УДК 621.369.9

ЭНТРОПИЙНЫЕ МОДЕЛИ И ЭТАЛОННЫЕ ОПИСАНИЯ ПОМЕХОУСТОЙЧИВЫХ КОДОВ

К. Б. Махров, В. О. Хацаюк

Военно-космическая академия имени А.Ф. Можайского,
197198, Санкт-Петербург, ул. Ждановская, д. 13

Статья поступила в редакцию 16 сентября 2019 г.

Аннотация. В статье рассмотрены вопросы идентификации помехоустойчивых кодов в радиолиниях с многопозиционной модуляцией в условиях неопределенности относительно используемой маркировки канальных символов. Представлена разработанная аналитическая модель, отражающая зависимость вносимой избыточности от параметров помехоустойчивого кода и позволяющая сформировать эталонные описания помехоустойчивых кодов различных классов (сигнатуры). Использование предлагаемых эталонных описаний при решении задачи идентификации помехоустойчивых кодов позволяет значительно сократить признаковое пространство, и как следствие, снизить вычислительную сложность процедур технического анализа и уменьшить требуемый объем выборки по сравнению с известными статистическими методами.

Ключевые слова: когнитивное радио, эталонные описания, помехоустойчивое кодирование, сигнально-кодовые конструкции.

Abstract. In this work, we make a look at the actual problems of blind recognition of error-correcting codes in non-cooperative context, assuming non-binary modulation and no manipulation code knowledge. In particular, we propose a novel analytical model that reflects the dependence of introduced redundancy on the error-correcting code parameters and allows making compact mapping-invariant signatures of different error-correcting codes. Proposed approach to codes identification provides a reduction in feature space, computational complexity and the required sample size compared to known statistical methods. The simulation results confirm the viability of the proposed model for blind recognition of error-correcting codes.

Key words: cognitive radio, blind recognition, error-correcting codes, coded modulation.

Введение

Широкое развитие систем беспроводной связи привело к появлению новых технологий доступа к радиочастотному спектру, например, так называемому, динамическому доступу, лежащему в основе систем когнитивного радио. Функционирование данных систем предполагает возможность динамического реконфигурирования тракта приема сигналов в условиях априорной неопределенности относительно процедур их формирования. Указанная неопределенность устраняется в результате проведения технического анализа сигналов, основной целью которого является обеспечение доступа к передаваемым сообщениям [1].

В настоящее время в радиолиниях активно используются согласованные комбинации многопозиционной модуляции и помехоустойчивого кодирования, известные как сигнально-кодовые конструкции (СКК) [2, 3]. Многообразие и высокая структурная сложность используемых СКК обуславливают высокий уровень априорной неопределенности при решении задач определения их параметров, что приводит к существенному увеличению временных затрат на проведение соответствующих этапов технического анализа [4].

Анализ существующих методов обнаружения и идентификации сигнально-кодовых конструкций показал, что в основе большинства применяемых на практике подходов к идентификации СКК лежит определение параметров используемого помехоустойчивого кода методом пробного декодирования, либо алгебраическими методами на основе решения систем линейных уравнений или свертки с проверочными полиномами [5, 6, 7].

Указанные методы требуют априорного знания правила отображения комплексных сигнальных точек на множество двоичных кодовых комбинаций (маркировки). При отсутствии такой информации, в процессе идентификации выполняется поиск маркировки с использованием переборных алгоритмов. В

силу факториальной зависимости числа возможных маркировок от размера сигнального созвездия, реализация перебора среди всех возможных вариантов не позволяет достичь приемлемой оперативности решения задачи идентификации СКК, поэтому выполняется перебор среди небольшого числа известных сигнально-кодовых конструкций с соответствующим снижением полноты идентификации.

Отмеченные ограничения свидетельствуют о необходимости разработки новых подходов к идентификации СКК с целью повышения оперативности процедур технического анализа без снижения полноты. В частности, в настоящей работе представлена энтропийная модель помехоустойчивого кодирования, позволяющая сформировать эталонные описания помехоустойчивых кодов для идентификации СКК, инвариантные к используемой маркировке.

1. Постановка задачи

Рассмотрим основные преобразования при передаче информации от источника сообщений в типовой системе связи с СКК.

Первым этапом обработки сигнала является кодирование источника, обеспечивающее преобразование аналогового сигнала в цифровой и снижение информационной избыточности за счет применения различных схем сжатия данных. С выхода кодера источника двоичная информационная последовательность поступает на вход канального кодера, но предварительно, как правило, подвергается скремблированию (рандомизации) и перемежению символов [8]. Скремблирование обеспечивает равномерность энергетического спектра излучаемого радиосигнала и повышает надежность восстановления тактовой синхронизации в демодуляторе за счет приближения статистики появления символов в передаваемой информационной последовательности к случайной. Использование перемежителя символов позволяет декоррелировать ошибки в канале, то есть преобразовывать пакетные ошибки в ряд одиночных, что существенно повышает эффективность помехоустойчивого кодирования для каналов связи с памятью. В каскадных схемах помехоустойчивого

кодирования, перемежитель обычно выполняет перемежение символов внешнего кода перед подачей на кодер внутреннего. Таким образом, можно полагать, что дискретная последовательность x_t на входе канального кодера имеет минимальную избыточность, а ее символы распределены равномерно и независимы.

Последующие преобразования дискретной последовательности x_t в канальном кодере включают в себя разделение (демультиплексирование) на k-мерные векторы $\mathrm{x}_{\langle k\rangle}$, соответствующие кодируемым символам, и векторы $\mathrm{x}_{\langle u\rangle}$ размерности u из некодируемых символов, а также процедуры помехоустойчивого и манипуляционного кодирования.

Помехоустойчивое кодирование, описываемое отображением $\{\mathrm{x}_{<k>}\} \to \{\mathrm{y}_{<n>}\}$, заключается во внесении избыточности в информационную последовательность за счет дополнительных проверочных символов, значение которых полностью определяется информационными. В результате, k-мерные входные векторы $\mathrm{x}_{\langle k\rangle}$ преобразуются в n-мерные выходные векторы $\mathrm{y}_{\langle n\rangle}$. Отношение $R = \frac{k}{n}$ называется относительной скоростью кода, а величина $\frac{n-k}{n} = (1-R)$ – относительной избыточностью. Они характеризуют, соответственно, количество информационных и избыточных символов в кодированной последовательности.

Используемые в настоящее время помехоустойчивые коды разделяются на два основных класса – блоковые, в которых значение выходных символов зависит только от текущих информационных, и непрерывные, в которых для вычисления выходных символов используются также предыдущие значения информационных (или выходных) символов, находящиеся в памяти кодера. Наиболее распространенными непрерывными кодами являются сверточные, обладающие свойством линейности. Задание связей между входными (информационными) и выходными символами, полностью определяющее

блоковый помехоустойчивый код, производится с помощью порождающей матрицы кода $\mathbf{G}$ размерности $k \times n$, а кодирование описывается как умножение вектора $\mathrm{x}_{\langle k \rangle}$ на порождающую матрицу $\mathbf{G}$. В случае сверточного кода, порождающая матрица является полубесконечной, но может быть получена периодическим повторением со сдвигом базисной порождающей матрицы $\mathbf{G}$. Таким образом, полная идентификация помехоустойчивого кода предполагает восстановление порождающей матрицы, а частичная включает в себя определение параметров n, k и памяти кодера для непрерывных кодов.

Манипуляционное кодирование представляет собой нелинейное преобразование, выполняемое в два этапа. На первом, дискретная последовательность, состоящая из векторов некодированных символов $\mathrm{x}_{\langle u \rangle}$ и векторов $\mathrm{y}_{\langle n \rangle}$ с выхода помехоустойчивого кодера разделяется на фрагменты длиной a двоичных разрядов $\left\{\left\langle \mathrm{x}_{<u>};\mathrm{y}_{<n>}\right\rangle_c\right\} \to \left\{\left\langle \mathrm{z}_{<a>}\right\rangle_L\right\}$, где $L = \dfrac{c(u+n)}{a}$, $c \in \mathbb{N}$ называется размерностью конструкции. На втором этапе двоичные комбинации отображаются на сигнальное созвездие – множество элементарных сигналов, представленных комплексными огибающими, Данное биективное отображение $\mu : \mathbb{F}_2^a \to \mathcal{C}$, где $\mathbb{F}_2^a$ – множество двоичных комбинаций длины a, $\mathcal{C}$ – сигнальное созвездие из 2^a элементов будем называть маркировкой. Полученная дискретная последовательность, соответствующая элементарным сигналам, преобразуется в непрерывный сигнал $z(t)$ в модуляторе.

В приемное устройство переданный полезный сигнал $z(t)$ поступает в смеси с шумом $e(t)$. После восстановления тактовой синхронизации в демодуляторе, из принятой смеси $\tilde{z}(t) = z(t) + e(t)$ получается дискретная последовательность оценок комплексных огибающих передаваемых элементарных сигналов. В результате их обратного отображения на множество двоичных комбинаций формируется последовательность оценок канальных

символов $\tilde{Z}_{\langle N \rangle} \triangleq \langle \tilde{z}_1, \ldots, \tilde{z}_N \rangle$. Решение задачи идентификации СКК в условиях априорной неопределённости предполагает, что используемая при передаче маркировка μ неизвестна, и двоичное представление элементарных сигналов при приеме выбирается произвольным образом.

Понимая под идентификацией СКК определение структуры и параметров кодера, задача идентификации на вербальном уровне может быть декомпозирована следующим образом:

- обнаружение помехоустойчивого кода в принятой последовательности $\tilde{Z}_{\langle N \rangle}$ и определение его параметров n, k;

- определение используемой маркировки μ и восстановление манипуляционного кода;

- восстановление порождающей матрицы помехоустойчивого кода $\mathbf{G}$.

Результаты решения данной задачи позволяют синтезировать декодер СКК и обеспечить доступ к передаваемой информационной последовательности.

Для реализации процедур обнаружения и определения параметров помехоустойчивого кода в СКК необходимо разработать инвариантную к выбранной маркировке обобщенную модель, адекватно отражающую зависимость свойств кодированной последовательности от параметров используемого помехоустойчивого кода и позволяющую сформировать эталонные описания минимальной размерности для помехоустойчивых кодов различных классов.

2. Статистический подход к идентификации помехоустойчивых кодов

Обнаружение и идентификация СКК в условиях априорной неопределённости относительно структуры и параметров кодирующего устройства требует выбора в качестве информативных наиболее общих (инвариантных) признаков, в наименьшей степени зависящих от уровня имеющейся неопределенности, характера передаваемой информации и вероятности ошибки в канале. Для алгебраических методов анализа в качестве информативных признаков выступает наличие и характер линейных связей

между символами кодированной последовательности, обусловленные используемым помехоустойчивым кодом. Достоинством таких методов является возможность вычисления порождающей матрицы кода непосредственно из принятой последовательности, а основные недостатки заключаются в чувствительности к ошибкам и необходимости точного восстановления маркировки μ при приеме.

Наименьшей чувствительностью к наличию ошибок в принятой последовательности и возможностью частичной идентификации помехоустойчивого кода при произвольном задании маркировки сигнальных точек отличаются статистические методы анализа [9, 10, 11]. Использование статистического подхода к обнаружению и оцениванию параметров помехоустойчивого кода основано на предположении, что в некодированной случайной дискретной последовательности любой фрагмент произвольной длины t появляется с одинаковой вероятностью, определяемой только длиной этого фрагмента. Избыточность, вносимая помехоустойчивым кодером, приводит к сокращению числа возможных последовательностей определенной длины и, как следствие, изменению вероятности появления различных комбинаций в кодированной последовательности.

Наиболее наглядным является случай блоковых кодов, например, простого кода с проверкой на четность, где проверочный разряд принимает значение 0 или 1 в зависимости от того, является ли четным или нечетным количество единиц в кодовом слове. Таким образом, значение проверочного разряда (бита четности) однозначно зависит от информационных. В результате, если число возможных двоичных комбинаций длины $(t-1)$ в случайной последовательности составляет 2^{t-1}, то при добавлении проверочного разряда и увеличении длины комбинации до t, число возможных («разрешенных») комбинаций на выходе кодера не изменится. Соответственно, число возможных двоичных комбинаций длины t при наличии в неискаженной исследуемой последовательности кода с проверкой на четность в два раза меньше возможного числа комбинаций той же длины в случайной последовательности.

Экспериментальные исследования, выполненные с применением различных статистических тестов, например, NIST [12] подтверждают, что помехоустойчивое кодирование практически не изменяет закона распределения отдельных элементов случайной дискретной последовательности (символов), но оказывает влияние на распределение блоков символов длины $t > t_{min}$, где t_{min} определяется параметрами кода.

Подход к обнаружению и оцениванию параметров помехоустойчивых кодов, основанный на анализе распределения фрагментов дискретной последовательности различной длины исследовался, например, в работе [9]. При этом ряд распределения различных комбинаций можно рассматривать как вектор количественных признаков конкретного помехоустойчивого кода, заданных в пространстве, размерность которого экспоненциально зависит от длины фрагмента. Основными недостатками данного подхода является зависимость распределения от выбранной маркировки и, как следствие, невозможность построения обобщенной аналитической модели, а также практическая нереализуемость процедуры оценивания эмпирического закона распределения фрагментов длиной более 24 двоичных символов из-за чрезвычайно большого (более 10^{12} бит) требуемого объема выборки. Таким образом, применение на практике статистических методов идентификации помехоустойчивых кодов требует снижения размерности вектора измерений и описания помехоустойчивых кодов.

3. Вносимая избыточность и энтропийные модели помехоустойчивых кодов

Наиболее рациональным представляется использование в математической модели помехоустойчивого кодирования не закона распределения кодовых комбинаций, а тесно связанной с ним фундаментальной характеристики, присущей помехоустойчивым кодам – избыточности, вносимой кодером. Данная характеристика является интегральной, то есть, характеризует соотношение между «разрешенными» и «запрещенными» кодовыми

комбинациями в целом, что позволяет существенно сократить пространство признаков. Поскольку отображение принятых сигнальных точек на множество двоичных комбинаций является взаимно однозначным (биективным), выбор произвольной маркировки на приеме изменяет конкретные значения «разрешенных» и «запрещенных» кодовых комбинаций, но не изменяет соотношения между ними. Следовательно, избыточность является характеристикой, инвариантной к выбранной маркировке.

Наряду с относительной избыточностью, определяемой выражением $\frac{n-k}{n}=(1-R)$, кодированная дискретная последовательность может быть охарактеризована абсолютной избыточностью $\mathcal{D}$ – количеством избыточных символов на блок фиксированной длины. С точки зрения теории информации, под абсолютной избыточностью понимается разность между максимально возможным и действительным количеством информации, передаваемым сообщением, состоящим из символов некоторого алфавита произвольной мощности. Для канала связи без памяти, абсолютная избыточность соответствует разности между максимальным уровнем энтропии (логарифмом мощности алфавита) и уровнем энтропии источника [13]. Обозначив через $r=\lim\limits_{n\to\infty}\frac{1}{n}H\left(M_1,M_2,\ldots M_n\right)$ среднюю энтропию на символ сообщения M_i, а через $H_{max}=\log\left|M\right|$ – максимальную энтропию, выражение для абсолютной избыточности на символ сообщения может быть записано как

$$\mathcal{D}=H_{max}-r\,. \qquad (1)$$

Рассматривая в качестве символов сообщения M_i двоичные комбинации длины t, и выбрав двойку в качестве основания логарифма для вычисления энтропии, получим $H_{max}=\log_2 2^t=t$. Тогда значение $\mathcal{D}=t-r$ будет соответствовать числу избыточных (зависимых) двоичных символов в кодовой комбинации длины t. Это число зависит от параметров кода, причем для непрерывных кодов, где кодовые слова потенциально могут иметь бесконечную

длину, более информативной является зависимость $\mathcal{D}(t)$ числа избыточных символов, содержащихся в произвольной двоичной комбинации длины t на выходе кодера от длины комбинации. Поскольку наряду с избыточностью, вносимой помехоустойчивым кодером, исходной информационной последовательности присуща некоторая собственная избыточность, значение $\mathcal{D}(t)$ можно представить как сумму $\mathcal{D}(t) = \mathcal{D}_s(t) + \mathcal{D}_c(t)$, где $\mathcal{D}_s$ – избыточность источника, а $\mathcal{D}_c$ – избыточность, вносимая помехоустойчивым кодером. После кодирования источника и рандомизации энтропия источника максимальна, поэтому можно считать, что для двоичной последовательности на входе помехоустойчивого кодера $\mathcal{D}_s(t) \approx 0$ и $\mathcal{D}(t) \approx \mathcal{D}_c(t)$. Так как в соответствии с выражением (1) зависимость $\mathcal{D}_c(t)$ характеризует уменьшение абсолютного уровня энтропии в зависимости от длины кодовых комбинаций t (разрядности символов алфавита), будем называть ее энтропийной моделью помехоустойчивого кода.

Зависимость $\mathcal{D}_c(t)$ имеет различный характер для блоковых и непрерывных (сверточных) кодов. Процесс блокового помехоустойчивого кодирования может быть представлен как инъективное отображение из множества информационных (входных) векторов размерности k, в множество кодовых слов длины n. Соответственно, число избыточных символов в кодовом слове равно $(n-k)$, и энтропийная модель $\mathcal{D}_{cB}(t)$ блочного кода с параметрами (n,k) определяется выражением:

$$\mathcal{D}_{cB}(t) = \begin{cases} (n-k)\left\lfloor \dfrac{t}{n} \right\rfloor, & t \bmod n \le k \\ (n-k)\left\lfloor \dfrac{t}{n} \right\rfloor + t \bmod n - k, & t \bmod n > k \end{cases}. \qquad (2)$$

Минимальная длина двоичных комбинаций с выхода кодера, на которой начинает проявляться вносимая кодером избыточность, на один символ больше числа информационных символов k, а длина комбинации t_{min}, на которой вносимая избыточность достигает значения $(n-k)$ равна длине кодового слова

n. Таким образом, зависимость $\mathcal{D}_{cB}(t)$ имеет вид ступенчатый кривой (рис. 1) и значение $\mathcal{D}_{cB}(t)$ увеличивается на $(n-k)$ бит при увеличении длины комбинации t на n символов.

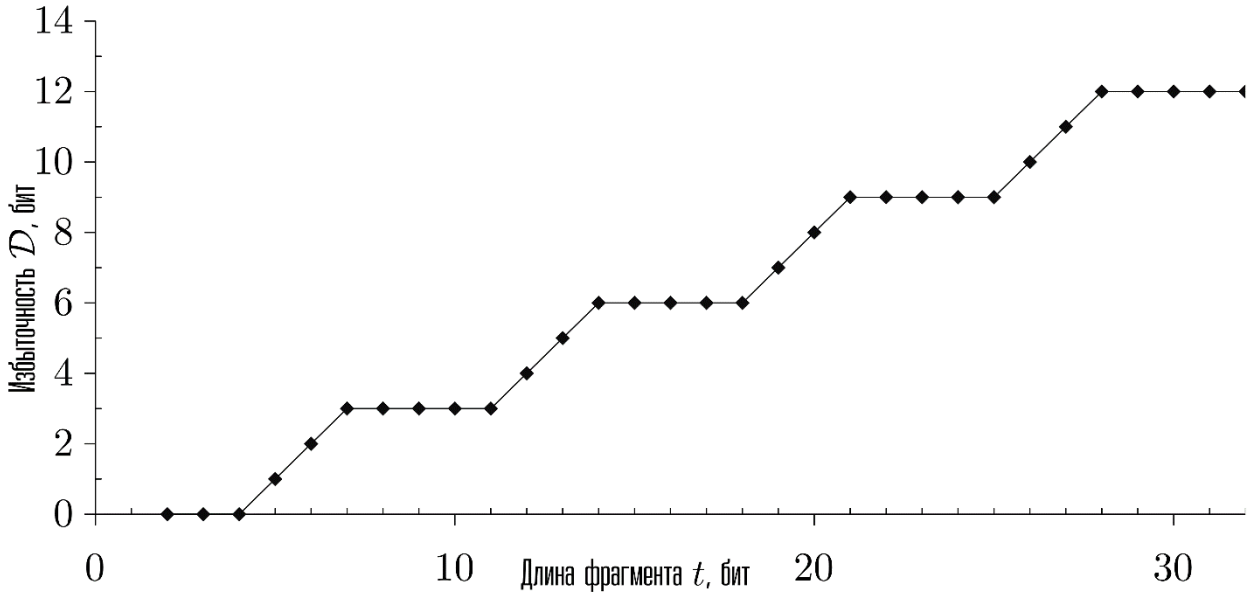

Рис. 1. Зависимость вносимой избыточности от длины фрагмента для блокового кода Хэмминга (7, 4).

В отличие от блокового, кодер непрерывного (сверточного) кода содержит память, объем которой определяет глубину линейных связей между символами кодовой последовательности [14]. Наиболее распространенная реализация кодера включает в себя регистры сдвига, число которых равно числу входных информационных символов (входов кодера) k, а n выходных символов вычисляется как линейные комбинации предшествующих входных, находящихся в ячейках регистров. При этом длина регистров (память) и номера ячеек, используемых для вычисления выходных символов определяются порождающими полиномами кода $g_{i,j}(D)$, где D – оператор задержки, i – номер входа, j – номер выхода.

Памятью i-го входа сверточного кодера ℓ_i называется максимальная степень порождающего полинома $g_{i,j}(D)$:

$$\ell_i = \max_{j=1,2,\ldots,n} \deg g_{i,j}(D), \forall i = 1,2,\ldots,k.$$

Память кода ℓ определяется максимальной степенью среди всех порождающих полиномов кода: $\ell = \max\limits_{i=1,\ldots,k} \ell_i = \nu - 1$, где ν – длина кодового ограничения. Общее число элементов (ячеек) памяти кодера $\ell^{\perp} = \sum\limits_i \ell_i$ будем называть полной памятью сверточного кодера.

В [15] показано, что множество начальных кодовых последовательностей сверточного кода образует линейный блоковый код с параметрами (n_0, k_0). Длина слова данного кода n_0 и количество информационных символов в слове k_0 пропорциональны кодовому ограничению ν: $n_0 = \nu n,\ k_0 = \nu k$, соответственно, из νn начальных символов с выхода сверточного кодера, $\nu(n-k)$ будут избыточными. Наличие вносимой избыточности (уменьшение энтропии) для сверточных кодов начинает проявляться на кодовых комбинациях, длина которых равна минимальной длине проверочного уравнения кода, при этом избыточность достигает $(n-k)$ бит на длине комбинации t_{min}, определяемой в соответствии со следующим выражением:

$$t_{min} = l_{min} \cdot n, \text{ где } l_{min} = \left\lfloor \frac{\ell^{\perp}}{n-k} + 1 \right\rfloor.$$

Значение t_{min} также соответствует минимальной длине фрагмента кодовой последовательности, на которой проявляются линейные связи между символами и тесно связано с явлением, называемым «дефект ранга», лежащим в основе большинства алгебраических методов идентификации сверточных кодов. Для наиболее распространенных кодов с относительной скоростью равной $\frac{n-1}{n}$, l_{min} совпадает с длиной кодового ограничения ν.

Для непрерывного (сверточного) кодера с параметрами (n, k, l_{min}), зависимость вносимой избыточности $\mathcal{D}_{cR}$ от длины кодовой комбинации t определяется следующим выражением:

$$\mathcal{D}_{cR}(t)=\begin{cases}0, & t<n(l_{min}-1)\\ \dfrac{n-k}{n}\left(t-n(l_{min}-1)-t \bmod n\right) & t\geq n(l_{min}-1), t \bmod n\leq k\\ \dfrac{n-k}{n}\left(t-n(l_{min}-1)-t \bmod n\right)+t \bmod n-k & t\geq n(l_{min}-1), t \bmod n>k\end{cases} \quad (3)$$

Начиная с $t>n\cdot(l_{min}-1)$, данная зависимость также имеет возрастающий ступенчатый характер (рис. 2). Произведение $t_Z=n\cdot(l_{min}-1)$ определяет длину начального «безызбыточного» (нулевого) участка на графике, а разность $(n-k)$ – число добавочных избыточных символов на каждый последующий n-символьный интервал. Другими словами, при увеличении длины кодовой комбинации $t>t_Z$, значение вносимой абсолютной избыточности возрастает с коэффициентом $\left(1-\dfrac{k}{n}\right)$ на участках подъема ступенчатой кривой.

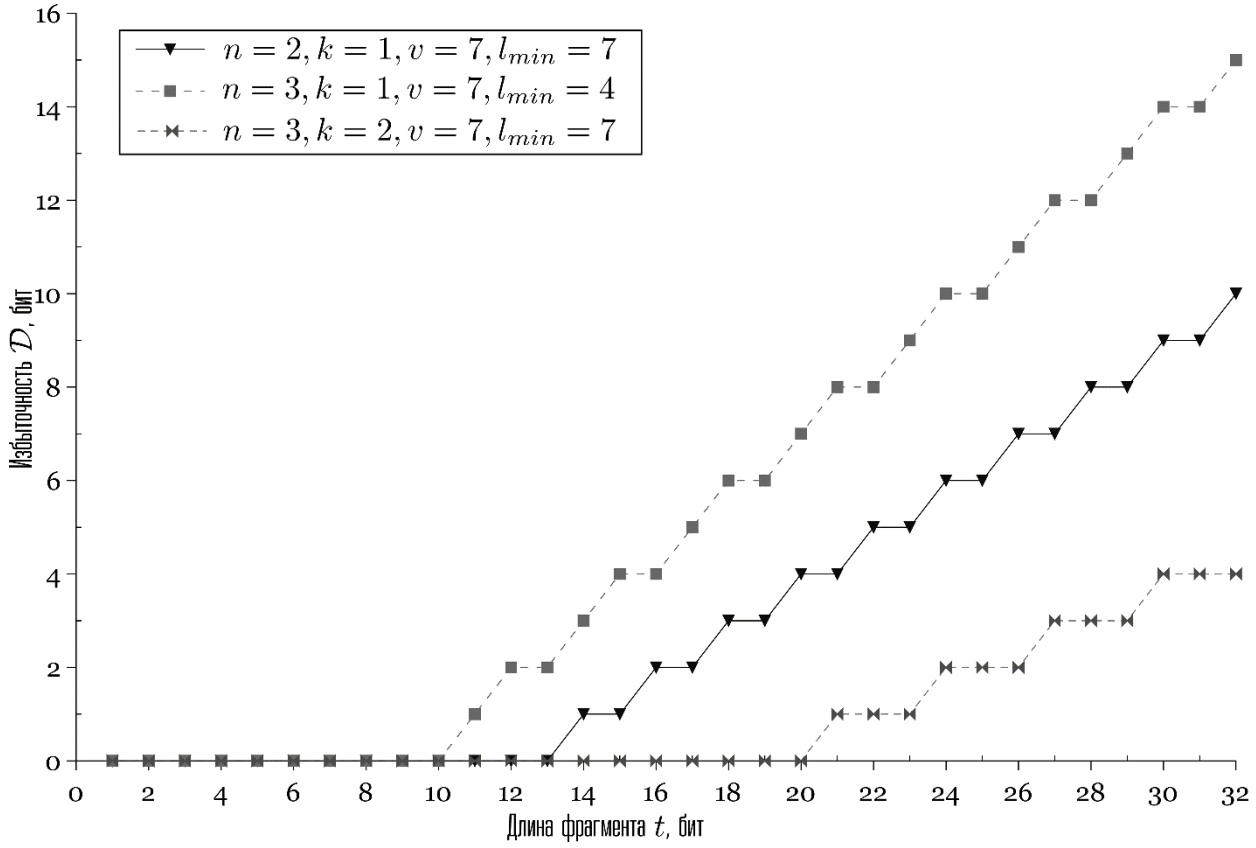

Рис. 2. Зависимость вносимой избыточности от длины фрагмента для различных сверточных кодов.

Поскольку характер зависимости вносимой абсолютной избыточности от длины кодовой комбинации t несет информацию о виде помехоустойчивого кода (блочный или непрерывный) и его параметрах (n,k,l_{min}), он может использоваться для формирования эталонных описаний кодов, используемых при решении задач обнаружения и идентификации сигнально-кодовых конструкций.

4. Энтропийные сигнатуры помехоустойчивых кодов

Для упрощения формализованного описания помехоустойчивых кодов различных классов на основе энтропийных моделей, предлагается вместо непосредственных значений вносимой избыточности $\mathcal{D}_c(t)$ использовать конечные разности первого порядка $\Delta\mathcal{D}_i = \mathcal{D}_c(t_{i+1}) - \mathcal{D}_c(t_i), i = 0,1,2,\ldots$. Достоинством такого подхода является то, что конечные разности первого порядка $\Delta\mathcal{D}_i$ принимают лишь два значения – 0 или 1 независимо от вида и параметров помехоустойчивого кода (рис. 3). Это обусловлено тем, что приращение избыточности в соответствии с (2) или (3) происходит дискретно и составляет один бит при увеличении длины кодовой комбинации t на один двоичный разряд.

Свойства конечных разностей функции избыточности вытекают из свойств соответствующего помехоустойчивого кода. Поскольку для блоковых кодов функция избыточности $\mathcal{D}_c(t)$ имеет вид ступенчатой кривой со ступеньками длины n, состоящими из горизонтального участка длины k и возрастающего участка длины $(n-k)$, зависимость $\Delta\mathcal{D}(t)$ является периодической с периодом, равным n, на котором присутствует k нулевых и $(n-k)$ единичных значений. Для помехоустойчивых кодов с памятью характерным является наличие начального нулевого участка, выделенного жирной линией на рис. 3, длина которого t_Z определяется полной памятью кодера. Начиная со значения с номером (t_Z+1), зависимость приобретает периодический характер, с периодом равным n, как и для блоковых кодов, что на рис. 3 показано дополнительным выделением одного периода.

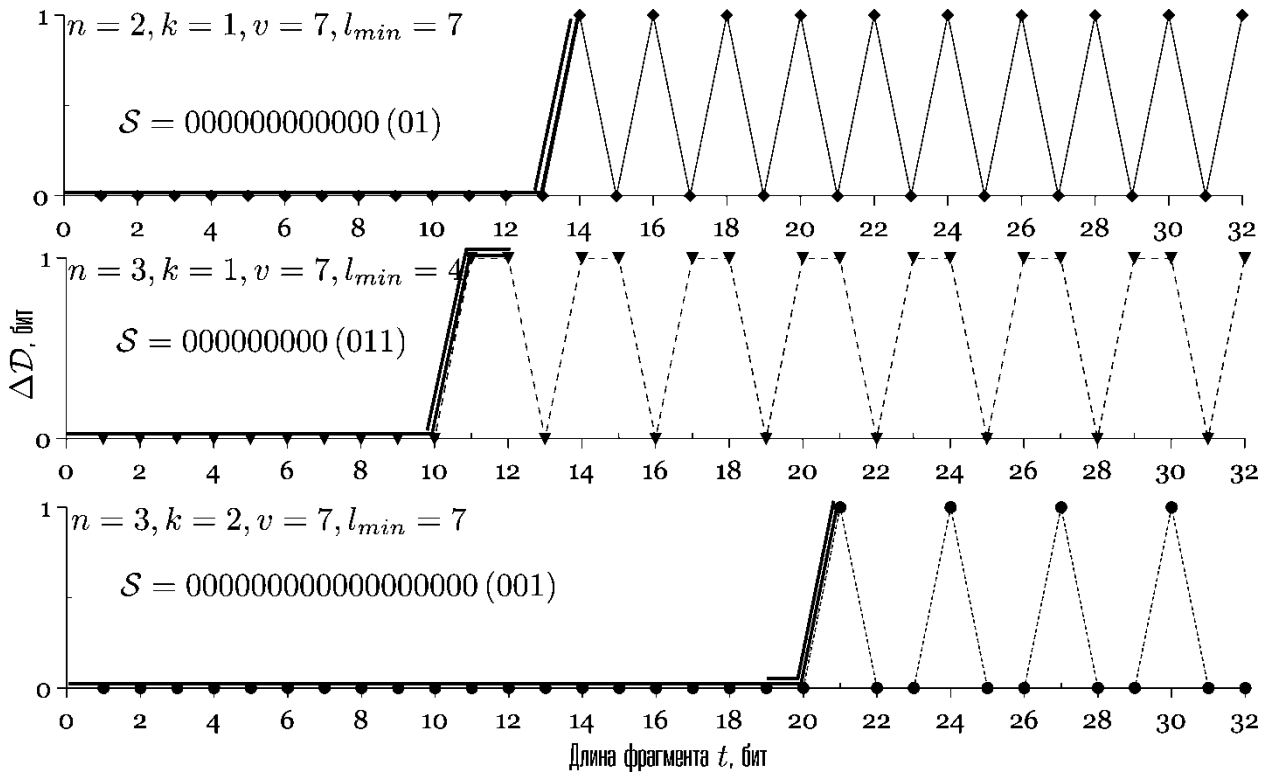

Рис. 3. Конечные разности функции избыточности и энтропийные сигнатуры различных сверточных кодов.

Таким образом, записав t_{min} первых значений $\Delta\mathcal{D}_i$, получим последовательность $\mathcal{S}=\left\langle \Delta\mathcal{D}_1,\ldots,\Delta\mathcal{D}_{t_{min}} \right\rangle$, соответствующую множеству помехоустойчивых кодов с одинаковыми параметрами (n,k) для блоковых или (n,k,l_{min}) для сверточных кодов. Основой для вычисления данной последовательности является аналитическая энтропийная модель, поэтому будем называть ее **энтропийной сигнатурой** (ЭС) помехоустойчивого кода.

Соответственно, ЭС минимальной размерности $\mathcal{S}$ представляет собой комбинацию начальной нулевой последовательности длины $t_Z = n\cdot\left(l_{min}-1\right)=n\cdot\left(\left\lfloor\frac{\ell^{\perp}}{n-k}+1\right\rfloor-1\right)$ и одного периода длины n (рис. 3). Поскольку она отражает лишь часть параметров помехоустойчивого кода, энтропийная сигнатура может рассматриваться как эталонное описание некоторого **класса** помехоустойчивых кодов с одинаковыми параметрами

(n,k,l_{min}) и, возможно, различными порождающими матрицами.

Предлагаемый подход не накладывает каких-либо ограничений на вид используемого помехоустойчивого кода. Рассмотренный способ задания кода с помощью порождающей матрицы **G** выбран с учетом возможного использования алгебраических методов на дальнейших этапах анализа и не является единственно возможным. Значения $\mathcal{D}_c(t)$ и $\Delta\mathcal{D}(t)$ не зависят от положения проверочных символов в кодовой последовательности, а зависят только от соотношения между числом информационных и проверочных символов на длине комбинации t. Таким образом, систематические и несистематические коды с одинаковыми параметрами (n,k,l_{min}) будут иметь одинаковые ЭС.

В процессе манипуляционного кодирования, двоичные комбинации с выхода помехоустойчивого кодера могут разделяться на части, дополняться некодированными символами и передаваться несколькими элементарными сигналами. В результате уменьшается значение вносимой избыточности и изменяется структура кодовой последовательности. Это не является препятствием к использованию ЭС при идентификации СКК, поскольку статистические связи между двоичными символами с выхода помехоустойчивого кодера обуславливают наличие таких связей между канальными символами разрядности a. Энтропийные сигнатуры для многомерных сигнально-кодовых конструкций могут строиться на основе выражений (2) или (3) с учетом увеличения длины кодового слова n и относительной скорости кодирования R.

В общем виде, задача идентификации СКК с использованием энтропийных сигнатур решается следующим образом. Принятая последовательность $\tilde{Z}_{\langle N\rangle}$ делится на фрагменты длины t, по которым вычисляются оценки $\tilde{\mathcal{D}}(t)$ и конечные разности $\Delta\tilde{\mathcal{D}}(t)$. Для непрерывных кодов фрагменты могут частично перекрываться. Диапазон значений t и сдвиг между фрагментами определяется исходя из имеющихся эталонных (модельных) ЭС для известных классов кодов

и СКК. Наилучшая оценка ЭС $\tilde{S}$ будет получена при совпадении начала первого фрагмента с началом кодового слова, поэтому в процессе анализа выполняется поиск границ кодовых слов. Полученная эмпирическая ЭС $\tilde{S}$ сравнивается с модельными и выбирается модельная ЭС, среднеквадратическая ошибка оценивания для которой минимальна. Для неизвестных классов кодов задача идентификации может быть решена путем подбора параметров (n,k,l_{min}) в выражениях (2) или (3) до получения наименьшего расхождения между модельной и эмпирической ЭС.

Для повышения достоверности идентификации в условиях наличия ошибок в принятой последовательности, число повторяющихся (периодических) элементов ЭС, используемых при вычислении $\tilde{S}$ может быть увеличено. Результаты экспериментальных исследований показали, что использование энтропийных моделей помехоустойчивых кодов в сочетании с разработанным методом оценивания вносимой избыточности позволяет сократить требуемый объем выборки и количество вычислений при решении задачи идентификации СКК на несколько порядков по сравнению с известными статистическими методами анализа.

Заключение

Использование эталонных описаний помехоустойчивых кодов на основе аналитической энтропийной модели позволяет значительно сократить признаковое пространство, и как следствие, снизить вычислительную сложность процедур технического анализа и уменьшить требуемый объем выборки по сравнению с известными статистическими методами. Существенное снижение требований к объему выборки для анализа обеспечивается за счет перехода от построения рядов распределения кодовых комбинаций различной длины к оцениванию избыточности, вносимой помехоустойчивым кодером. Предложенный подход может быть положен в основу метода обнаружения и идентификации СКК. На первом этапе метода производится обнаружение и оценивание параметров используемого

помехоустойчивого кода на основе сравнения эмпирических ЭС с модельными. На последующих этапах полученные оценки используются для восстановления манипуляционного кода и порождающей матрицы помехоустойчивого кода алгебраическими методами.

Литература

1. Замарин А.И., Атакищев О.И., Тавалинский Д.А., Рюмшин К.Ю. Последетекторный технический анализ цифровых последовательностей при идентификации сложных структур // Известия юго-западного государственного университета, 2014, № 1 (52), с. 14 – 21.
2. Григорьев В.А., Григорьев С.В. Сигнально-кодовые конструкции – СПб.: ВАС, 1997. – 147 с.
3. Баушев С.В., Зайцев И.Е., Яковлев А.А. Перспективы развития сигнально-кодовых конструкций для гауссовского канала связи // Зарубежная радиоэлектроника, 1990, № 1, с. 15 – 31.
4. Кадуков Е.П. Модель радиосигналов с модуляцией с непрерывным изменением в пространстве параметров фазовых диаграмм и комплекс информативных признаков для распознавания видов модуляции излучений зарубежных спутниковых систем связи // Журнал радиоэлектроники [электронный журнал]. 2019. № 7. Режим доступа: http://jre.cplire.ru/jre/jul19/12/text.pdf
5. Marazin M., Gautier R., Burel G. Dual Code Method for Blind Identification of Convolutional Encoder for Cognitive Radio Receiver Design // 2009 IEEE Globecom Workshops, Honolulu, HI, 2009, pp. 1-6. DOI 10.1109/GLOCOMW.2009.5360726
6. Zrelli Y., Marazin M., Rannou E., Gautier, R. Blind Identification of Convolutional Encoder Parameters over GF(2m) in the Noiseless Case // Computer Communications and Networks (ICCCN) 2011 Proceedings of 20th International Conference on, pp. 1-5.

7. Xie H., Chai X., Wang F., Huang Z. A method for blind identification of rate 1/2 convolutional code based on improved Euclidean algorithm // Signal Processing (ICSP) 2012 IEEE 11th International Conference on, vol. 2, pp. 1307-1310.
8. Баринов А. Ю. Перемежение в канальном кодировании: свойства, структура, специфика применения // Журнал радиоэлектроники [электронный журнал]. 2019. №1. Режим доступа: http://jre.cplire.ru/jre/jan19/13/text.pdf
9. Ратушин А.П., Рачинский Е.В., Балунин Е.И. Признаки распознавания непрерывных корректирующих кодов в дискретных последовательностях // Наукоемкие технологии, 2010, № 9, с. 20 – 23.
10. Sicot G., Houcke S. Etude statistique du seuil dans la detection d'entrelaceur // Proc. GRETSI 2005, Belgium, pp. 231-254.
11. Bellard M., Tillich J.P. Detecting and reconstructing an unknown convolutional code by counting collisions // 2014 IEEE International Symposium on Information Theory (ISIT 2014), pp. 2967-2971.
12. Rukhin A., Soto J. A Statistical Test Suite for Random and Pseudorandom Number Generators for Cryptographic Applications. [Электронный ресурс]. Режим доступа: https://csrc.nist.gov/publications/detail/sp/800-22/rev-1a/final.
13. MacKay D. Information Theory, Inference, and Learning Algorithms. – Cambridge University Press, 2003 – 640 p.
14. Rosenthal J. Connections between linear systems and convolutional codes // Codes, Systems and Graphical Models. The IMA Volumes in Mathematics and its Applications, 2001, vol. 123, pp. 39 – 66.
15. Месси Дж. Пороговое декодирование – М.: Мир, 1966. – 208 с.

Для цитирования:
К. Б. Махров, В. О. Хацаюк. Энтропийные модели и эталонные описания помехоустойчивых кодов. Журнал радиоэлектроники [электронный журнал]. 2019. № 10. Режим доступа: http://jre.cplire.ru/jre/oct19/1/text.pdf
DOI 10.30898/1684-1719.2019.10.1