УДК 621.37:621.391

С.Г. Рассомахин, Т.В. Лавровская

Харьковский национальный университет имени В.Н. Каразина, Харьков

МАТЕМАТИЧЕСКИЕ МОДЕЛИ СЛУЧАЙНЫХ И ПСЕВДОСЛУЧАЙНЫХ КОДОВ

Проведен анализ причин кризиса помехоустойчивого кодирования. Обоснована перспективность применения псевдослучайных кодов и разработана их математическая модель. Рассмотрены сравнительные вероятностные характеристики равномерных и нормальных случайных кодов. Проведена оценка свойств простейших равномерных псевдослучайных кодов, полученных методом линейной конгруэнтной генерации. Предложены рекомендации по выбору параметров линейных конгруэнтных генераторов кодовыхсимволов.

Ключевые слова: *случайное кодирование, псевдослучайные коды, евклидово пространство кода, линейная конгруэнтная генерация, статистические модели процесса декодирования.*

Введение

Постановка проблемы. В теории систем передачи информации известна историческая роль методологии, основанной на использовании случайно выбираемых кодов, в доказательстве фундаментальных теорем для зашумленных каналов [1, 2]. Однако, доказательства на основе случайного выбора кода обычно называются неконструктивными, поскольку до сих пор ни одна попытка разработки конкретных кодов и методов их обработки не увенчалась успехом.

Основная причина этого заключается в отсутствии вычислительно реализуемых методов неслучайного построения псевдослучайных кодов, а также методов их декодирования.

Постановки проблем и постулирование важнейших положений теории информации дали толчок для поиска решения задач с использованием исключительно детерминированных (неслучайных) сигналов и алгебраических методов канального кодирования[3]. В этой области достигнуты серьезные результаты, имеющие важное теоретическое значение и дающие живые приложения абстрактным разделам дискретной математики. Математический прогресс привел к кризису практики – объем фундаментальных исследований явно превышает требуемые прикладные потребности и является несопоставимым с недостаточным, на наш взгляд, приращением показателей удельной эффективности систем передачи информации (СПИ). Основной причиной этого является почти эквивалентный обмен приращения энергетической эффективности на проигрыш в частотной эффективности, характерный для систем с комбинаторным алгебраическим кодированием. Перспективным направлением развития теории построения кодов следует считать исследование методов псевдослучайного кодирования (модуляции), которые приближают статистические характеристики канальных сигнально-кодовых конструкций к характеристикам реализаций шумовых последовательностей.

Анализ существующих результатов. Основным результатом, который дает объективные предпосылки для развития технологий случайного кодирования, является трактовка Шеннона физической сущности пропускной способности непрерывного канала C при ограничении средней мощности передатчика P. Процесс приближения к пропускной способности в этих условиях описывается следующим образом [1]. «… Пусть созданы $M = 2^k$ выборок белого шума, каждая длительности T. Им приписываются двоичные числа от 0 до $(M-1)$. В передатчике последовательности сообщений разбиваются на группы по k двоичных знаков и для каждой группы в качестве сигнала передается соответствующая выборка шума. Приемнику эти M выборок известны, и принятый искаженный шумом сигнал сравнивается с каждой из них. Выборка, которая имеет наименьшее среднеквадратичное расстояние от принятого зашумленного сигнала, принимается за переданный сигнал, по которому восстанавливается соответствующее двоичное число. Этот прием эквивалентен выбору наиболее вероятного (апостериори) сигнала.

Число используемых выборок шума M будет зависеть от допустимой частоты ошибок p, но для почти всех наборов выборок имеем

$$\lim_{p\to 0}\lim_{T\to\infty}\frac{\log M(p,T)}{T} = F\log\frac{P+N}{N}, \qquad (1)$$

где F – полоса частот канала, N – спектральная плотность мощности аддитивного шума.

Таким образом, независимо от того, насколько малым выбрано p, можно, выбирая T достаточно большим, приблизиться сколь угодно близко к пере-

даче $TF\log\frac{P+N}{N}$ двоичных единиц за время T …». Фактически, данное описание соответствует процессу декодирования случайно выбранного кода по правилу максимального правдоподобия [3], а пропускная способность – это предельно достижимая скорость передачи информации при помощи использования лучшего кода.

Следовательно, лучшим является код, полученный как математическое ожидание по ансамблю случайно выбираемых кодов.

В работах [4] и [5] разработаны элементы теории, способной придать идеям случайного кодирования определенный конструктивизм за счет применения эвристических методов коррекции свойств кодовых книг случайных кодов. Однако, этого недостаточно для разработки робастных алгоритмов кодирования и, особенно, декодирования таких кодов. Поскольку вычислительно реализуемыми алгоритмами декодирования (со сложностью не выше полиномиальной от длины блока) могут быть только алгоритмы декодирования кодов, построенных с использованием детерминированных методов генерации кодовых слов, речь может идти исключительно о «псевдослучайном» кодировании. При определенных условиях такие коды могут приближаться по своим вероятностным характеристикам декодирования к случайным кодам, допуская при этом возможность разработки алгебраических алгоритмов декодирования.

Основными причинами, сдерживающими практическую реализацию псевдослучайных кодов (ПСК), являются: во-первых, отсутствие формализованного математического аппарата получения кодов с хорошими корректирующими свойствами, а, во-вторых, – отсутствие практических способов не переборного декодирования. В работе [6] обоснована перспективность простейшего алгебраического метода генерации равномерных кодов, основанного на методе линейной конгруэнтной генерации, однако не получены рекомендации по выбору параметров генераторов.

Кроме того, рассмотренный в работе [6] метод декодирования на основе решения задачи целочисленного линейного программирования по методу отсекающих плоскостей Гомори, оказался нестойким к зацикливанию, что явилось причиной отсутствия возможности гарантированного получения решения.

Указанные причины не позволяют применять метод для практически требуемых значений длин блоков кодовых слов.

Целью статьи является разработка математической модели случайных и псевдослучайных кодов для проведения анализа и обоснования требований к методу получения кодовых книг, а также к числовым параметрам линейной конгруэнтной генерации. Сопутствующей целью работы, кроме этого, является разработка статистических моделей для оценки вероятностных свойств случайных и псевдослучайных кодов.

Основная часть

1. Математическая модель и пространственная структура случайных кодов

Процедура получения кодовой книги произвольного случайного кода может быть представлена следующим образом. Пусть кодированию подлежит последовательность двоичных символов, разбитая на блоки длиной k бит. Каждому из блоков может быть сопоставлено целое число, обозначающее условный лексикографический номер комбинации двоичных символов.

Таким образом, все возможные сообщения источника оказываются пронумерованными числами i из диапазона от 0 до $M = 2^k$.

Кодовая книга случайного кода Kb формируется, как набор векторов:

$$Kb = \left\{\overrightarrow{kb_i}\right\},\ i \in \left[0,(M-1)\right];$$
$$\overrightarrow{kb_i} = \left\{\underbrace{x_0^{\,i}, x_1^{\,i}, \ldots, x_{n-1}^{\,i}}_{n}\right\}, \qquad (2)$$

где n – длина блока кода;

$x_k^{\,i}$, $\left\{i \in \left[0,(M-1)\right]; k \in \left[0,(n-1)\right]\right\}$ – независимые, одинаково распределенные случайные величины с функцией плотности распределения вероятностей $f(x)$ и нулевым средним $m_x = 0$.

Обозначим $R = \frac{k}{n}$ – скорость кода (которая может быть как больше, так и меньше единицы). Дисперсия $\sigma_x^{\,2}$ распределения $f(x)$ при кодировании информативных параметров сигналов в канале определяется выражением:

$$\sigma_x^{\,2} = \alpha \cdot R \cdot P, \qquad (3)$$

где P – бюджет мощности, выделенный на передачу одного бита сообщения источника;

α – коэффициент, зависящий от вида модуляции в канале.

Выбор распределения $f(x)$ определяет тип укладки точек кодовых слов случайного кода в n-мерном евклидовом пространстве. Если функция $f(x)$ непрерывна в некотором диапазоне x, то укладка объемная; если $f(x)$ отлична от нуля только на некотором конечном множестве значений x, то укладка поверхностная. Поскольку поверхностное

расположение точек в многомерном пространстве является частным (вырожденным) случаем объемной укладки, то для получения лучших характеристик кодовых книг следует использовать непрерывные функции распределения $f(x)$.

Ключевым вопросом построения случайного кода является выбор геометрии подпространства кода в n-мерном пространстве. Эта геометрия полностью определяется видом непрерывного распределения $f(x)$.

Имеет смысл, без какого-либо снижения общности, рассмотреть всего лишь два варианта: нормальное (гауссово) и равномерное в заданном диапазоне распределения символов кодовых слов x_k^i.

Нормальный случайный код. Обозначим кодовую книгу (набор из M векторов) нормального кода, как $Kbn = \{\overrightarrow{kbn_i}\}$.

Символы кодовых слов векторов $\overrightarrow{kbn_i}$ независимы и распределены по закону:

$$f_n(x) = \left(2\pi\sigma_x^2\right)^{-\frac{1}{2}} \exp\left(\frac{x^2}{2\sigma_x^2}\right). \qquad (4)$$

Поскольку при построении модели для статистических испытаний и использовании программных датчиков случайных чисел средняя по реализации ансамбля кодовых слов мощность (дисперсия) символов может оказаться отличной от требуемого значения, определяемого выражением (3), имеет смысл проведение нормировки кодовой книги:

$$\overline{Kbn} = Kbn \cdot \left\{\frac{\alpha RP}{M}\sum_{m=0}^{M-1}\frac{\left|\overrightarrow{kbn_m}\right|^2}{n}\right\}^{-\frac{1}{2}}. \qquad (5)$$

Выполнение (5) гарантирует фиксированную среднюю мощность (3), приходящуюся на один символ каждого из M равновероятных кодовых слов нормального случайно выбранного кода. Использование гауссова распределения (4) при $n \to \infty$ обеспечивает формирование кода объемно-сферической укладки: кодовые точки асимптотически располагаются внутри гиперсферы с центром в начале координат и радиусом $r = \sqrt{n\sigma_x^2}$.

На рис. 1 приведен вид сечения подпространства нормального случайного кода с кодовой книгой, определяемой (4), (5) произвольной плоскостью, проходящей через центр гиперсферы. При этом приняты значения

$$P = \alpha = R = 1,\ n = 12. \qquad (6)$$

Ввиду особенностей геометрии подпространства кодовой книги нормальный случайный код можно называть *гиперсферическим кодом*.

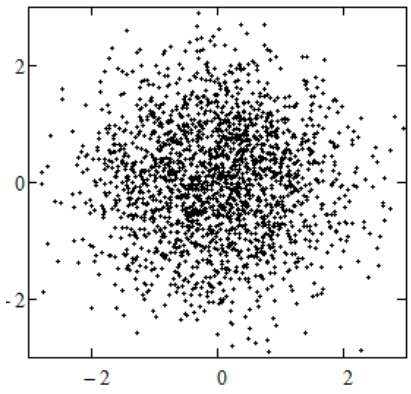

Рис. 1. Плоскостное сечение подпространства нормального случайного кода

Равномерный случайный код. Используем обозначение $Kbu = \left\{\overrightarrow{kbu_i}\right\}$ для кодовой книги кода с равномерным распределением символов слов. Функция плотности распределения вероятностей символов кодовых слов имеет вид:

$$f_u(x) = \begin{cases} \dfrac{1}{2\sqrt{3\sigma_x^2}}, & \text{при } |x| \le \sqrt{3\sigma_x^2}; \\ 0 \quad \text{в остальных случаях.} \end{cases} \qquad (7)$$

Условие нормировки случайно выбранного кода имеет вид:

$$\overline{Kbu} = Kbu \cdot \left\{\frac{\alpha RP}{M}\sum_{m=0}^{M-1}\frac{\left|\overrightarrow{kbu_m}\right|^2}{n}\right\}^{-\frac{1}{2}}. \qquad (8)$$

Точки кодовых слов равномерного случайного кода располагаются внутри n-мерного гиперкуба с размером ребра, равным

$$r = 2\sqrt{3\sigma_x^2}.$$

Сечение подпространства кода произвольной плоскостью, проходящей через пару координатных осей n мерного пространства при фиксированных условиях (6) показано на рис. 2.

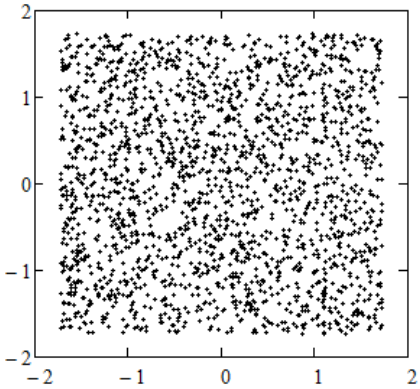

Рис. 2. Плоскостное сечение подпространства равномерного случайного кода

Расположение точек кодовых слов равномерного случайного кода делает возможным применение для него названия *гиперкубический код.*

2. Статистическая модель для исследования вероятностных свойств случайных кодов

Объективной характеристикой помехоустойчивости случайных кодов является величина вероятности ошибки при декодировании, приведенная к блоку кода при заданном отношении сигнал/шум. Ввиду объективной невозможности аналитической оценки вероятностных характеристик случайно выбираемых кодов, единственным способом определения предпочтения между гиперсферическим и гиперкубическим кодами является их статистическое исследование. Структура статистической модели для исследования вероятностных свойств случайных кодов в гауссовом канале приведена на рис. 3.

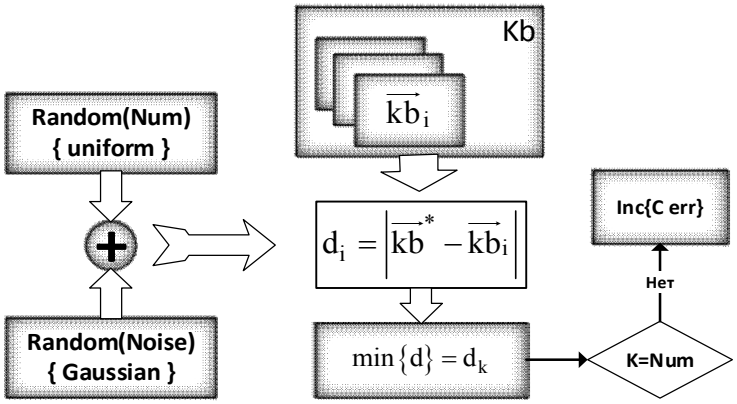

Рис. 3. Структура одного цикла статистической модели

На каждом цикле испытаний модели производится равномерно случайный выбор номера Num одного из M кодовых слов $\overrightarrow{kbn}_{Num}$ или $\overrightarrow{kbu}_{Num}$, вектор которого складывается с вектором гауссовой помехи $Random\{Noise\}$, обладающей мощностью, вычисляемой при заданном отношении сигнал/шум h. Затем, методом последовательного перебора вычисляются элементы массива взаимных расстояний $d_i,\ i \in [0,(M-1)]$ между искаженным кодовым словом и всеми словами кодовой книги. За истинный номер принимается номер кодового слова, ближайшего к анализируемому и соответствующего минимальному расстоянию d_k. В случае, если найденный номер ближайшего слова не совпадает с величиной Num, производится инкрементация счетчика ошибок $Cerr$. Реализуемый алгоритм соответствует декодированию по правилу максимального правдоподобия. Количество циклов испытаний подбирается, исходя из требуемой точности оценки и величины задаваемого отношения сигнал/шум. После проведения серии из K испытаний вероятность ошибки вычисляется отношением $p = Cerr/K$.

На рис. 4 представлены результаты статистического моделирования гиперсферического и гиперкубического кодов в равных энергетических условиях при трех различных отношениях сигнал/шум $h = 2,\ 3,\ 4$. Вероятности декодирования с ошибкой p представлены, как функции длины блока случайных кодов n при фиксированной скорости кодирования $R = 1$. На рис. 4 кривые, отмеченные как Hs, соответствуют гиперсферическим кодам, а кривые Hc – гиперкубическим.

Анализ результатов статистических испытаний позволяет сделать заключение о несомненном преимуществе гиперкубических случайных кодов, получаемых при равномерном распределении символов кода в пределах заданного числового диапазона (7).

3. Математическая модель и пространственная структура псевдослучайных кодов линейной конгруэнтной генерации

Поскольку равномерное распределение символов обеспечивает лучшие характеристики случайных кодов, то для получения псевдослучайных кодовых книг целесообразно использовать детерминированные алгоритмы, обеспечивающие аппроксимацию (7), (8). Наиболее простым и распространенным алгоритмом генерации равномерно распределенных псевдослучайных последовательностей (ПСП) является использование линейного конгруэнтного генератора (ЛКГ) [6]. При формировании канального кода получаемые числа, равномерно распределенные в заданном диапазоне, отождествляются с некоторыми значениями одного из информативных параметров сигнала (амплитудой, частотой или фазой). Подходящими переносчиками для построения числовых кодов являются амплитудно-фазовые, частотные, амплитудно-частотные и времяимпульсные методы модуляции, допускающие многоуровневую шкалу градации одного или нескольких информативных параметров.

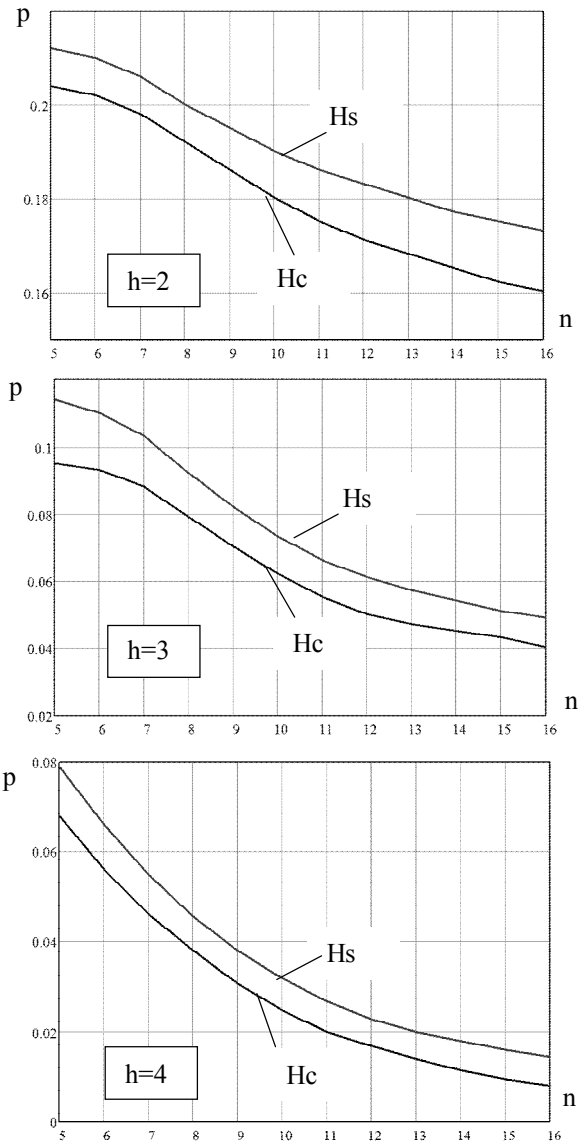

Рис. 4. Результаты статистических исследований случайных кодов

Представление физической модели случайных кодов в канале не является предметом данной статьи, поэтому в дальнейшем рассматриваться не будет.

В соответствии со свойствами линейных конгруэнтных последовательностей[6] элементы векторов кодовых слов $\overrightarrow{\mathrm{kb}_i} = \left\{ x_0{}^i, x_1{}^i, \ldots x_{n-1}{}^i \right\}$ имеют следующие значения:

- $x_0{}^i = i \in \left[0 \ldots (M-1) \right]$ – число, определяющее порядковый номер i блокадвоичных символов источника и являющееся порождающим числом ПСП;

$$x_k{}^i = \mathrm{mod}\left[a x_{k-1}{}^i + b, m \right], k \in 1 \ldots n-1 \qquad (9)$$

числа ПСП, порождаемые $x_0{}^i$ порекуррентному алгоритму ЛКГ;

- a, b, m – целые положительные константы, удовлетворяющие условиям: $m \geq M$, b и m – взаимно простые числа, величина $(a-1)$ кратна любому простому числу, которое меньше m и является его делителем;
- $(a-1)$ кратно четырем, если m кратно четырем.

Очевидно, что при выполнении данных условий произвольное k-тое число $i-$того кодового вектора связано с порождающим числом ПСП $x_0{}^i$ зависимостью:

$$x_k{}^i = \mathrm{mod}\left[a^k x_0{}^i + \frac{a^k - 1}{a-1} b, m \right], \qquad (10)$$
$$k \in 1 \ldots n-1.$$

Таким образом, для получения любого кодового слова ПСК ЛКГ достаточно задание номера i и осуществления $(n-1)$ рекуррентных вычислений по правилу (9) или (10). Это, в отличие от случайных кодов, избавляет от необходимости хранения в памяти передатчика и приемника полных кодовых книг. Каждая итерация (9), (10) содержит только одну нелинейную математическую операцию вычисления по модулю m. Это потенциально облегчает нахождение простого математического алгоритма не переборного декодирования, при этом процесс декодирования, по сути, может быть охарактеризован, как *факторизация* ПСП.Детерминированность алгоритма генерации псевдослучайных кодовых слов является причиной регулярной пространственной структуры кодовых книг ПСК ЛКГ. Это является естественным, поскольку получаемый код относится к классу линейных. Для примера, на рис. 5 представлены сечения пространства ПСК ЛКГ девятью перпендикулярными плоскостями при

$$k = 10,\ n = 10,\ R = 1,\ a = 5,\ b = 19,\ m = 2^k.$$

Геометрия кодовой книги ПСК ЛКГ демонстрирует неравномерность распределения взаимных расстояний кодовых точек в различных плоскостях. Несмотря на это корректно отнести ПСК ЛКГ к классу гиперкубических кодов.

Для выравнивания энергетических условий с условиями, в которых рассматривались случайные коды Hs и Hc, после первичного формирования по правилам (9), (10), кодовая книга ПСК ЛКГ $\mathrm{KbL} = \left\{ \overrightarrow{\mathrm{kbL}} \right\}$, перед получением сечений, показанных на рис. 5, подвергнута центрированию и нормировке:

$$\overline{\mathrm{KbL}} = \left(\mathrm{KbL} - \frac{M-1}{2} \right) \cdot \left(\frac{M^2 - 1}{12} \right)^{-\frac{1}{2}}. \qquad (11).$$

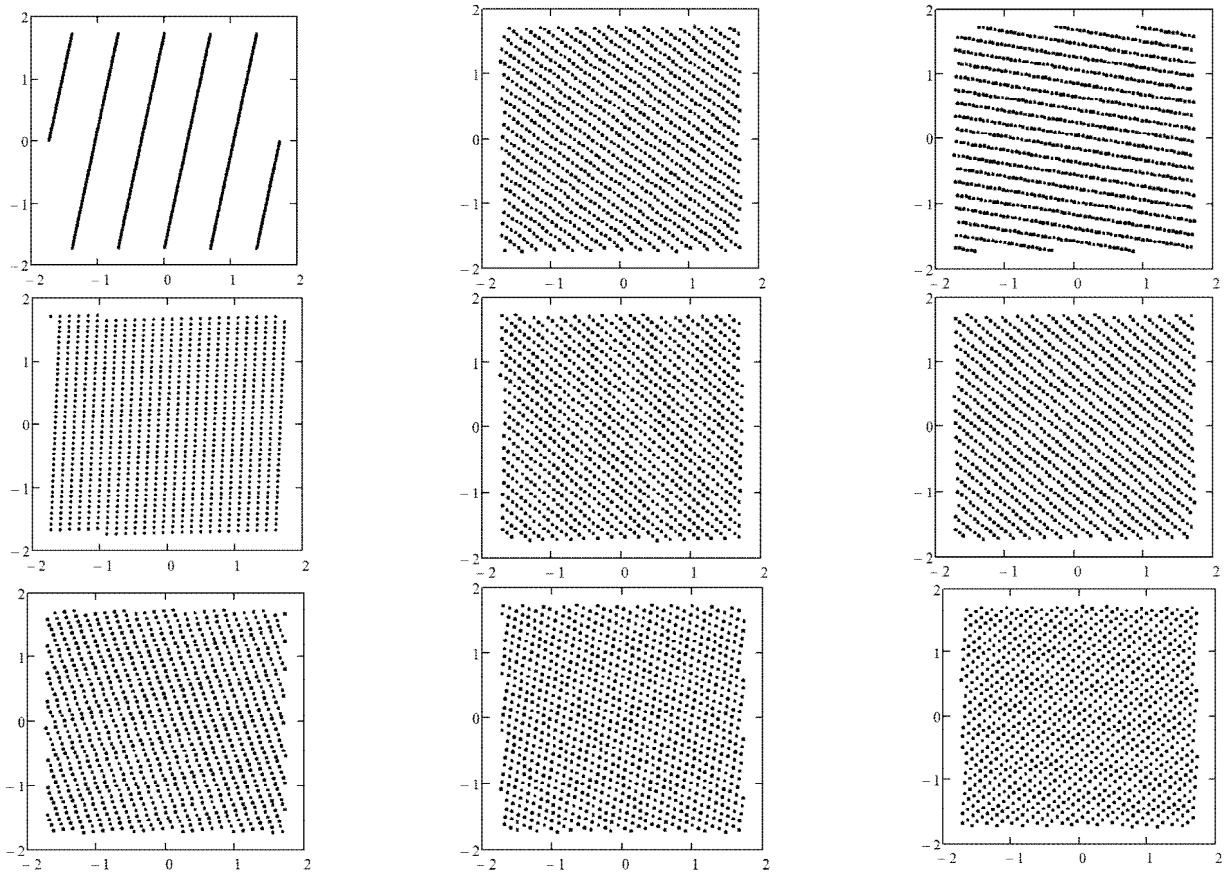

Рис. 5. Сечения пространства ПСК линейной конгруэнтной генерации

4. Статистическое исследование ПСК ЛКГ

Представляет интерес сравнение вероятностных свойств псевдослучайного кода ЛКГ с полученными ранее свойствами равномерно случайных гиперкубических кодов Hc .

Для статистической оценки вероятности декодирования с ошибкой использовалось правило максимального правдоподобия и алгоритм, показанный на рис. 3.

Отсутствие свойства равномерно случайного распределения числовых символов кодовых слов ПСК ЛКГ, которое является следствием использования детерминированного способа генерации этих символов, проявляется в скачкообразном изменении остаточной вероятности декодирования с ошибкой p при изменении длины блока кода n .

Результаты вычислительного эксперимента по исследованию вероятностных свойств ПСК ЛКГ для трех значений отношения сигнал/шум в гауссовом канале представлены на рис. 6.

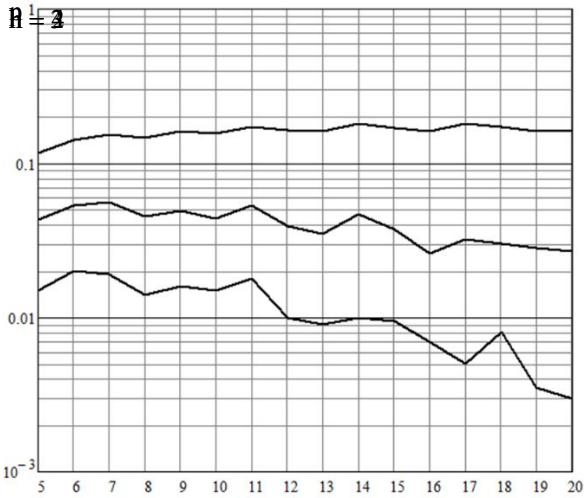

Рис. 6. Результаты статистических исследований ПСК ЛКГ

Анализ полученных на рис. 6 зависимостей позволяет выявить несколько особенностей псевдослучайных кодов по сравнению со случайными кодами, характеристики которых показаны выше на рис. 4. Данные особенности можно описать следующим образом:

– функции $p(n)$ не являются монотонными;

– при сохранении порядка величины p на уровне, соответствующем гиперкубическим случайным кодам Hc , наблюдается некоторое замедление снижения средней вероятности ошибки с ростом длины блока n . Это означает, что для достижения требуемой величины вероятности ошибки при заданном отношении сигнал/шум при использовании ПСК ЛКГ в условиях фиксированной скорости R необходимо использовать коды с большей длиной блока n , чем при использовании кодов Hc .

Кроме того, статистическое исследование ПСК ЛКГ позволило сформулировать два дополнительных требования к выбору параметров конгруэнтной генерации.

Во-первых, наилучшее (в среднестатистическом смысле) значением аддитивной константы b рекуррентного алгоритма (9) является следующее правило:

$$b = 2^{n-1} \pm \beta, \qquad (12)$$

где $\beta = 1,\ 3,\ 5, \ldots$ – нечетное.

Во вторых, недопустимые значения мультипликативной константы a в (9) описываются выражением:

$$a = 2^{Q} + 1, \text{ при } Q > 3. \qquad (13)$$

Полученные выражения (12) и (13) дополняют набор правил по выбору параметров стандартной конгруэнтной генерации в приложениях построения псевдослучайных кодов.

Выводы

Основной результат данной статьи заключается в получении универсальных математических моделей случайных кодов гиперсферической и гаперкубической укладки, а также модели псевдослучайного кода линейной конгруэнтной генерации.

Результаты статистического исследования полученных моделей позволили оценить вероятностные свойства кодов и сделать однозначное заключение о предпочтительности кодов гиперкубической укладки.

Использование детерминированного метода генерации кодовых слов ПСК на основе метода ЛКГ позволяет получить гиперкубические коды, практически совпадающие по своим свойствам со случайными кодами.

Еще одним аргументом в пользу выбора линейного конгруэнтного генератора для генерации ПСК является то, что данный метод обладает n нелинейными операциями на длину блока n и, как следствие, позволяетреализовать линейный алгоритм декодирования ПСК.

В ходе статистических исследований были получены дополнительные требования по выбору параметров для генерации ПСК с применением линейного конгруэнтного генератора.

Список литературы

1. Shannon C.E. A Mathematical Theory of Communication / C.E. Shannon // Bell Syst. Tech. J., Julay-Oct. 1948. – Vol. 27. – P. 379 – 423, 623 – 656.

2. Shannon C.E. Communication in the presence of noise / C.E. Shannon // Proc. IRE. – Jan. 1949. – Vol. 37. – P. 10 – 21.

3. Хэмминг Р.В. Теория кодирования и теория информации / Р.В. Хэмминг; пер. с англ. – М.: Радио и связь, 1983. – 176 с

4. Shulman N. Random Coding Techniques for Nonrandom Codes / N. Shulman // IEEE Trans. Inf. Theory, Sep. 1999. – Vol. 45, № 6. – P. 2101 – 2104.

5. Флейшман Б.С. Конструктивные методы оптимального кодирования для каналов с шумами / Б.С. Флейшман. – М.: Изд. АН СССР, 1963.– 224 с.

6.Рассомахин С.Г. Линейное целочисленное декодирование псевдослучайных кодов на основе метода отсечений Гомори / С.Г. Рассомахин // Системи обробки інформації. – Х.: ХУПС, 2011. – Вип. 5 (95). – С. 93 – 98.

Надійшла до редколегії 26.05.2016

Рецензент: д-р техн. наук, проф. В.А. Краснобаєв, Харківський національний університет ім. В.Н. Каразіна, Харків.

МАТЕМАТИЧНІ МОДЕЛІ ВИПАДКОВИХ ТА ПСЕВДОВИПАДКОВИХ КОДІВ

С.Г. Рассомахін. Т.В. Лавровська

Проведено аналіз причин кризи завадостійкого кодування. Обґрунтовано перспективність використання завадостійких кодів і розроблена їх математична модель. Розглянуті порівняльні імовірнісні характеристики рівномірних і нормальних випадкових кодів. Проведена оцінка властивостей найпростіших рівномірних псевдовипадкових кодів, отриманих методом лінійної конгруентної генерації. Запропоновані рекомендації по вибору параметрів лінійних конгруентних генераторів кодових слів.

Ключові слова: *випадкове кодування, псевдовипадкові коди, евклідів простір коду, лінійна конгруентністю генерація, статистичні моделі процесу декодування.*

MATHEMATICAL MODELS OF RANDOM AND PSEUDO-RANDOM CODE

S.G. Rassomakhin, T.V. Lavrovska

Analysis of the reasons of crisis of error-correcting coding. Grounded prospects for application of pseudorandom codes and developed their mathematical model. Considered comparative probabilistic characteristics of equal-length and normal casual codes. The evaluation of the properties of the simplest equal-length pseudo-random codes, which were received by method of the linear congruent generation. Proposed of recommendations on the choice of parameters of linear congruent generators code symbols.

Keywords: random *encryption, pseudo-random codes, Euclidean space code generation linear congruent, statistical models of the decoding process.*