M\left[(M_x)^0\sum_{z=0}^{i_0-1}\sigma_{z,0}x^{\lambda_0=1}(t+z)\sum_{z=0}^{i_1-1}\sigma_{z,1}x^{\lambda_1=1}(t+\tau_1+z)-\right.$$
$$-(M_x)^1\sum_{z=0}^{i_0-1}\sigma_{z,0}x^{\lambda_0=1}(t+z)\sum_{z=0}^{i_1-1}\sigma_{z,1}x^{\lambda_1=0}(t+\tau_1+z)- \quad (3.3)$$
$$\left.-(M_x)^1\sum_{z=0}^{i_0-1}\sigma_{z,0}x^{\lambda_0=0}(t+z)\sum_{z=0}^{i_1-1}\sigma_{z,1}x^{\lambda_1=1}(t+\tau_1+z)+(M_x)^2\sum_{z=0}^{i_0-1}\sigma_{z,0}x^{\lambda_0=0}(t+z)\sum_{z=0}^{i_1-1}\sigma_{z,1}x^{\lambda_1=0}(t+\tau_1+z)\right].$$

Из (3.3) следует

$$K_x(\tau)=M\left[\sum_{g=0}^{3}(-M_x)^{2-\sum_{k=0}^{1}\lambda_k}\prod_{k=0}^{1}\left(\sum_{z=0}^{i_k-1}\sigma_{z,k}x^{\lambda_k}(t+\tau_k+z)\right)\right], \quad (3.4)$$

где λ_k – (0,1)-коэффициенты, входящие в состав двоичного разложения числа g; в нашем случае $g=\lambda_1\cdot 2^1+\lambda_0\cdot 2^0$; $\sigma_{z,k}$ – также (0,1)-коэффициенты, применяемые для включения аргументов в z-й момент времени в k-е слово АКФ, значение $\tau_k=0$ при $k=0$.

Очевидно, что обобщение равенств (3.1)–(3.4) на w подмножеств аргументов приведет к зависимости вида

$$K_x(\tau_w)=M\left[\sum_{g=0}^{2^w-1}(-M_x)^{w-\sum_{k=0}^{w-1}\lambda_k}\prod_{k=0}^{w-1}\left(\sum_{z=0}^{i_k-1}\sigma_{z,k}x^{\lambda_k}(t+\tau_k+z)\right)\right], \quad (3.5)$$

где $g=\lambda_{w-1}2^{w-1}+\lambda_{w-2}2^{w-2}+\ldots+\lambda_0 2^0$.

Правомерность соотношения (3.5) может быть достаточно просто доказана по индукции. Так, умножая правую часть (3.5) на сумму вида

$$(-M_x)^0\sum_{z=0}^{i_w-1}\sigma_{z,w}x^{\lambda_w=1}(t+\tau_w+z)+(-M_x)^1\sum_{z=0}^{i_w-1}\sigma_{z,w}x^{\lambda_w=0}(t+\tau_w+z),$$

что соответствует w+1 слову в АКФ, нетрудно получить выражение

$$K_x(\tau_{w+1}) = M\left[\sum_{g=0}^{2^w-1}(-M_x)^{w-\sum_{k=0}^{w-1}\lambda_k}\prod_{k=0}^{w-1}\left(\sum_{z=0}^{i_k-1}\sigma_{z,k}x^{\lambda_k}(t+\tau_k+z)\right)\times\right.$$

$$\times(-M_x)^{\lambda_w=0}\sum_{z=0}^{i_w-1}\sigma_{z,w}x^{\lambda_w=1}(t+\tau_w+z) + \sum_{g=0}^{2^w-1}(-M_x)^{w-\sum_{k=0}^{w-1}\lambda_k}\prod_{k=0}^{w-1}\left(\sum_{z=0}^{i_k-1}\sigma_{z,k}x^{\lambda_k}(t+\tau_k+z)\right)\times$$

$$\left.\times(-M_x)^{\lambda_w=1}\sum_{z=0}^{i_w-1}\sigma_{z,w}x^{\lambda_w=0}(t+\tau_w+z)\right] =$$

$$= M\left[\sum_{g=2^w}^{2^{w+1}-1}(-M_x)^{w-\sum_{k=0}^{w}\lambda_k}\prod_{k=0}^{w}\left(\sum_{z=0}^{i_k-1}\sigma_{z,k}x^{\lambda_k}(t+\tau_k+z)\right)\Bigg|_{\lambda_w=1} +\right.$$

$$\left. + \sum_{g=0}^{2^w-1}(-M_x)^{w-\sum_{k=0}^{w}\lambda_k}\prod_{k=0}^{w}\left(\sum_{z=0}^{i_k-1}\sigma_{z,k}x^{\lambda_k}(t+\tau_k+z)\right)\Bigg|_{\lambda_w=0}\right].$$

или

$$K_x(\tau_{w+1}) = M\left[\sum_{g=0}^{2^{w+1}-1}(-M_x)^{w+1-\sum_{k=0}^{w}\lambda_k}\prod_{k=0}^{w}\left(\sum_{z=0}^{i_k-1}\sigma_{z,k}x^{\lambda_k}(t+\tau_k+z)\right)\right], \qquad (3.6)$$

где $g = \lambda_w \cdot 2^w + \lambda_{w-1}\cdot 2^{w-1} + ... + \lambda_0 \cdot 2^0$.

Так как равенство (3.5) и полученное соотношение эквивалентны, то доказательство по индукции можно считать достоверным.

Рассмотрим вычислительный аспект исследуемой методологии применительно к вопросам хеширования случайной выборки. Очевидно, что, с точки зрения скоростных свойств расчетных алгоритмов (при наблюдении событий посредством компьютера), равенство (3.5) удобно представить с учетом только двух слов, что минимизирует число перестановок аргументов:

$$K_x(\tau) = M\left[\sum_{\omega=0}^{i-1}\sum_{\eta=0}^{j-1}\sigma_\omega x(t+\omega)\sigma_\eta x(t+\tau+\eta)\right]. \qquad (3.7)$$

Данное соотношение позволяет избежать ряда вычислительных процедур, связанных с классическим центрированием событий, что существенно повышает производительность процессора (и других аппаратных средств) при регистрации случайных событий.

Теоретически, равенство (3.7) позволяет определить оптимальные значения i и j, если предположить, что для элементарных событий $x(t) \in \{X\}$ при $n = \infty$ выполняется равномерный закон распределения. При решении данной задачи будем использовать соотношение для математического ожидания удельного веса слова в выборке длиной n вида $M_x = \frac{i}{n}\frac{1}{2}\left(2^l - 1\right)$. С учетом приведенного значения на базе формулы (3.7) можно сформировать равенство

$$K_x(\tau) = \frac{i\,j}{4n^2}(2^l - 1)^2 \left[\sum_{\omega=0}^{i-1} \sum_{\eta=0}^{j-1} \sigma_\omega \sigma_\eta \right]. \qquad (3.8)$$

Очевидно, что из (3.8) легко определяется минимальное значение функции, которое достигается, если только для одной пары значений ω и η выполняется равенство $\sigma_\omega \sigma_\eta = 1$.

Схема, поясняющая принцип формирования ВКФ-кодов, представлена на рис. 3.1.

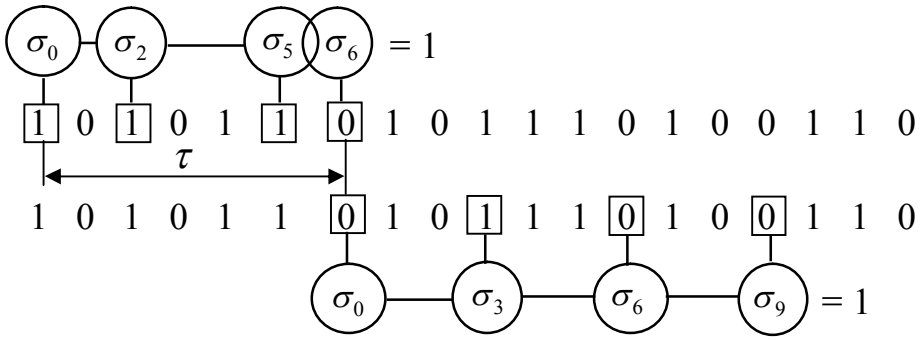

Рис. 3.1

Здесь при $k = 0$ значения $\sigma_0 = \sigma_2 = \sigma_5 = \sigma_6 = 1$, $\sigma_1 = \sigma_3 = \sigma_4 = 0$, при $k = 1$ коэффициенты второго слова равны $\sigma_0 = \sigma_3 = \sigma_6 = \sigma_9 = 1$, $\sigma_1 = \sigma_2 = \sigma_4 = \sigma_5 = \sigma_7 = \sigma_8 = 0$.

Пример. Пусть требуется вычислить значение функции автокорреляции у последовательности вида

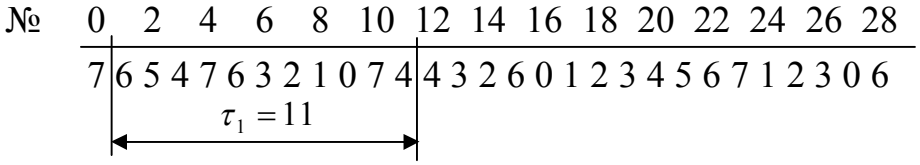

Рис. 3.2

В целях формирования статистики для данной выборки будем использовать алгоритм, следующий из соотношения (3.5) при наборе параметров вида *w=2*, $\tau_0 = 0$, $\tau_1 = 11$, $i = 5$, $j = 6$ с коэффициентами $\sigma_{z,k}$, представленными следующими наборами:

$$\sigma_{z,0} : \underset{1\ \ 1\ \ 0\ \ 1\ \ 1}{\sigma_{0,0}\sigma_{1,0}\sigma_{2,0}\sigma_{3,0}\sigma_{4,0}}, \quad \sigma_{z,1} : \underset{1\ \ 1\ \ 1\ \ 0\ \ 1\ \ 0}{\sigma_{0,1}\sigma_{1,1}\sigma_{2,1}\sigma_{3,1}\sigma_{4,1}\sigma_{5,1}}.$$

Определим численное значение выборочного параметра с учетом длины выборки $n = 29$ по формуле

$$K_x(\tau) = \frac{1}{29} \sum_{t=0}^{28} \left[\left(\sum_{z=0}^{4} \sigma_{z,0} x(t+z) - \bar{x} \sum_{z=0}^{4} \sigma_{z,0} \right) \left(\sum_{z=0}^{5} \sigma_{z,1} x(t+11+z) - \bar{x} \sum_{z=0}^{5} \sigma_{z,1} \right) \right].$$

Для данного примера среднее значение $\bar{x} = 3,69$. Соответственно, величина эмпирической ковариации при этом будет равна

$$K_x(11) = \frac{1}{29}\sum_{t=0}^{28}\left[\left(\sum_{z=0}^{4}\sigma_{z,0}x(t+z) - 14,76\right)\left(\sum_{z=0}^{5}\sigma_{z,1}x(t+11+z) - 14,76\right)\right] = 44.$$

В общем случае рассмотренные механизмы формирования ВКФ (АКФ) имеют целью развитие методов наблюдения последовательностей в рамках статистического анализа наборов чисел. Получаемые в результате применения алгоритмов расшифрования данные проверяются на возможную принадлежность слов в предложениях определенному языку с учетом известных, наиболее часто используемых конструкций.

3.2. Включение элементарных событий в АКФ (ВКФ) с помощью набора (0,1)-коэффициентов

С точки зрения приложений теории корреляции, в частности при наблюдении последовательностей, формируемых на базе регистра сдвига, при передаче данных по w независимым каналам и т. д. существенный интерес вызывает задача исследования автокорреляционной зависимости, возникающей между элементарными событиями $x_\omega(t \pm \tau_\omega), x_\eta(t \pm \tau_\eta), ..., x_\psi(t \pm \tau_\psi)$, циркулирующими в системах связи между абонентами. При этом наличие любого преобразования между элементарными событиями каналов отождествляется с функцией ковариации (корреляции) между переменными вида

$$x_\omega(t) = f\left[x_\eta(t \pm \tau_\eta)\right],\ x_\eta(t \pm \tau_\eta) = f\left[x_\psi(t \pm \tau_\psi)\right], ..., x_\psi(t \pm \tau_\psi) = f\left[x_\omega(t \pm \tau_\omega)\right] ... .$$

Аналогичная методология анализа может быть также использована и при поиске семантически взаимосвязанных объектов, например в информационно-справочных системах. В данном случае необходимо решать задачу соответствия w подмножеств заданных аргументов и словосочетаний в хранимой информации. При этом в результате поиска может быть составлен каталог или список первоисточников содержащих объекты, принадлежащие морфологической категории.

В общем случае при программировании формулы, устанавливающей степень зависимости между подмножествами аргументов, рассмотренное выше соотношение (3.5) для функции автокорреляции между w элементарными событиями следует привести к виду

$$K_x(\tau_w) = M\left[\prod_{k=0}^{w-1}\left(\sum_{z=0}^{i_k-1}\sigma_{x_z,k}x_z(t+\tau_k) - \sum_{z=0}^{i_k-1}\sigma_{x_z,k}M_{x_z}\right)\right], \qquad (3.9)$$

здесь i_k – длина k-го слова; $\sigma_{x_z,k} \in \{0,1\}$ – коэффициенты включения аргументов x_z в k-е слово функции, значение $\tau_k = 0$ при $k = 0$, $n \to \infty$.

Очевидно, что равенство (3.9) представляет собой произведение подмножеств событий x_z, реализуемое следующей скобочной формой:

$$K_x(\tau_w) = \lim_{n\to\infty}\frac{1}{n}\sum_{t=0}^{n-1}\left[\sum_{z=0}^{i_0-1}\sigma_{x_z,0}x_z(t+\tau_0) - \sum_{z=0}^{i_0-1}\sigma_{x_z,0}M_{x_z}\right]\times$$

$$\times\left[\sum_{z=0}^{i_1-1}\sigma_{x_z,1}x_z(t+\tau_1) - \sum_{z=0}^{i_1-1}\sigma_{x_z,1}M_{x_z}\right]\times\ldots\times\left[\sum_{z=0}^{i_{w-1}-1}\sigma_{x_z,w-1}x_z(t+\tau_{w-1}) - \sum_{z=0}^{i_{w-1}-1}\sigma_{x_z,w-1}M_{x_z}\right].$$

Раскроем в приведенном произведении скобки и сгруппируем члены в соответствии с законом включения и исключения переменных $x_z^{\lambda_k}(t+\tau_k)$ в функцию ковариации. Иными словами, будем учитывать закон изменения двоичных коэффициентов λ_k в разложении некоторой переменной g при естественном порядке следования номеров индексов $g = \lambda_{w-1}2^{w-1} + \lambda_{w-2}2^{w-2} + \ldots + \lambda_0 2^0$. В результате для автокорреляции w подмножеств аргументов можно записать соотношение

$$K_x(\tau_w) = \lim_{n\to\infty}\frac{1}{n}\sum_{t=0}^{n-1}\left[\sum_{g=0}^{2^w-1}(-M_x)^{w-\sum_{k=0}^{w-1}\lambda_k}\prod_{k=0}^{w-1}\sum_{z=0}^{i_k-1}\sigma_{x_z,k}x_z^{\lambda_k}(t+\tau_k)\right], \qquad (3.10)$$

где M_x принято равным среднему по всем $M_{x_z} = x_z\frac{1}{2^l}$ при $n=\infty$.

Достоверность соотношения (3.10), как и в случае автокорреляции между аргументами времени, следует из рассуждений.

Умножим правую часть полученного равенства на разность

$$\sum_{z=0}^{i_w-1}\sigma_{x_z,w}x_z(t+\tau_w) - M_x\sum_{z=0}^{i_w-1}\sigma_{x_z,w}, \quad t = 0,1,\ldots,n-1,$$

тогда можно записать соотношение

$$K_x(\tau_w) = \lim_{n\to\infty}\frac{1}{n}\sum_{t=0}^{n-1}\sum_{g=0}^{2^w-1}\left[(-M_x)^{w-\sum_{k=0}^{w-1}\lambda_k}\prod_{k=0}^{w-1}\sum_{z=0}^{i_k-1}\sigma_{x_z,k}x_z^{\lambda_k}(t+\tau_k)\right]\times$$

$$\times(-M_x)^{\lambda_w=0}\sum_{z=0}^{i_w-1}\sigma_{x_z,w}x_z^{\lambda_w=1}(t+\tau_w) +$$

$$+\lim_{n\to\infty}\frac{1}{n}\sum_{t=0}^{n-1}\sum_{g=0}^{2^w-1}\left[(-M_x)^{w-\sum_{k=0}^{w-1}\lambda_k}\prod_{k=0}^{w-1}\sum_{z=0}^{i_k-1}\sigma_{x_z,k}x_z^{\lambda_k}(t+\tau_k)\right](-M_x)^{\lambda_w=1}\sum_{z=0}^{i_w-1}\sigma_{x_z,w}x_z^{\lambda_w=0}(t+\tau_w).$$

Отсюда следует

$$K_x(\tau_{w+1}) = \lim_{n\to\infty}\frac{1}{n}\sum_{t=0}^{n-1}\left[\sum_{g=0}^{2^{w+1}-1}(-M_x)^{w+1-\sum_{k=0}^{w}\lambda_k}\prod_{k=0}^{w}\sum_{z=0}^{i_k-1}\sigma_{x_z,k}x_z^{\lambda_k}(t+\tau_k)\right], \qquad (3.11)$$

где $g = \lambda_w 2^w + \lambda_{w-1}2^{w-1} + \ldots + \lambda_0 2^0$.

Правые части в формулах (3.10) и (3.11) эквивалентны. Следовательно, соотношение (3.10) можно считать доказанным.

Очевидно, что равенство (3.10), хотя и является достаточно общим, не учитывает ряда составляющих, включение которых в формулу диктуется сущностью АКФ, зависящей от аргументов времени. Так, например, если функция имеет параметр $w = const$, то использование указанного соотношения позволяет определить степень связности некоторого набора переменных

$$x_\omega(t+\tau_\omega),\ x_\gamma(t+\tau_\gamma),\ x_\psi(t+\tau_\psi)$$

при длине выборки, равной n. Однако (3.11) не учитывает других перестановок аргументов, например таких как

$$x_\psi(t+\tau_\psi)\, x_\omega(t+\tau_\omega)\, x_\gamma(t+\tau_\gamma),$$
$$x_\gamma(t+\tau_\gamma)\, x_\psi(t+\tau_\psi)\, x_\omega(t+\tau_\omega)\ldots.$$

Данный факт, естественно, влечет за собой неверное представление о сложившейся системе связей с точки зрения чисто временных аргументов и, как следствие, неверное значение статистической АКФ (подразд. 3.1). Таким образом, обобщая формулу (3.9) с учетом формы записи, использованной в равенстве (3.10), можно записать соотношение вида

$$K_x(\tau_h) = \lim_{n\to\infty} \frac{1}{n} \sum_{t=0}^{n-1} \sum_{\eta=0}^{h-1} \left[\sum_{g=0}^{2^{w_\eta}-1} (-M_x)^{w_\eta - \sum_{k=0}^{w_\eta-1} \lambda_k} \prod_{k=0}^{w_\eta-1} \sum_{z=0}^{i_k-1} \sigma_{x_z,k} x_z^{\lambda_k}(t+\tau_k) \right]. \qquad (3.12)$$

В данном выражении h – число различных предложений, а w_η – число наборов переменных в предложении с номером η (или ранг произведения).

Используя в качестве отправного пункта алгоритм формирования оценок по формуле (3.12), факт оптимальности вероятности пропуска ошибки АКФ с аргументами времени при $\sigma_\omega \sigma_\lambda = 1$ (см. равенства выше) без учета центрирующих констант M_x и, наконец, определив $h = 1$, что также минимизирует значение функции, из зависимости (3.12) можно получить выражение

$$K_x(\tau_w) = \lim_{n\to\infty} \frac{1}{n} \sum_{t=0}^{n-1} \prod_{k=0}^{w-1} x_z(t+\tau_k), \quad 0 \le z \le i_k - 1,\ \forall k, \qquad (3.13)$$

где только одно $\sigma_{x_z,k} = 1$ в k-м слове для всех $k = \overline{0, w-1}$.

Приведенное соотношение дает минимальное значение автокорреляционной функции, что, при конечной длине выборки *n,* будет определять и минимум числа перестановок элементарных событий, дающих соответствующий класс эквивалентностей. Данный факт приводит к минимальному изоморфизму в случае формирования хеш-функции и, следовательно, к минимуму вероятности пропуска ошибки при формировании экспериментальных сверточных кодов.

Из равенства (3.13) следует

$$K_x(\tau_w) = p\left[x_z(t+\tau_{w-1}) \cup x_z(t+\tau_{w-2}) \cup \ldots \cup x_z(t+\tau_0)\right] \prod_{k=0}^{w-1} x_z\left(t+\tau_k\right) = \\ = p\left[\bigcup_{k=0}^{w-1} x_z(t+\tau_k)\right] \prod_{k=0}^{w-1} x_z(t+\tau_k), \tag{3.14}$$

где события $x_z(t+\tau_k) \neq 0$ могут быть равны между собой при различных значениях τ_k, $p\left[x_z(t+\tau_{w-1}) \cup x_z(t+\tau_{w-2}) \cup \ldots \cup x_z\left(t+\tau_0\right)\right]$ – вероятность совместного наступления w событий.

Рассмотрим алгоритм идентификации случайной выборки с помощью функции автокорреляции применительно к рассмотренному выше примеру.

Пусть длина последовательности, как и в предыдущем случае, равна $n = 29$.

0	2	4	6	8	10	12	14	16	18	20	22	24	26	28
7 6 5 4 7 6 3 2 1 0 7 4	4 3 2 6 0 1 2 3 4 5 6 7 1 2 3 0 6													

$\tau_1 = 11$

Рис. 3.3

Значения $w = 2,\ \tau_0 = 0,\ \tau_1 = 11,$ а наборы коэффициентов $\sigma_{x_z,k}$ численно равны

$\sigma_{x_z,0}$:

$\sigma_{x_0,0}$	$\sigma_{x_1,0}$	$\sigma_{x_2,0}$	$\sigma_{x_3,0}$	$\sigma_{x_4,0}$
1	0	1	1	1
7	-	6	3	2

$\sigma_{x_z,1}$:

$\sigma_{x_0,1}$	$\sigma_{x_1,1}$	$\sigma_{x_2,1}$	$\sigma_{x_3,1}$	$\sigma_{x_4,1}$	$\sigma_{x_5,1}$
1	1	1	0	1	0
4	4	3	-	5	-

В данной постановке задачи предполагается, что коэффициенты $\sigma_{x_z,k} = 0$, если в очередном такте испытаний значения x_z в z-м разряде k-го слова оказывается отличным от значения, определенного условиями эксперимента.

Воспользуемся соотношением (3.9) и определим эмпирическую АКФ с учетом конечной длины выборки и $w = 2$:

$$\hat{K}_x(\tau_w) = \frac{1}{29} \sum_{t=0}^{28} \left[\sum_{z=0}^{4} \sigma_{x_z,0} x_z(t) - \sum_{z=0}^{4} \sigma_{x_z,0} \bar{x}_z\right] \left[\sum_{z=0}^{5} \sigma_{x_z,1} x_z(t+11) - \sum_{z=0}^{5} \sigma_{x_z,1} \bar{x}_z\right]. \tag{3.15}$$

Для рассматриваемого примера средние значения $\bar{x}_z$ могут быть представлены набором вида

x_z	0	1	2	3	4	5	6	7
$\sum x_z$	0	3	8	12	16	10	30	28

$\bar{x} = 13{,}375$.

Очевидно, что исходные данные в приведенном примере свидетельствуют о том, что количество нулевых сомножителей в формуле (3.15) будет достаточно велико. И только при $t = 0,16,21$ получим отличные от нуля слагаемые вида

$$\hat{K}_x(\tau_w) = \frac{1}{29}\left[(7-13,4)(4+4+3-40,2)+(3-13,4)(5-13,4)+(2-13,4)(4-13,4)\right] \approx 13,15.$$

Теоретическое значение для независимых событий $K_x[\tau] \to 0$ при $n \to \infty$.

Таким образом, можно сделать вывод о том, что степень зависимости между указанными в данной задаче подмножествами дискретных аргументов имеет существенное отклонение от асимптотического параметра. Полученное значение также достаточно отвлеченно характеризует произвольную выборку случайных событий. Вероятно, более точной и информативной характеристикой последовательности многоразрядных векторов со стохастической природой явилось бы максимальное значение АКФ или количество семантических конструкций (без учета произведений x_z в (3.14)), удовлетворяющих заданному набору коэффициентов $\sigma_{x_z,k}$. При этом сформированное значение с определенной степенью достоверности позволило бы судить о структуре анализируемого информационного блока, о принадлежности его к определенной категории информационных объектов, а в экспериментах с длиной выборки, близкой к бесконечности, об устойчивой зависимости между элементарными событиями базового источника чисел.

3.3. ГПСЧ в задачах инъективного отображения выборки. Алгебраическое преобразование данных

Среди применяемых алгоритмов инъективного отображения множеств в подмножества меньшего размера принцип сжатия данных на ГПСЧ является наиболее эффективным. В связи с этим требуемое решение будем искать в форме векторно-матричных преобразований, описывающих поведение цифровой системы на уровне двоичных состояний.

На начальном этапе решения поставленной задачи будем использовать известный принцип построения цифровых схем по таблице истинности. Данная методология предполагает, что зависимость значений некоторой функции $f_i[x_\lambda(t)]$ от входных наборов аргументов $x_\lambda(t)$ в некоторой схеме может быть представлена соотношением

$$f_i[x_\lambda(t)] = \bigvee_{j=0}^{2^l-1} a_{j,i} \bigwedge_{\lambda=0}^{l-1} \left[\sigma_\lambda \oplus \bar{x}_\lambda(t)\right], \qquad (3.16)$$

где $a_{j,i} \in \{0,1\}$ – конституенты j-й строки таблицы истинности i-й булевой функции; значения σ_λ соответствуют разложению индекса j:

$$j = \sigma_{l-1}2^{l-1} + \sigma_{l-2}2^{l-2} + \ldots + \sigma_0 2^0.$$

Входные переменные для функции (3.16) формируются в синхронном устройстве счета по закону

$$\begin{cases} x_0(t) \equiv x_0(t-1) \oplus 1, \\ x_\lambda(t) \equiv x_\lambda(t-1) \oplus \bigwedge\limits_{u=0}^{\lambda-1} x_u(t-1), \qquad \lambda = \overline{1, l-1}. \end{cases}$$

Например при $l = 5$ и $j = 11$ коэффициенту $a_{11,i}$ в (3.16) соответствует конъюнкция переменных $x_0 x_1 \overline{x}_2 x_3 \overline{x}_4$.

Очевидно, что, используя метод временных производных для цифровых последовательностей, можно получить описания схем для функций $f_i\left[x_\lambda(t)\right]$ в полиномиальной форме:

$$f_i\left[x_\lambda(t)\right] \equiv \sum_{q=0}^{2^l-1} c_{q,i} \bigwedge_{\lambda=0}^{l\text{-}1} x_\lambda^{\sigma_\lambda}(t) \quad (\mathrm{mod}\, 2), \tag{3.17}$$

$$q = \sum_{\lambda=0}^{l\text{-}1} 2^\lambda \sigma_\lambda,$$

где $c_{q,i} \in \{0,1\}$ – коэффициенты, определяющие вид полиномиальной схемы.

В приведенном выражении значения $c_{q,i}$ могут быть вычислены по формулам (2.37) на основании СДНФ:

$$c_{q,i} \equiv \sum_{j=0}^{q} C_q^j a_{j,i} \qquad (\mathrm{mod}\, 2), \quad q = \overline{0, 2^l - 1}. \tag{3.18}$$

При этом применение формулы обращения к соотношению (3.18), аналогично (2.37), позволяет выполнить обратный переход, то есть из полиномиальной формы (3.17) получить функцию в СДНФ (3.16), используя выражение

$$a_{j,i} \equiv \sum_{q=0}^{j} C_j^q c_{q,i} \qquad (\mathrm{mod}\, 2), \quad j = \overline{0, 2^l - 1}. \tag{3.19}$$

Рассмотрим теперь принцип инъективного отображения с точки зрения аналитических преобразований (3.17) – (3.19). Очевидно, что устройство-генератор, предположительно осуществляющее обратное преобразование сжатых двоичных сигналов, должно выполнять преобразование $x_\lambda(t) = f\left[x_\omega(t-1)\right]$. С учётом равенства (3.16) данное соотношение представимо в виде

$$f_i\left[x_\lambda(t-1)\right] = \bigvee_{j=0}^{2^l-1} a_{j,i} \bigwedge_{\lambda=0}^{l\text{-}1} \left[\sigma_\lambda \oplus \overline{x}_\lambda(t-1)\right], \qquad i = \overline{1, l}. \tag{3.20}$$

Подстановка в (3.20) результатов преобразования (3.19) показывает, что любой последовательности случайных событий длиной $n = 2^l \neq \infty$ можно поставить в соответствие генератор марковской цепи, синтез которого производится по таблице истинности и формулам

$$f_i\left[x_\lambda(t)\right] \equiv \sum_{j=0}^{2^l-1}\left(\sum_{q=0}^{j} C_j^q a_{q,i}\right)\bigwedge_{\lambda=0}^{l-1} x_\lambda^{\sigma_\lambda}(t-1) \quad (\mathrm{mod}\ 2). \tag{3.21}$$

В случае когда априорно известны коэффициенты $c_{q,i}$, соотношение, аналогичное (3.21), может быть записано в виде (3.22):

$$f_i\left[x_\lambda(t)\right] \equiv \bigvee_{j=0}^{2^l-1}\left(\sum_{q=0}^{j} C_j^q c_{q,i}\right)\bigwedge_{\lambda=0}^{l-1}\left[\sigma_\lambda \oplus \overline{x}_\lambda(t-1)\right] \quad (\mathrm{mod}\,2),$$
$$j=\sigma_{l-1}2^{l-1}+\sigma_{l-2}2^{l-2}+\ldots+\sigma_0 2^0. \tag{3.22}$$

В векторно-матричной форме преобразования (3.21) имеют вид

$$\begin{vmatrix} x_0(t) \\ x_1(t) \\ \ldots \\ x_{l-1}(t) \end{vmatrix} \equiv \begin{vmatrix} a_{0,0} & a_{0,1} & \ldots & a_{0,l-1} \\ a_{1,0} & a_{1,1} & \ldots & a_{1,l-1} \\ . & . & . & . \\ a_{2^l-1,0} & a_{2^l-1,1} & \ldots & a_{2^l-1,l-1} \end{vmatrix}^T \begin{vmatrix} C_0^0 & 0 & 0 & \ldots & 0 \\ C_1^0 & C_1^1 & 0 & \ldots & 0 \\ C_2^0 & C_2^1 & C_2^2 & \ldots & 0 \\ . & . & . & . & . \\ C_{2^l-1}^0 & C_{2^l-1}^1 & C_{2^l-1}^2 & \ldots & C_{2^l-1}^{2^l-1} \end{vmatrix}^T \begin{vmatrix} 1 \\ x_0(t-1) \\ x_1(t-1) \\ x_0(t-1)x_1(t-1) \\ \ldots \\ x_0(t-1)\ldots x_{l-1}(t-1) \end{vmatrix}, \tag{3.23}$$

или $X^t = Q^T C^T X^{t-1}$ при $F\left(X^{t-1}\right) = X^t$ над $GF(2)$.

Очевидно, что умножение первых двух матриц в (3.23) приводит к соотношениям вида

$$\begin{vmatrix} x_0(t) \\ x_1(t) \\ \ldots \\ x_{l-1}(t) \end{vmatrix} \equiv \begin{vmatrix} \sum_{q=0}^{0} a_{q,0}C_0^q & \sum_{q=0}^{0} a_{q,1}C_0^q & \ldots & \sum_{q=0}^{0} a_{q,l-1}C_0^q \\ \sum_{q=0}^{1} a_{q,0}C_1^q & \sum_{q=0}^{1} a_{q,1}C_1^q & \ldots & \sum_{q=0}^{1} a_{q,l-1}C_1^q \\ . & . & . & . \\ \sum_{q=0}^{2^l-1} a_{q,0}C_{2^l-1}^q & \sum_{q=0}^{2^l-1} a_{q,1}C_{2^l-1}^q & \ldots & \sum_{q=0}^{2^l-1} a_{q,l-1}C_{2^l-1}^q \end{vmatrix}^T \begin{vmatrix} 1 \\ x_0(t-1) \\ x_1(t-1) \\ x_0(t-1)x_1(t-1) \\ \ldots \\ x_0(t-1)\ldots x_{l-1}(t-1) \end{vmatrix} \quad (\mathrm{mod}\ 2).$$

Далее получаем

$$\begin{vmatrix} x_0(t) \\ x_1(t) \\ \ldots \\ x_{l-1}(t) \end{vmatrix} \equiv \begin{vmatrix} \sum_{j=0}^{2^l-1}\sum_{q=0}^{j} x_0^{\sigma_0}(t-1)\ldots x_{l-1}^{\sigma_{l-1}}(t-1)a_{q,0}C_j^q \\ \sum_{j=0}^{2^l-1}\sum_{q=0}^{j} x_0^{\sigma_0}(t-1)\ldots x_{l-1}^{\sigma_{l-1}}(t-1)a_{q,1}C_j^q \\ \ldots \\ \sum_{j=0}^{2^l-1}\sum_{q=0}^{j} x_0^{\sigma_0}(t-1)\ldots x_{l-1}^{\sigma_{l-1}}(t-1)a_{q,l-1}C_j^q \end{vmatrix} \quad (\mathrm{mod}\ 2). \tag{3.24}$$

$$j=\sigma_{l-1}2^{l-1}+\sigma_{l-2}2^{l-2}+\ldots+\sigma_0 2^0.$$

Остановимся на данном описании генерирующей части системы отображения. Главное достоинство преобразований (3.23)–(3.24) заключается в использовании операции суммирования по модулю два при формировании очередного состояния схемы. Это позволяет предположить возможность линеаризации цепи обратной связи устройства, в случае если выполняется условие $a_{j,i} \in \{0,1\}$ с равной вероятностью 0,5. По сути, невыполнение указанного условия и является основным препятствием при сжатии данных и определяет общую трудоемкость всего механизма интеграции.

В общем случае преобразование (3.23) определяет следующие действия при решении задачи сжатия. Во-первых, учитывая, что регистрируемая системой информация носит в целом статистический характер, в ходе расчетов инъективного отображения проверяется гипотеза о соответствии бернуллиевской выборки длиной $n < 2^l$ некоторой подпоследовательности совокупности объектов, формируемых с помощью линейного полинома $\varphi(x)$. Во-вторых, если сравнение выборки и всех подпоследовательностей не приводит к желаемому результату, то подбор полиномиального (линейного) описания следует продолжить путем замены примитивного многочлена степени l на другой полином с большим значением l. В-третьих, при невозможности решения задачи в обозримое время необходимо исследовать целесообразность применения методик, суть которых состоит в придании свойств случайности сжимаемой двоичной последовательности. Данная методология может быть реализована поэлементным суммированием по модулю два исходной последовательности и некоторой рекуррентной реализации, формируемой полиномом $\varphi'(x)$, отличным от $\varphi(x)$. При этом, очевидно, потребуется доработка системы, выполняющей обратное преобразование, так как хранение в памяти двух многочленов наряду со сжатым кодом и другой вспомогательной информацией для рассматриваемого алгоритма обработки является обязательным.

Дальнейшее рассмотрение алгоритма инъективного преобразования базируется на следующих рассуждениях и зависимостях, описывающих генераторы марковской цепи.

Матрица Q в соотношении (3.23) описывает все состояния проектируемого генератора данных. Следовательно, используя принцип дешифрации состояний, легко произвести замену (0,1)-таблицы истинности размером $l \times 2^l$ на матрицу переходов T, отождествляющую последовательность состояний некоторой цепи Маркова:

$$T = \left| b_{\lambda,\omega} \right|, \quad \lambda, \omega = \overline{0, 2^l - 1}, \quad b_{\lambda,\omega} \in \{0,1\}.$$

Известно, что полиномиальная форма системы $X^t = F\left[X^{t-1} \right]$ может быть получена в результате умножения матриц $S_n^T T\, S_n^T$:

$$L \equiv \begin{vmatrix} S_{n-1} & 0 \\ S_{n-1} & S_{n-1} \end{vmatrix}^T \cdot |T| \cdot \begin{vmatrix} S_{n-1} & 0 \\ S_{n-1} & S_{n-1} \end{vmatrix}^T, \quad S_0 = 1. \tag{3.25}$$

Выполненные дополнительные исследования в рамках данной задачи доказывают, что матрица биномиальных коэффициентов C_n в соотношении (3.23) является естественным обобщением матрицы S_n. В связи с этим, обобщая равенство (3.25), можно перейти к удобному для программирования соотношению следующего вида:

$$L \equiv C_n^T T\, C_n^T \qquad (\text{mod } 2). \tag{3.26}$$

Очевидно, что для элементов матрицы $S_n^T T$ можно записать обобщение в виде $C_n^T T$:

$$C_n^T T \equiv \begin{vmatrix} \sum\limits_{\lambda=0}^{2^l-1} C_\lambda^0 b_{\lambda,0} & \sum\limits_{\lambda=0}^{2^l-1} C_\lambda^0 b_{\lambda,1} & \cdots & \sum\limits_{\lambda=0}^{2^l-1} C_\lambda^0 b_{\lambda,2^l-1} \\ \sum\limits_{\lambda=1}^{2^l-1} C_\lambda^1 b_{\lambda,0} & \sum\limits_{\lambda=1}^{2^l-1} C_\lambda^1 b_{\lambda,1} & \cdots & \sum\limits_{\lambda=1}^{2^l-1} C_\lambda^1 b_{\lambda,2^l-1} \\ \cdot & \cdot & \cdot & \cdot \\ \sum\limits_{\lambda=2^l-1}^{2^l-1} C_\lambda^{2^l-1} b_{\lambda,0} & \sum\limits_{\lambda=2^l-1}^{2^l-1} C_\lambda^{2^l-1} b_{\lambda,1} & \cdots & \sum\limits_{\lambda=2^l-1}^{2^l-1} C_\lambda^{2^l-1} b_{\lambda,2^l-1} \end{vmatrix} \quad (\text{mod } 2)$$

или

$$C_n^T T \equiv \left| \sum_{\lambda=k}^{2^l-1} C_\lambda^k b_{\lambda,\omega} \right| \qquad (\text{mod } 2) \quad .$$

$$0 \le k, \omega \le 2^l - 1.$$

Соответственно, векторно-матричное произведение (3.26) может быть представлено зависимостью

$$C_n^T T C_n^T \equiv \left| \sum_{\omega=0}^{s} C_s^\omega \sum_{\lambda=k}^{2^l-1} C_\lambda^k b_{\lambda,\omega} \right| \qquad (\text{mod } 2), \tag{3.27}$$

$$0 \le k, s \le 2^l - 1,$$

где k – номер строки, s – номер столбца.

Полная запись формулы (3.27) будет содержать матрицу элементов

$$\begin{vmatrix} \sum\limits_{\omega=0}^{0} C_0^\omega \sum\limits_{\lambda=0}^{2^l-1} C_\lambda^0 b_{\lambda,\omega} & \sum\limits_{\omega=0}^{1} C_1^\omega \sum\limits_{\lambda=0}^{2^l-1} C_\lambda^0 b_{\lambda,\omega} & \cdots & \sum\limits_{\omega=0}^{2^l-1} C_{2^l-1}^\omega \sum\limits_{\lambda=0}^{2^l-1} C_\lambda^0 b_{\lambda,\omega} \\ \sum\limits_{\omega=0}^{0} C_0^\omega \sum\limits_{\lambda=1}^{2^l-1} C_\lambda^1 b_{\lambda,\omega} & \sum\limits_{\omega=0}^{1} C_1^\omega \sum\limits_{\lambda=1}^{2^l-1} C_\lambda^1 b_{\lambda,\omega} & \cdots & \sum\limits_{\omega=0}^{2^l-1} C_{2^l-1}^\omega \sum\limits_{\lambda=1}^{2^l-1} C_\lambda^1 b_{\lambda,\omega} \\ \cdot & \cdot & \cdot & \cdot \\ \sum\limits_{\omega=0}^{0} C_0^\omega \sum\limits_{\lambda=2^l-1}^{2^l-1} C_\lambda^{2^l-1} b_{\lambda,\omega} & \sum\limits_{\omega=0}^{1} C_1^\omega \sum\limits_{\lambda=2^l-1}^{2^l-1} C_\lambda^{2^l-1} b_{\lambda,\omega} & \cdots & \sum\limits_{\omega=0}^{2^l-1} C_{2^l-1}^\omega \sum\limits_{\lambda=2^l-1}^{2^l-1} C_\lambda^{2^l-1} b_{\lambda,\omega} \end{vmatrix}.$$

Таким образом, соотношение (3.27) позволяет установить соответствие между матрицей переходов T, формируемой на основании элементов массива Q (или конституент таблицы истинности), и структурой отображающего устройства, реализующего обратное инъективное преобразование.

Полученное преобразование позволяет получить принципиальное решение задачи инъективного отображения элементов множества $\{X\}$ (двоичной последовательности) в элементы меньшего множества $\{Y\}$, однако недостаток данной информационной технологии заключается в том, что, в самом общем случае, сформированная цепь обратной связи устройства будет иметь нелинейный характер. Это приводит к увеличению средней сложности схемы, а также средней тождественности набора исходных данных и полиномиальной формы решений поставленной задачи.

С целью линеаризации обратной связи формируемого устройства (что эквивалентно сокращению затрат аппаратуры на развёртку состояний цепи Маркова) рассмотрим структуру матрицы L и местоположение коэффициентов полиномиального описания преобразователя в пределах данного массива.

Анализ расчетов матрицы $L = C_n^T T C_n^T$ показывает, что попытка линеаризации системы является удачной, если проверка на наличие унитарной единицы элементов столбцов с номерами $2^1,2^2,2^3,\ldots,2^{l-1}$, принадлежащих строкам с номером $2^0,2^1,2^2,\ldots,2^{l-1}$, а также проверка строки с номером 2^{l-1} на наличие единичных значений только в столбцах с номерами $2^0,2^1,\ldots,2^{l-1}$ является удовлетворительной. Если сформулированные условия выполняются, то процесс линеаризации считается выполненным. В случае невыполнения хотя бы одного из условий следует перейти к алгоритму формирования нового шифротекста (или выбору нового полинома) и, достроив выборку событий до требуемой величины 2^l (последовательно подбирая биты необязательных событий), осуществить программный поиск примитивного полинома в соответствии с преобразованием $L = C_n^T T C_n^T$.

В более сложном случае может быть рассмотрен метод выравнивания вероятностей с помощью линейных рекуррентных последовательностей, формируемых другими неприводимыми полиномами над $GF(2)$.

Очевидно, что в общем случае лучшие варианты сжатия могут привести к образованию кода длиной $3\log_2 n$, где коэффициент 3 характеризует сумму:

1) $\log_2 n$ бит описания генератора;

2) $\log_2 n$ бит начальной загрузки устройства;

3) $\approx \log_2 n$ бит, указывающих на длину сжатой информации.

Полученные результаты носят выраженный прикладной характер. Возможное внедрение изложенного принципа отображения связано с использованием интеллектуальных систем, реализующих алгоритмы передачи информации в вычислительных сетях. Использование данного подхода в криптографии позволяет существенно сократить пребывание абонента в канале связи.

3.4. Пример кодирования и расчета сжатого кода последовательности двоичных событий

Рассмотрим пример формирования кодового набора двоичных символов, обладающего способностью к 100-процентному воспроизведению информации. Пусть задана последовательность элементарных событий, образующая текст сообщения вида «α КАССИОПЕИ».

Применение тривиального способа кодирования к данному тексту приводит к образованию двоичной псевдореализации случайной выборки длиной $n = 2^l - 1$ (рис. 3.4).

Однако при данном методе кодирования в синтезированной последовательности не выполняется условие равновероятности для бернуллиевских элементарных событий. И хотя $l_{\min}$ – ширина «окна», дающая при сдвиге неповторяющиеся наборы символов, равна 6 (это же значение определяет и степень генераторного полинома), синтез линейной системы будет существенно затруднен из-за большого числа (33) доопределяемых символов.

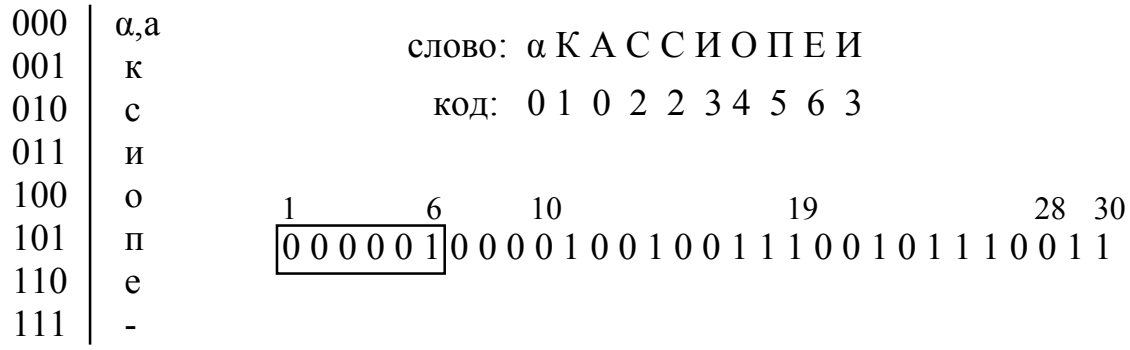

Рис. 3.4

Другой способ кодирования основывается на применении кодов символов, ориентированных на выравнивание вероятностей нуля и единицы. Это даёт возможность получить реализацию последовательности, достаточно близкую к случайной, и с набором двоичных чисел, приведенным на рис. 3.5. Однако недостатком сформированной двоичной последовательности является наличие «окна» разрядностью 7, что, так же как и в предыдущем примере, требует долговременных расчетов.

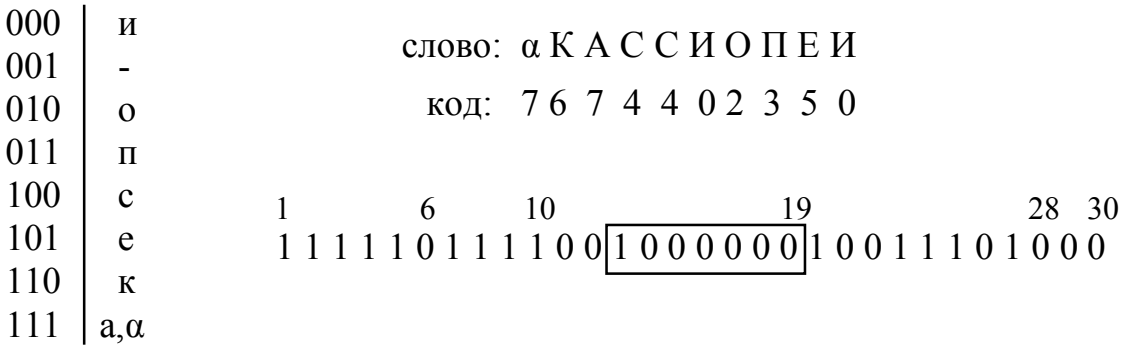

Рис. 3.5

Специальным образом подобранные коды (триады, соответствующие отсчетам 5-разрядного ГПСЧ) дают оптимальный шифротекст для рассматриваемого сообщения в виде набора слов, представленного рис. 3.6. При этом набор символов оказывается практически равновероятен, а размер «окна», очевидно, равен пяти битам.

Учтём далее тот факт, что для исходного не двоичного текста циклическая перестановка букв не изменяет смыслового содержания запоминаемой информации. Это позволяет в результате некоторого числа сдвигов получить шифротекст с теми же высокоточными вероятностными характеристиками, однако весьма удобный для реализации процедуры сжатия (рис. 3.7).

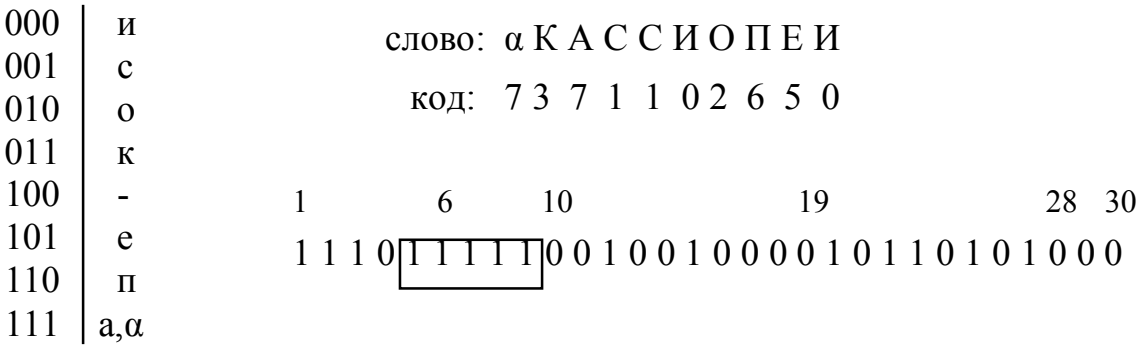

Рис. 3.6

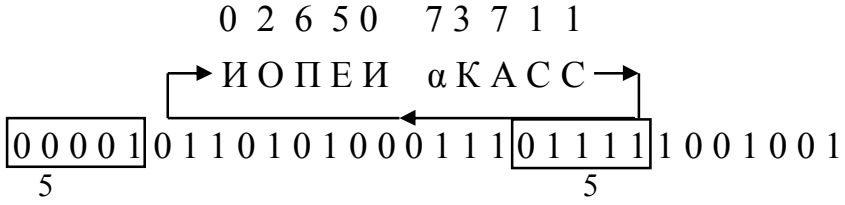

Рис. 3.7

Если полученный набор доопределить до значения $2^5 - 1 = 31$ единичным или нулевым битом справа (в нашем случае удобно использовать единичное значение), то, выполнив 5 сдвигов последовательности вверх и вправо, можно построить следующую T матрицу переходов марковской цепи с 16-ричными символами для обозначения тетрад в строках (верхняя строка нулевая):

$$T_1 = \begin{vmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 4 \\ 0 & 2 \end{vmatrix}_{16}, \quad T_2 = \begin{vmatrix} 4 & 0 \\ 2 & 0 \\ 0 & 8 \\ 0 & 1 \end{vmatrix}_{16}, \quad T_3 = \begin{vmatrix} 8 & 0 \\ 1 & 0 \\ 0 & 4 \\ 0 & 2 \end{vmatrix}_{16},$$

$$T = \begin{vmatrix} T_1 & 0 & 0 & 0 \\ 0 & T_2 & 0 & 0 \\ 0 & 0 & T_3 & 0 \\ 0 & 0 & 0 & T_2 \\ T_2 & 0 & 0 & 0 \\ 0 & T_3 & 0 & 0 \\ 0 & 0 & T_2 & 0 \\ 0 & 0 & 0 & T_3 \end{vmatrix}_{16}.$$

Векторно-матричное произведение $C_{32}^{T}T$ для данной интерпретации запоминаемой информации будет иметь вид (3.28). Очевидно, что нарушение регулярности в подмассиве R_{1_1} обусловлено известным отклонением числа состояний регистра сдвига с ожидаемой линейной обратной связью от значения 2^l. Искусственное введение нулевой вершины в последовательность состояний генератора хотя и даст равенство R_{1_1} и R_1, однако однозначно приведёт к невозможности линеаризации системы и увеличению длины сжатого кода.

$$C_{32}^{T}T = \begin{vmatrix} R_{1_1} & R_1 & R_1 & R_1 \\ 0 & R_1 & 0 & R_1 \\ 0 & 0 & R_1 & R_1 \\ 0 & 0 & 0 & R_1 \\ R_2 & R_3 & R_2 & R_3 \\ 0 & R_3 & 0 & R_3 \\ 0 & 0 & R_2 & R_3 \\ 0 & 0 & 0 & R_3 \end{vmatrix}. \quad R_1 = \begin{vmatrix} F & F \\ 3 & 3 \\ 0 & F \\ 0 & 3 \end{vmatrix}_{16}, \quad R_2 = \begin{vmatrix} 6 & 9 \\ 2 & 1 \\ 0 & 9 \\ 0 & 1 \end{vmatrix}_{16}, \quad R_3 = \begin{vmatrix} 9 & 6 \\ 1 & 2 \\ 0 & 6 \\ 0 & 2 \end{vmatrix}_{16}, \quad R_{1_1} = \begin{vmatrix} 7 & F \\ 3 & 3 \\ 0 & F \\ 0 & 3 \end{vmatrix}_{16}. \tag{3.28}$$

Для нашего примера (см. рис. 3.7) произведение $C_{32}^{T}TC_{32}^{T}$ будет равно

$$C_{32}^{T}TC_{32}^{T} = \begin{array}{c|cccccccc} 0 & 7 & F & F & F & F & F & F & F \\ 1 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 2 & 0 & 8 & 0 & 0 & 0 & 0 & 0 & 0 \\ & . & . & . & . & . & . & . & . \\ 4 & 0 & 0 & 8 & 0 & 0 & 0 & 0 & 0 \\ & . & . & . & . & . & . & . & . \\ 8 & 0 & 0 & 0 & 0 & 8 & 0 & 0 & 0 \\ & . & . & . & . & . & . & . & . \\ 16 & 6 & 8 & 8 & 0 & 0 & 0 & 0 & 0 \\ & . & . & . & . & . & . & . & . \end{array}_{16}.$$

Рассмотрим строки с номерами 1, 2, 4, 8 и 16 данного векторно-матричного произведения, характеризующие связи в устройстве с 5-ю элементами памяти. В частности, строка с номером 1 содержит единственный единичный бит в столбце с номером два. Это означает, что информация из разряда с весом 2^1 сдвигается в разряд с весом 2^0 без каких-либо изменений.

Вторая строка произведения содержит единичное значение только в столбце с номером четыре. Следовательно, в формируемой системе имеет место сдвиг данных из разряда с весом 2^2 в разряд с весом 2^1 и также без какакихлибо преобразований.

Аналогичные явления характерны и для строк с номерами 4 и 8. Здесь единичный бит присутствует в столбцах с номерами 8 и 16 соответственно. Это обусловливает сдвиг в генераторе из разряда с весом 2^3 в разряд с весом 2^2, а также перезапись информации из разряда с весом 2^4 в разряд с весом 2^3.

Обратная связь формируемой системы будет описываться 16 строкой произведения $C_{32}^T T C_{32}^T$. В данной строке единичные биты расположены в 1,2,4 и 8 столбцах. Это означает, что цепь обратной связи схемы, воспроизводящей исходную последовательность, может быть представлена полиномом

$$\varphi(x) = 1 + x^2 + x^3 + x^4 + x^5 .$$

Учитывая, что степень переменной x в многочлене $\varphi(x)$ выбирается из условия l минус степень веса разряда, получаем

$$5-0=5, \quad 5-1=4, \quad 5-2=3, \quad 5-3=2.$$

Данный набор и дает степени переменной x вида x^2, x^3, x^4, x^5.

Схема, позволяющая воспроизвести заданное закодированное сообщение, окончательно может быть представлена в следующем виде (рис. 3.8).

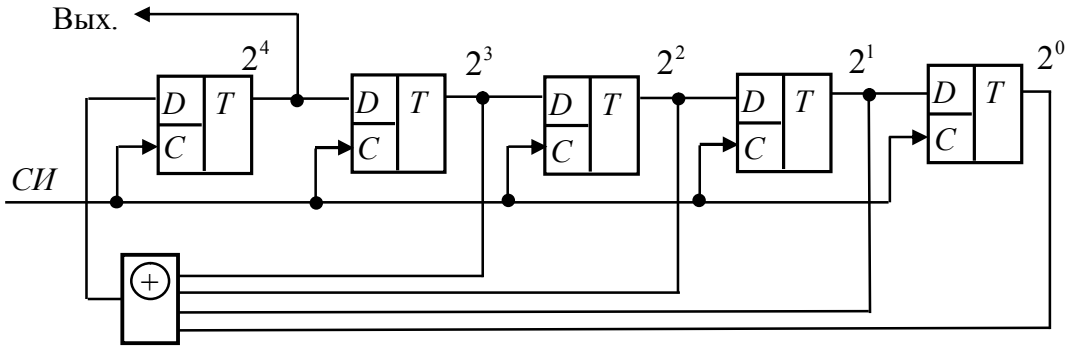

Рис. 3.8

Таким образом, вместо 30-разрядного двоичного сообщения сформированное описание предполагает хранение 15-разрядного слова 01100 01111 и 11110, где первые 5 бит указывают на стартовую загрузку генератора, вторые – на организацию обратной связи, позволяющую генерировать восстанавливаемую информацию, а третьи – на длину формируемой последовательности.

ЛИТЕРАТУРА

1. Романцев, Ю. В. Защита информации в компьютерных системах и сетях / Ю. В. Романцев, П. А. Тимофеев, В. Ф. Шаньгин. – М. : Радио и связь, 1999.
2. Зима, В. Ф. Безопасность глобальных сетевых технологий / В. Ф. Зима, А. А. Молдовян, Н. А. Молдовян. – СПб. : Питер, 2001.
3. Математические и компьютерные основы криптографии / Ю. С. Харин [и др.]. – Минск, 2003.
4. Мафтик, С. Механизмы защиты в сетях ЭВМ / С. Мафтик. – М. : Мир, 1993.
5. Олифер, В. Г. Сетевые информационные системы / В. Г. Олифер, Н. А. Олифер. – СПб. : Питер, 2006.
6. Правовые и организационно-технические методы защиты информации : учеб. пособие / В. Ф. Голиков [и др.]. – Минск, БГУИР, 2004.
7. Хореев, П. Б. Методы и средства защиты информации в компьютерных системах / П. Б. Хореев. – М. : АСАДЕМА, 2005.
8. Петраков, А. В. Основы практической защиты информации : учеб. пособие / А. В. Петраков. – М. : СОЛОН_ПРЕСС, 2005.
9. Герасименко, В. А. Основы защиты информации / В. А. Герасименко, А. А. Малюк. – М. : МГИФИ, 1997.
10. Зегжда, П. Д. Как построить защищенную систему. Технология создания безопасных систем / П. Д. Зегжда, А. М. Ивашко. – СПб. : Питер, 1998.
11. Леонов, А. П. Безопасность автоматизированных банковских и офисных технологий / А. П. Леонов, К. П. Леонов, Г. В. Фролов. – Минск : Национальная книжная палата Беларуси, 1996.
12. Лукацкий, А. В. Обнаружение атак / А. В. Лукацкий. – СПб. : Питер, 1998.
13. Оценка безопасности информационных технологий. – М. : СИП РИА, 2001.
14. Голиков, В. Ф. Криптографическое кодирование информации : метод. указания к лабораторным работам / В. Ф. Голиков, А. В. Курилович. – БГУИР, 1996.
15. Кобяк, И. П. Включение и исключение аргументов в автокорреляционной функции с помощью набора (0,1)-коэффициентов / И. П. Кобяк // АВТ. – 2001. – № 3. – С. 64–74.
16. Кобяк, И. П. О методе включения и исключения как базовом алгоритме синтеза контролепригодных микроэлектронных схем / И. П. Кобяк // Микроэлектроника. – 1997. – Т.26. – №2. – С. 42–48.
17. Кобяк, И. П. Принцип включения и исключения в задачах идентификации случайных последовательностей / И. П. Кобяк // АВТ. – 1997. – №2. – С. 53–64.
18. Федоров, Р. Ф. Стохастические преобразователи информации / Р. Ф. Федоров, В. В. Яковлев, Г. В. Добрис. – Л. : Машиностроение, 1978.

19. Яковлев, В. В. Стохастические вычислительные машины / В. В. Яковлев, Р. Ф. Федоров. – Л. : Машиностроение, 1974.

20. Варакин, Л. Е. Системы связи с шумоподобными сигналами / Л. Е. Варакин. – М. : Радио и связь, 1985.

21. Кобяк, И. П. Системные средства для сжатия\восстановления данных в подсистемах запоминающих устройств интеллектуальных вычислительных систем / И. П. Кобяк // АВТ. – 2001. – №2. – С. 51–61.

22. Основы криптографии : учеб. пособие / А. П. Алферов [и др.]. – М. : Гелиос АРВ, 2001.

23.Теория кодирования / Т. Кассами [и др.] ; пер. с яп. – М. : Мир, 1978.

24. Кларк, Дж. Кодирование с исправлением ошибок в системах цифровой связи / Дж. Кларк, Дж. Кейн; пер. с англ. – М. : Радио и связь, 1987.

25. Горбатов, В. А. Основы дискретной математики : учеб. пособие / В. А. Горбатов. – М. : Высш. шк., 1986.

26. Бендат, Дж. Прикладной анализ случайных данных / Дж. Бендат, А. Пирсол; пер. с англ. – М. : Мир, 1989.

27. Бертсекас, Д. Сети передачи данных / Д. Бертсекас, Р. Галлагер; пер. с англ. – М. : Мир, 1989.

28. Нечаев, В. И. Элементы криптографии. Основы теории защиты информации : учеб. пособие / В. И. Нечаев; под ред. В. А. Садовничего. – М. : Высш. шк., 1999.

29. Блейхут, Р. Теория и практика кодов, контролирующих ошибки / Р. Блейхут; пер. с англ. – М. : Мир, 1986.

Учебное издание

Кобяк Игорь Петрович

ЗАЩИТА ИНФОРМАЦИИ В ВЫЧИСЛИТЕЛЬНЫХ СЕТЯХ

УЧЕБНО-МЕТОДИЧЕСКОЕ ПОСОБИЕ

Редактор *И. П. Острикова*
Корректор *Е. Н. Батурчик*
Компьютерная верстка *А. В. Бас*

Подписано в печать 15.12.2011. Формат 60x84 1/16. Бумага офсетная. Гарнитура «Таймс».
Отпечатано на ризографе. Усл. печ. л. 5,23. Уч.-изд. л. 5,0. Тираж 100 экз. Заказ 37.

Издатель и полиграфическое исполнение: учреждение образования
«Белорусский государственный университет информатики и радиоэлектроники»
ЛИ №02330/0494371 от 16.03.2009. ЛП №02330/0494175 от 03.04.2009.
220013, Минск, П. Бровки, 6