

УДК 621.37:621.391

Т.В. Лавровська, С.Г. Рассомахин

Харьковский национальный университет имени В.Н. Каразина, Харьков

ФИЗИЧЕСКАЯ МОДЕЛЬ ПСЕВДОСЛУЧАЙНЫХ КОДОВ В МНОГОМЕРНОМ ЭВКЛИДОВОМ ПРОСТРАНСТВЕ

В статье рассмотрены основные проблемы практического использования псевдослучайных кодов в современных системах передачи информации. Проанализированы существующие методы генерации псевдослучайных последовательностей с учетом практических особенностей их применения. Разработана физическая модель псевдослучайных кодов в евклидовом пространстве, которая представляет переносчики кодовых слов в виде отрезков реализаций псевдослучайных процессов с ограниченным спектром.

Ключевые слова: псевдослучайные коды, линейный конгруэнтный генератор, регистр сдвига с линейной обратной связью, евклидово пространство кода, физическая модель сигналов.

Введение

Постановка проблемы. В результате теоретических и практически реализованных работ, начиная с работ Шеннона и до сегодняшних дней, созданы информационные технологии, которые полностью изменили жизнь людей, к ним можно отнести глобальную сеть Internet, мобильную и спутниковую телефонную связь. В связи с этим из года в год возрастает количество технических средств обработки и передачи информации, работающие в сетях беспроводной связи. В связи с этим постоянно возникают проблемы рационального использования радиочастотного ресурса, появляются потребности в увеличении скорости передачи, а также обеспечение достаточно высокого уровня помехозащищенности данных в сетях со случайными и умышленными угрозами. Решением, перечисленных выше проблем, является построение эффективных систем передачи информации (СПИ).

Для улучшения качественных характеристик СПИ необходимо повышение показателей удельной частотной и энергетической эффективности. Одной из значительных проблем создания эффективных систем телекоммуникации является правильный выбор методов защиты данных от случайных ошибок. Применение алгебраических кодов для исправления ошибок влечет за собой необходимость введения избыточности и, как следствие, — снижение реальной скорости передачи. Кроме того, вычислительная простота построения и декодирования таких кодов достигается за счет ухудшения их корректирующих свойств в результате отказа от метода максимального правдоподобия при декодировании.

Математическим аппаратом теоретических исследований Шеннона [1] в области надежной связи является использование случайно выбираемых кодов, которые позволяют достичь монотонного снижения вероятности декодирования с ошибкой при увеличении длины блока кода.

Практически до сегодняшнего дня случайные (псевдослучайные) коды (ПСК) для обеспечения помехоустойчивости и секретности не используются. Это является следствием отсутствия приемлемых по вычислительной сложности методов декодирования, обеспечивающих корректирующую способность, близкую к максимальному правдоподобию. Кроме того, реализация конструктивных алгоритмов построения ПСК может быть получена только при использовании детерминированных алгоритмов генерации псевдослучайных символов кодовых слов.

Привлекательность технологий ПСК заключается в возможности создания сигнально-кодовых конструкций, которые позволяют одновременно повысить как частотную, так и энергетическую эффективность СПИ. Основной преградой для широкого использования ПСК является отсутствие небрежных методов декодирования.

Вычислительная сложность алгоритмов, основанных на вычислении евклидовых расстояний возрастает экспоненциально с увеличением длины блока кода и при практически требуемых значениях длины блока является неприемлемой. Получение простых линейных алгебраических методов декодирования наталкивается на трудности, вытекающие из нелинейности детерминированных алгоритмов генерации ПСП. Таким образом, можно утверждать, что идеи применения ПСК могут найти конструктивное воплощение в случае, если будут найдены линейные (линеаризованные) методы декодирования таких кодов.

Цель работы: анализ, существующих методов генерации ПСП и обоснование наиболее эффективных методов построения и декодирования ПСК.

Изложение основного материала

1. Основные свойства случайных кодов

Общая модель случайных кодов К. Шеннона [1] дает возможность построения СПИ, достигающих од-

новременно высоких значений показателей удельной частотной и энергетической эффективности.

Рассмотрим модель построения ПСК. Пусть передаче подлежит двоичная последовательность, предварительно разбитая на блоки длиной k символов.

Каждый из таких блоков может содержать $m = 2^k$ различных комбинаций, которые могут быть пронумерованы в лексикографическом порядке числами i из диапазона $i \in [0, m-1]$.

Каждому блоку информационной последовательности длиной k бит ставится в соответствие кодовое слово длиной n недвоичных целых чисел из фиксированного диапазона $[0, m-1]$. При построении физической модели ПСК значения этих символов кодового слова ПСК определяют (с учетом масштабирования по выделенному бюджету энергии на передачу k бит) значение соответствующего информативного параметра сигнала. В соответствии с данным описанием появляется возможность определения скорости кода в виде:

$$V = \frac{k}{n}. \quad (1)$$

Поскольку длина кодового слова ПСК n , фактически, может выбираться независимо от длины исходного блока двоичных символов k , то скорость кода (1) может быть как больше, так и меньше единицы:

$$V \lesseqgtr 1. \quad (2)$$

При передаче информации, в случае появления на выходе источника блока из k двоичных символов с номером i , в канал выдается последовательность из n недвоичных символов, получаемых по некоторому псевдослучайному алгоритму генерации. В этом случае номер i удобно рассматривать, как некоторое число инициализации алгоритма генерации ПСП.

Процесс декодирования ПСП сводится к очередному сравнению полученных канальных кодовых слов, искаженных помехами, с эталонами, хранящимися в кодовой книге приемника или генерируемыми в реальном масштабе времени по известному алгоритму, идентичному алгоритму кодера. Эталонное кодовое слово, расположенное на наименьшем евклидовом расстоянии от принятого, полагается за истинное. Исходный блок источника, содержащий k бит, однозначно восстанавливается записью в двоичном виде первого символа найденного истинного кодового слова, поскольку данный символ является лексикографическим номером последовательности источника и использовался для инициализации алгоритма генерации ПСП. Число используемых выборок ПСП будет зависеть от допустимой частоты ошибок ε :

$$\lim_{\varepsilon \rightarrow 0} \lim_{n \rightarrow \infty} \frac{\log 2^k(\varepsilon, n)}{n} = F \cdot \log \frac{P+N}{N}, \quad (3)$$

где F – ширина полосы пропускания; P – средняя мощность сигнала; N – средняя мощность помех.

Описанный выше алгоритм кодирования/декодирования ПСК представляет собой конструктивную реализацию наилучшего кода, который, при увеличении длины блока n и фиксированной скорости V обеспечивает асимптотическое приближение к пропускной способности (3).

Следует отметить, подобное приближение невозможно при использовании традиционных методов построения и декодирования регулярных алгебраических кодов.

Для практического осуществления описанного алгоритма кодирования необходимо обосновать выбор способа генерации последовательностей псевдослучайных чисел на основе использования какого-либо детерминированного алгоритма.

2. Анализ методов построения ПСК с использованием детерминированным способов генерации ПСП

Рассмотрим два основных, наиболее распространенных метода генерации псевдослучайных последовательностей, нашедших свое применение в приложениях криптографии.

Способ, основанный на применении многотактных линейных регистров сдвига (LinearFeedbackShiftRegister, LFSR) [2].

В качестве начальной инициализации для элементов вектора выбираются числа из диапазона $R_i \in [0, 2^k - 1]$, где k – это степень примитивного многочлена $P(x)$, который определяет структуру LFSR генератора. Следует заметить, что степень этого многочлена удобно выбирать равной количеству бит в исходном кодированном сообщении. В регистрсдвига при его инициализации помещается двоичный код k -битового блока символов источника. В результате осуществления n тактов сдвига генерируется последовательность из n канальных символов, которая является соответствующим n -символьным отрезком полной ПСП максимального периода.

Реализация генератора ПСП на регистре сдвига, текущее состояние которого обозначается в дальнейшем вектором $R = \{R_0, \dots, R_{n-1}\}$, основывается на применении LFSR с базовой конфигурацией Фибоначчи.

Пусть $C = \{c_0, c_1, \dots, c_k\}$ – вектор, задающий структуру генератора ПСП, тогда порождающий ПСП многочлен можно записать в виде:

$$P(x) = \sum_{i=0}^k C_i \cdot h_i; \quad C_i \in [0, 1]. \quad (3)$$

Введем вектор S , содержащий $q+1$ элементов из диапазона $[0, k]$, каждый из которых равен индексу ненулевого элемента вектора C . Величина q определяется числом резидентных обратных связей в структуре LFSR. Вектор S всегда содержит не

менее двух элементов, так как нулевая и старшая степень переменной x всегда присутствует в полиноме (3). Оправданным для S является название "вектор отводов", используемое в дальнейшем. Используемая для построения LFSR ПСК схема сдвигового генератора представлена на рис. 1.

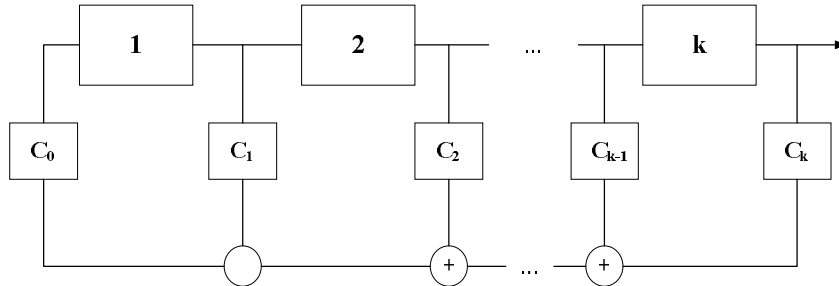


Рис. 1. Схема генератора ПСК с помощью LFSR

В процессе генерации происходит потактовый сдвиг содержимого ячеек памяти регистра право на 1 позицию.

Содержимое младшей ячейки памяти регистра формируется в соответствии с вектором отводов s . Текущее значение символа ПСК считывается на каждом такте из разрядов регистра R после завершения очередного сдвига. В младшую ячейку заносится результат вычисления R_0 в соответствии со следующей формулой:

$$R_0 = \sum_{j=1}^q R_{(S_j-1)} \cdot \quad (4)$$

В результате, инициализации вектора R двоичным числом i за конечное количество тактов n , генерируется вектор, координатами которого являются псевдослучайные числа $Z_i = \{z_{i,1}, z_{i,2}, \dots, z_{i,n}\}$. Перебор i от 0 до $2^k - 1$ дает возможность получить кортеж последовательностей $Z = \{Z_0, Z_1, \dots, Z_{2^k-1}\}$, составляющих кодовую книгу ПСК.

Декодирование ПСК, генерируемого на основе применения изложенного метода, сводится к нахождению лексикографического номера ПСП. Этот номер $x_0 = Z_0 = \{z_{0,1}, z_{0,2}, \dots, z_{0,n}\}$ (в двоичном виде) выступает в качестве инициализирующей битовой последовательности для регистра сдвига. Таким образом, декодирование кодового слова сводится к корректному определению порождающего числа путем факторизации принятой ПСП. Состояния регистра сдвига на каждом такте можно трактовать, как двоичную форму записи кодового символа ПСК. При инициализации LFS R нулевой последовательностью $x_0 = \{0, \dots, 0\}$ будет получено нулевое ко-

вое слово. Наличие нулевого кодового слова в книге ПСК является, в принципе, допустимым, однако, не желательным по причине возможного уменьшения

взаимного евклидова расстояния между кодовыми словами. Отмеченное можно отнести к недостаткам метода, использующего линейные регистры сдвига.

Аналитическая модель формирования кодовой книги ПСК LFSR может быть представлена в виде рекуррентного правила:

$$x_i = \text{mod} \left(2x_{i-1} + \sum_{q=1}^{N-1} C_q \text{mod} \left(\left(\text{mod} \left(\frac{x_{i-1}}{2^{N-q}}, 1 \right), 2 \right), 2 \right), 2 \right), \quad (5)$$

$$i \in [1, 2^k - 1]$$

где $C_i = [0, 1]$ – коэффициенты при степенях порождающего полинома; N – степень присутствующая в порождающем полиноме. Для получения полной кодовой книги набор кодовых слов, формируемых по правилу (5), должен быть дополнен нулевым кодовым словом.

Исходя из (5) удельное число нелинейных операций необходимых для генерации одного символа кодового слова равно трем, а значит для генерации кодового слова длиной n составляется $3 \cdot n$.

Способ, основанный на применении линейной конгруэнтной генерации [3]. Способ генерации ПСП с применением линейного конгруэнтного генератора (ЛКГ) является наиболее распространенным алгоритмом генерации ПСП в силу сочетания качеств простоты и эффективности.

Символами ПСК являются числа, выбираемые равномерно случайно из заданного диапазона $[0, 2^k - 1]$. Величина этого диапазона по-прежнему определяется длиной исходного блока двоичных символов источника k . Числовые значения символов ПСК при построении переносчиков в физическом канале пропорциональны значениям соответствующих информативных параметров сигналов (фаза, частота, амплитуда).

Генерация ПСП с использованием ЛКГ по сути осуществляемых действий похожа на рассмотренный процесс получения кодовых книг ПСК LFSR.

Построение кодового слова основывается на определении порождающего числа ЛКГ x_0 , которое является лексикографическим номером кодируемой двоичной последовательности источника, а также, одновременно, первым символом кодового слова ПСК. По рекуррентной цепочке вычислений генерируется $n-1$ кодовых символов блока кода:

$$x_{i+1} = \text{mod}(a \cdot x_{i-1} + b, m) \quad (6)$$

где x_{i+1} – очередное псевдослучайное число.

В соответствии с основными свойствами линейных конгруэнтных последовательностей параметры в выражении (6) должны удовлетворять требованиям:

- a, b, m – должны быть целыми положительными константами на всем периоде генерации ПСП;
- величина кратна любому простому числу и является делителем m ;
- $(a-1)$ кратно 4-м, если m кратно четырем;
- b и m – взаимно простые числа, причем $m \geq 2^k$.

Изменяя параметры a, b, m можно влиять на длину периода и на сами генерируемые значения символов ПСК x_i .

Произвольное i -ое кодовое слово имеет вид: $X_i = \{x_{i,0}, x_{i,1}, \dots, x_{i,n-1}\}$. Процесс декодирования кодовых слов, полученных по рассмотренному алгоритму ЛКГ ПСК также, как и для алгоритма LFSRPCK основан на оценке элементов вектора X_i для точного определения порождающего числа кодового слова $x_{i,0}$.

Параметр i однозначно определяется лексикографическим номером кодового блока источника. В отличие от рассмотренного ранее метода LFSR ПСК, выбор порождающего числа $x_0 = 0$ не приводит к генерации нулевого кодового слова, кодовая книга является полной и содержит все возможные ненулевые варианты последовательностей.

Удельное число нелинейных операций на вычисление одного символа блока ЛКГ ПСК, как следует из выражения (6), равно единице, соответственно для вычисления блока длиной n необходимо выполнить всего n операций вычисления.

Сравнение рассмотренных методов LFSR и ЛКГ ПСК, позволяет отдать предпочтение методу ЛКГ, поскольку количество нелинейных операций при получении кодового слова в 3 раза меньше.

3. Сравнительная оценка спектров взаимных расстояний для LFSR и ЛКГ ПСК

При передачи данных в канале с помехами, генерируемый код должен обладать достаточно высоким уровнем помехозащищенности. Данный пара-

метр напрямую зависит от величины взаимных расстояний кодовых слов, чем больше это значение, тем точнее декодер оценит принятое кодовое слово.

Оценка взаимных расстояний кодовых слов, сгенерированных на основе LFSR и ЛКГ, была проведена на основе алгоритмов, разработанных в среде MathCad.

По предложенным способам генерации кодовых слов генерируются кодовые книги с размерностью 2^k кодовых слов. Оценка взаимных кодовых расстояний основана на попарном сравнении всех возможных значений кодовых слов из диапазона $[0, 2^k - 1]$:

$$d_{i,j} = |X_i - X_j|, \quad (7)$$

где $d_{i,j}$ – взаимное расстояние между словами с номерами i и j в кодовых книгах LFSR и ЛКГ ПСК.

Результаты тестирования рассматриваемых моделей представлены на рис. 2 и 3 в виде гистограмм, которые отображают распределение взаимных расстояний кодовых слов LFSR и ЛКГ ПСК.

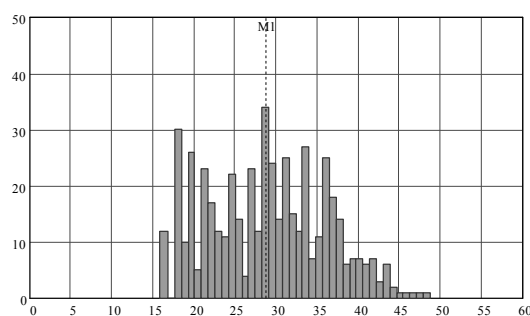


Рис. 2. Распределение кодовых слов ЛКГ ПСК при $n = 5$

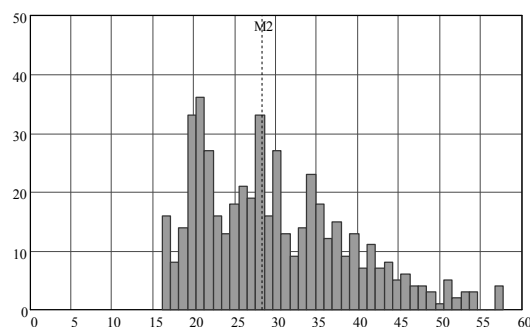


Рис. 3. Распределение кодовых слов LFSR ПСК при $n = 5$

Сопоставление полученных результатов статистического исследования ПСК LFSR и ЛКГ свидетельствует об их достаточной близости, поскольку медианы распределений, практически, совпадают. Учитывая рассмотренное ранее преимущество метода ЛКГ по числу выполняемых нелинейных операций, можно сделать вывод, о том, что использование ЛКГ в качестве генератора ПСК, является оптимальным решением [4].

4. Физическая модель ПСК ЛКГ

Обозначим $\overline{\text{KbL}} = \{\overline{\text{kbl}}_0, \overline{\text{kbl}}_1, \dots, \overline{\text{kbl}}_{2^k-1}\}$ – массив 2^k кодовых слов ПСК ЛКГ, где $\overline{\text{kbl}}_i = \{kw_0^i, \dots, kw_{n-1}^i\}$ – i -е кодовое слово, содержащее n символов. Каждое из кодовых слов массива является числовым кортежем для формирования информативных параметров сигналов. Рассмотрим наиболее простой способ формирования канальной формы сигналов ПСК, основанный на амплитудно-фазовой модуляции. Перед процедурой модуляции кодовые слова подвергаются преобразованию – центрированию и масштабированию числовых символов.

Операция центрирования производится для минимизации канальных затрат энергии и выполняется в соответствии со следующей формулой:

$$\overline{\text{KbL}} = \text{KbL} - \left(\frac{2^k - 1}{2} \right). \quad (8)$$

Средняя мощность (дисперсия) реализации кортежей кодовых слов может оказаться отличной от выделенного бюджета энергии на передачу одного символа (в случае, если $V=1$ – одного бита). Поэтому центрированный массив кодовой книги необходимо масштабировать, таким образом, чтобы среднее значение квадрата числовых символов кодовой книги не превышало выделенного бюджета мощности. Поскольку числа в равновероятных кодовых словах имеют дискретное равномерное распределение в диапазоне $\pm(2^k-1)/2$, то процесс масштабирования имеет вид:

$$\overline{\text{KbL}}' = \overline{\text{KbL}} \cdot \sqrt{\frac{12 \cdot P}{2^{2k} - 1}}. \quad (9)$$

Выполнение (8) и (9) гарантирует фиксированную среднюю мощность, приходящуюся на один символ каждого из 2^k кодовых слов, равную P .

Комплексная низкочастотная огибающая сигнала, передающего i -е кодовое слово, обладает длительностью T и спектром, ограниченным верхним значением $F = n/(2T)$. Эта огибающая может быть представлена рядом Фурье:

$$S_i(t) = \sum_{q=0}^{\frac{n}{2}-1} kwn_q^i \cdot \cos \left[2\pi \frac{q+1}{T} t \right] + \sum_{q=\frac{n}{2}}^n kwn_q^i \cdot \sin \left[2\pi \frac{q-\frac{n}{2}+1}{T} t \right], \quad t \in T, \quad (10)$$

где kwn_q^i – q -й числовой символ i -го кодового слова после проведения преобразований (8) и (9). Для примера полный набор отрезков комплексных огибающих сигналов для ПСК ЛКГ при $n=4$ и

$P=1$ всех кодовых слов представлен на рис. 4. Внешний вид сигналов ПСК напоминает отрезки реализаций случайных процессов.

При увеличении длины блока n спектр огибающих будет расширяться, а их вид приближаться к виду теплового шума. Таким образом, физическое представление переносчиков псевдослучайных кодов в канале соответствует описанию, приведенному в [1] и представлено на рис. 5, например, для случая длины блока $n=16$.

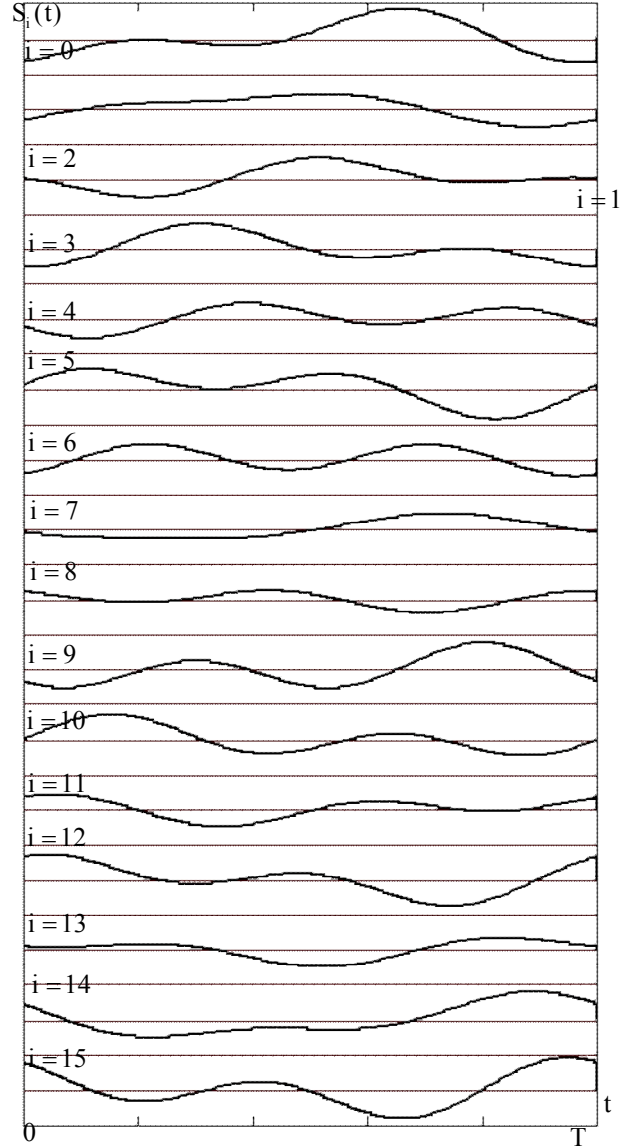


Рис. 4. Комплексные огибающие сигналов кодовой книги ПСК ЛКГ при $n=4$.

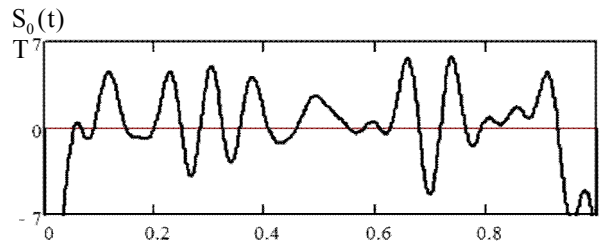


Рис. 5. Комплексная огибающая сигнала $S_0(t)$ при $n=16$

Для выделения символов кодовых слов из сигнала необходимо применить оптимальный когерентный прием[5]. Данный метод позволяет восстанавливать посимвольно числа кодовых слов ПСК на основе вычисления корреляционного интеграла:

$$\widetilde{kw}_q^i = \frac{2}{T} \int_0^T kw_q^i(t) \cdot Q_q(t) dt, \quad (11)$$

где

$$Q_q(t) = \begin{cases} \cos\left(2\pi \frac{q+1}{T} t\right), & \text{при } q = \left[0, \dots, \frac{n}{2}-1\right]; \\ \sin\left(2\pi \frac{q+1}{T} t\right), & \text{при } q = \left[\frac{n}{2}, \dots, n-1\right]. \end{cases} \quad (12)$$

Здесь $Q_q(t)$ – q -я квадратурная компонента в спектре сигнала. Набор квадратур фиксирован для всех кодовых слов ПСК.

Восстановленные по формуле (11) символы кодовых слов ПСК необходимо подвергнуть операциям, обратным (9) и (8), соответственно. При этом происходит восстановление истинного диапазона чисел ПСП ЛКГ. Стоит отметить, что в случае отсутствия помех в канале, символы кодового слова $\widetilde{kw}_q^i$, получаемые на основе (11), (12), полностью совпадают с исходными $\widetilde{kw}_q^i \equiv kw_q^i$. В случае присутствия помех в канале идеального восстановления не происходит. При этом, естественно, кодовые слова могут в значительной мере искажаться. Даже небольшая помеха, может вызвать искажение всех оценок символов кодового слова. Это приводит к невозможности применения алгебраических методов декодирования, обладающих ограниченными (конечными) значениями, так называемых, обнаруживающих и исправляющих способностей.

Выводы

Основным результатом данной статьи является обоснование перспективности практического при-

менения ПСК в современных системах передачи информации. Получена физическая модель процесса построения и посимвольной обработки кодовых слов. Обоснован выбор метода генерации ПСК на основе ЛКГ, поскольку способ ЛКГ обладает наименьшим удельным числом нелинейных операций. Проведена сравнительная оценка спектра взаимных расстояний, кодов ЛКГ и LFSR, которая показала практическую близость оценок спектральных свойств. Получены комплексные огибающие сигналов, ПСКЛКГ, которые описывают физическую модель представления кодов в многомерном евклидовом пространстве.

Данная модель позволяет предъявить требования к реализации модемов ПСК и сформулировать математическую задачу реализации вычислительно эффективных методов декодирования.

Список литературы

1. Шеннон К. Работы по теории информации и кибернетике / К. Шеннон. – М.: ИЛ, 1963. – 830 с.
2. Рассомахин С.Г. Оценка эффективности псевдослучайных кодов, сгенерированных с помощью LFSR / С.Г. Рассомахин, Т.В. Лавровская, О.И. Вотяков // Прикладная радиоэлектроника. – 2016. – Том. 14. – Вып. 4. – С. 339.
3. Алиев Т.И. Основы моделирования дискретных систем / Т.И. Алиев. – СПб.: СПбГУ ИТМО, 2009. – 363 с.
4. Рассомахин С.Г. Анализ применения правила простого округления для получения вычислительно реализуемых методов декодирования / С.Г. Рассомахин, Т.В. Лавровская // Системи обробки інформації. – 2015. – Вип. 5 (151). – С. 115-117.
5. Долгов В.І. Основи статичної теорії прийому дискретних сигналів: Підручник для студентів вищих навчальних закладів денної та заочної форми навчання / В.І. Долгов. – Х.: Форт, 2013. – 520 с.

Поступила в редколлегию 23.06.2016

Рецензент: д-р техн. наук, проф. В.А. Краснобаев, Харьковский национальный университет им. В.Н. Каразина, Харьков.

ФІЗИЧНА МОДЕЛЬ ПСЕВДОВИПАДКОВИХ КОДІВ В БАГАТОМІРНОМУ ЕВКЛІДОВОМУ ПРОСТОРІ

Т.В. Лавровська, С.Г. Рассомахін

В статті розглянуто основні проблеми практичного використання псевдовипадкових кодів в сучасних системах передачі інформації. Проаналізовані існуючі методи генерації псевдовипадкових послідовностей з урахуванням практичних особливостей їх застосування. Розроблена фізична модель псевдовипадкових кодів в евклидовому просторі, котра представляє переносники кодових слів у вигляді відрізків псевдовипадкових процесів з обмеженим спектром.

Ключові слова: псевдовипадкові коди, лінійний конгруентний генератор, регістр зрушення з лінійним зворотним зв'язком, евклидовий простір коду, фізична модель сигналів.

PHYSICAL MODEL OF PSEUDO-RANDOM CODES IN THE MANY-DIMENSIONAL EUCLIDEAN SPACE

T.V. Lavrovska, S.G. Rassomakhin

In the article are considered the main problems of practical use of pseudo-random codes in modern communication systems. There were analyzed the existing methods of generating pseudorandom sequences taking into account the characteristics of their practical application. There was developed physical model of the pseudo-random code in Euclidean space, which consist of code words vectors in the form of segments of implementations of pseudorandom processes with limited range.

Keywords: pseudocausal codes, linear congruous generator, shift register with a linear feed-back, Euclidean space of code, physical model of signals.